Strategic Listening: Examining the Efficacy of a Pedagogical Cycle Intended to Teach the Listening Process, Increase Metacognitive Awareness and Improve Comprehension

A thesis presented to
the faculty of
the College of Arts and Sciences of Ohio University

In partial fulfillment
of the requirements for the degree
Master of Arts

Mary E. Freday

June 2011

© 2011 Mary E. Freday. All Rights Reserved.
This thesis titled
Strategic Listening: Examining the Efficacy of a Pedagogical Cycle Intended to Teach
the Listening Process, Increase Metacognitive Awareness and Improve Comprehension

by

MARY E. FREDAY

has been approved for
the Department of Linguistics
and the College of Arts and Sciences by

__________________________

Gregory Kessler
Assistant Professor of Linguistics

__________________________

Benjamin M. Ogles
Dean, College of Arts and Sciences
ABSTRACT

FREDAY, MARY E., M.A., June 2011, Linguistics

Strategic Listening: Examining the Efficacy of a Pedagogical Cycle Intended to Teach the Listening Process, Increase Metacognitive Awareness and Improve Comprehension

Director of Thesis: Gregory Kessler

Successful L2 listeners use metacognitive strategies more often and in more effective combinations than novice listeners. Raising learners’ awareness of metacognition and training them to self-select and use strategy clusters is hypothesized to improve L2 listening abilities. This study examined the effect of 10 weeks of explicit strategy instruction, delivered in a cyclical format to sensitize learners to the metacognitive processes underlying listening, on the listening comprehension, metacognitive awareness, and opinions of intermediate ESL learners. These constructs were examined with an alternate group for comparative purposes. The constructs of listening comprehension and metacognitive awareness were analyzed for the presence of a correlation; they were not correlated at the beginning of the treatment but were significantly correlated at the end. The results regarding listening comprehension were inconclusive and no significant improvement was made. The treatment resulted in a significant increase in the experimental participants’ self-reported metacognitive awareness. Half the experimental participants had positive opinions about the treatment.

Approved: ________________________________

Gregory Kessler

Assistant Professor of Linguistics
To Scott
For listening, for supporting, for encouraging, and for loving me

To Eve
For laughing, for singing, for cuddling, and for understanding
ACKNOWLEDGMENTS

I would like to thank Dr. Gregory Kessler, my committee chair, for his constant and unwavering support of and excitement about this project. Your door was always open, and you were always ready with creative suggestions, inspiration, and reassurance when I needed it.

I would also like to express my gratitude to Dr. Dawn Bikowski for her keen insight and sharp attention to detail. Your advice and constructive criticism have made this a better investigation and a more sophisticated piece of literature.

My sincere appreciation also goes to Dr. Scott Jarvis for his all of his assistance with the empirical aspects of this examination. Your knowledge, encouragement, and fascination with research have made it an honor to have worked with you.

I would also like to express my deepest thanks to Dr. Christopher Thompson and Dr. Gerald Krzic for their willingness to work with me and with each other to make this project possible. Without your advocacy and collaboration, this investigation would never have taken place.

Finally, I would like to thank all of the participants in this project. Your interest in these strategies, your intense focus during listening activities, and your insightful reflections kept me going and constantly reminded me of the underlying purpose of this entire investigation.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>3</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>5</td>
</tr>
<tr>
<td>List of Tables</td>
<td>9</td>
</tr>
<tr>
<td>List of Figures</td>
<td>10</td>
</tr>
<tr>
<td>Chapter 1: Introduction</td>
<td>11</td>
</tr>
<tr>
<td>Research Area</td>
<td>11</td>
</tr>
<tr>
<td>Purpose of this Study</td>
<td>17</td>
</tr>
<tr>
<td>Chapter 2: Literature Review</td>
<td>18</td>
</tr>
<tr>
<td>Historical Background</td>
<td>18</td>
</tr>
<tr>
<td>Learners’ Perceptions of L2 Listening</td>
<td>18</td>
</tr>
<tr>
<td>Importance of a Process-Based Approach</td>
<td>19</td>
</tr>
<tr>
<td>Instructional Models</td>
<td>21</td>
</tr>
<tr>
<td>Empirical Investigations</td>
<td>26</td>
</tr>
<tr>
<td>Conclusion</td>
<td>35</td>
</tr>
<tr>
<td>Chapter 3: Methodology</td>
<td>37</td>
</tr>
<tr>
<td>Research Questions</td>
<td>37</td>
</tr>
<tr>
<td>Participants</td>
<td>38</td>
</tr>
<tr>
<td>Selection</td>
<td>38</td>
</tr>
<tr>
<td>Placement</td>
<td>41</td>
</tr>
<tr>
<td>Instructors</td>
<td>42</td>
</tr>
<tr>
<td>Materials</td>
<td>43</td>
</tr>
<tr>
<td>Measuring Listening Comprehension</td>
<td>43</td>
</tr>
<tr>
<td>Measuring Metacognitive Awareness</td>
<td>43</td>
</tr>
<tr>
<td>Measuring Learners’ Perceptions of the Instruction They Received</td>
<td>45</td>
</tr>
<tr>
<td>Procedure</td>
<td>46</td>
</tr>
<tr>
<td>Experimental Group Treatment</td>
<td>46</td>
</tr>
<tr>
<td>Documentation of Comparison Group Treatment</td>
<td>53</td>
</tr>
<tr>
<td>Comparison Group Treatment</td>
<td>56</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Metacognitive, Cognitive, and Socio-Affective Listening Strategies</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>Participants’ Demographic Information and Experimental Status</td>
<td>39</td>
</tr>
<tr>
<td>3</td>
<td>Stages of Instruction and Related Metacognitive Processes and Strategies</td>
<td>48</td>
</tr>
<tr>
<td>4</td>
<td>Strategy Presentation Schedule</td>
<td>51</td>
</tr>
<tr>
<td>5</td>
<td>Experimental Participants’ Pre- and Post-Test MPT Listening Scores</td>
<td>78</td>
</tr>
<tr>
<td>6</td>
<td>Comparison Participants’ Pre- and Post-Test MPT Listening Scores</td>
<td>79</td>
</tr>
<tr>
<td>7</td>
<td>Results of Paired-Samples T-Tests on MPT Listening Pre- and Post-Test Mean Scores</td>
<td>82</td>
</tr>
<tr>
<td>8</td>
<td>Experimental Participants’ Pre- and Post-Test MALQ Total Scores</td>
<td>85</td>
</tr>
<tr>
<td>9</td>
<td>Comparison Participants’ Pre- and Post-Test MALQ Total Scores</td>
<td>86</td>
</tr>
<tr>
<td>10</td>
<td>Results of Independent-Samples T-Tests on the Experimental and Comparison Groups’ Pre-Test MALQ Category Mean Scores</td>
<td>92</td>
</tr>
<tr>
<td>11</td>
<td>Results of Independent-Samples T-Tests on the Experimental and Comparison Groups’ Post-Test MALQ Category Mean Scores</td>
<td>93</td>
</tr>
<tr>
<td>12</td>
<td>Results of Independent-Samples T-Tests on MALQ Category Gain Mean Scores of Both Groups</td>
<td>94</td>
</tr>
<tr>
<td>13</td>
<td>Participants’ Self-Reported Amounts of Independent Listening Practice</td>
<td>103</td>
</tr>
<tr>
<td>14</td>
<td>Participants’ Least Favorite Instructional Activities</td>
<td>123</td>
</tr>
<tr>
<td>15</td>
<td>Comparison and Contrast of <em>Fast Listening</em> and the Experimental Pedagogical Cycle</td>
<td>168</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>MPT listening pre-and post-test mean scores</td>
<td>80</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Distribution of MPT listening post-test scores</td>
<td>83</td>
</tr>
<tr>
<td>Figure 3</td>
<td>MALQ pre- and post-test total mean scores</td>
<td>87</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Distribution of MALQ post-test total scores</td>
<td>89</td>
</tr>
<tr>
<td>Figure 5</td>
<td>MALQ pre- and post-test category mean scores</td>
<td>90</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Scatter plot of all participants’ MPT listening and MALQ pre-test scores</td>
<td>96</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Scatter plot of all participants’ MPT listening and MALQ post-test scores</td>
<td>98</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Participants’ perceived levels of L2 listening improvement</td>
<td>101</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Elements participants perceived to have influenced their L2 listening</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>improvement</td>
<td></td>
</tr>
<tr>
<td>Figure 10</td>
<td>Self-reported reasons why listening instruction had a positive effect on</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>L2 listening improvement</td>
<td></td>
</tr>
<tr>
<td>Figure 11</td>
<td>Specific instructional components perceived to be the most beneficial</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>for improving L2 listening</td>
<td></td>
</tr>
<tr>
<td>Figure 12</td>
<td>Participants’ favorite instructional activities</td>
<td>118</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

Research Area

The present study seeks to examine the efficacy of a pedagogical cycle intended to increase learners’ metacognitive awareness of the listening process with the goal of improving their listening comprehension. Researchers in the field of second language acquisition (SLA) have acknowledged that learners’ awareness of their learning processes, cognitive functions, and strategy use can have positive outcomes on learning tasks (Bolitho et al., 2003; Chamot, 2004; Jessner, 2006; Wenden, 1987; Vandergrift, Goh, Mareschal, & Tafaghodtari, 2006). A pedagogical cycle capable of increasing metacognitive awareness of the listening process, then, is hypothesized to affect increases in listening comprehension abilities (Field, 2000; Goh, 2008; Vandergrift, 2002; 2004; Vandergrift & Tafaghodtari, 2010).

Instruction specifically aimed at increasing learners’ awareness of their metacognitive processes is based on an understanding of metacognition as a regulator, or manager, of cognition. It involves the ability to “[think] about one’s thinking” (Vandergrift et al., 2006, p. 433), or plan for, monitor and evaluate one’s thought processes. The term was originally coined by Flavell (1976), who defined it as “one’s knowledge concerning one’s own cognitive processes and products or anything related to them” (p. 232). Flavell (1979) highlighted the potential importance of metacognition in almost all aspects of education, and stated that learners with advanced levels of control over their metacognition would have better attention, problem-solving skills, be more autonomous, be more apt at language acquisition, and have improved reading and...
listening comprehension, among other advantages. He went so far as to state that future educators might want to train learners to “[increase] the quantity and quality of [their] metacognitive knowledge and monitoring skills” (p. 910) in order to improve their abilities to succeed in learning tasks.

Metacognitive awareness of language and second language acquisition (SLA) is known as metalinguistic awareness, which was succinctly defined by Jessner (2006) as having an understanding of language in the abstract. This understanding develops from an ability to reflect upon language, to manipulate it, to make plans for improving comprehension, and to monitor it in real time (Gombert, 1992). Throughout this document, the term metacognitive awareness will be used, with an understanding that it refers to metacognitive awareness of language, or metalinguistic awareness. Metacognitive awareness also involves the ability to employ metacognitive strategies in order to accomplish a learning objective. Metacognitive strategies function to oversee, manage, or control more the task-specific cognitive strategies. There are hundreds of cognitive strategies available to learners (such as inferencing, using mnemonic techniques, or repetition), but the metacognitive strategies fall into three primary categories, or processes: planning, monitoring, and evaluating (Vandergrift, 1997; Wenden, 1991). It should be noted that problem-solving has more recently come to be acknowledged as a metacognitive process (Chamot, Barnhardt, El-Dinary, & Robbins, 1999; Vandergrift et al., 2006), and there remains a good deal of debate in the field regarding the classification of learning strategies (Macaro, 2006). The categories provided by Vandergrift (1997) remain valuable, despite the debate, and will be accepted
as the standard for the purposes of this research. Vandergrift’s (1997) categories are presented in Table 1.

Table 1

*Metacognitive, Cognitive, and Socio-Affective Listening Strategies*

<table>
<thead>
<tr>
<th>METACOGNITIVE STRATEGIES</th>
<th>Metacognitive Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning</td>
<td>Advance Organization</td>
</tr>
<tr>
<td></td>
<td>Directed Attention</td>
</tr>
<tr>
<td></td>
<td>Selective Attention</td>
</tr>
<tr>
<td></td>
<td>Goal Setting</td>
</tr>
<tr>
<td>Monitoring</td>
<td>Comprehension Monitoring</td>
</tr>
<tr>
<td></td>
<td>Double-check Monitoring</td>
</tr>
<tr>
<td></td>
<td>Auditory Monitoring</td>
</tr>
<tr>
<td>Evaluating</td>
<td>Performance Evaluation</td>
</tr>
<tr>
<td></td>
<td>Strategy Evaluation</td>
</tr>
<tr>
<td></td>
<td>Problem-Identification</td>
</tr>
<tr>
<td></td>
<td>Goal checking</td>
</tr>
</tbody>
</table>

**COGNITIVE STRATEGIES**

*Involve the handling, organization, and use of new information*

<table>
<thead>
<tr>
<th>Cognitive Process</th>
<th>Cognitive Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inferencing</td>
<td>Linguistic Inferencing</td>
</tr>
<tr>
<td></td>
<td>Kinesic Inferencing</td>
</tr>
<tr>
<td></td>
<td>Extralinguistic Inferencing</td>
</tr>
<tr>
<td></td>
<td>Between-Parts Inferencing</td>
</tr>
<tr>
<td></td>
<td>Voice &amp; Paralinguistic Inferencing</td>
</tr>
</tbody>
</table>
Table 1 provides a list of those metacognitive strategies which relate to each primary metacognitive process, as well as a representative sampling of both cognitive and socio-affective strategies, which are those strategies involving interaction with peers or assisting learners in gaining control of the emotional aspects of language learning (Vandergrift, 1997). Successful learners have been found to utilize all three types of strategies in effective, task-dependent combinations (Chamot, 2004; Macaro, 2006; Murphy, 1985; O’Malley & Chamot, 1990; O’Malley, Chamot, & Küpper, 1989; Rubin, 1975; Vandergrift, 1997; 2003a), and together these three categories comprise the widely-accepted taxonomies of learner strategies developed by Oxford (1990), O’Malley and Chamot (1990), and Vandergrift (1997).
The relationship between metacognitive awareness, strategy use, and improved listening comprehension has only recently begun to be explored, but the investigation is built on both an understanding of the efficacy of instruction in metacognitive strategy use for improving reading comprehension (Chamot et al.,1999; Grabe, 2004; Jimenez & Gamez, 1998; Muniz-Swicegood, 1994; Schoonen, Hulstijn, & Bossers, 1998), and on a series of descriptive studies aimed at identifying listeners’ strategy use, which found that more successful listeners deploy metacognitive strategies more often and in more effective combinations with cognitive strategies than their less-successful counterparts (Chamot, 2004; Murphy, 1985; O’Malley, Chamot, & Küpper, 1989; O’Malley & Chamot, 1990; Rubin, 1975; Vandergrift, 1997; 2003a). Though these lines of research have informed investigations into strategy instruction for improving learner control of the listening process, they are not without their caveats.

First, while the link between metacognitive strategy use and improved reading comprehension has led researchers to hypothesize that the same might be true for the listening skill (Imhof, 2001; Mendelsohn, 1998; Thompson & Rubin, 1996), inherent differences between reading and listening demand that the hypothesis be empirically examined. The most salient of these differences relates to the speed of the input and the resultant demands on cognitive processing. Metacognitive strategy use also places demands on cognitive processing, and therefore runs the risk of over-taxing learners who are below a given linguistic threshold and diminishing listening performance (Cross, 2009; Imhof, 2001; Macaro, 2006; Vandergrift, 1997; Winne, 1995). While this linguistic threshold has yet to be definitively defined, and may be somewhat learner-
specific, Vandergrift (1997) found preliminary evidence that it might be as low as the level of the true-beginner or beginner.

Second, while the rich descriptive studies of the successful language learner were informative and enlightening, they were not able to address the pivotal hypothesis concerning whether or not less-successful learners are capable of being taught how to use the strategies of their more successful peers. The direction of the relationship (does increased strategy use make learners more successful, or does higher proficiency create the conditions necessary for increased strategy use?) needs to be examined in empirical, classroom-based intervention studies (Chamot et al., 1999; Graham & Macaro, 2008; Lai, 2009).

A number of researchers have conducted such empirical investigations (Carrier, 2003; Cross, 2009; Goh & Taib, 2006; Graham & Macaro, 2008; O’Malley, 1987; Thompson & Rubin, 1996; Vandergrift & Tafaghodtari, 2010). The results of these have been largely inconclusive, either due to flawed methodologies (O’Malley, 1987), measurable success only possible though the use of specially designed assessment instruments (Cross, 2009; Graham & Macaro, 2008; Thompson & Rubin), sample sizes too small to produce generalizable results (Carrier, 2003; Goh & Taib, 2006), or artificially controlled comparison groups (Vandergrift & Tafaghodtari, 2010). However, each of these investigations also provided some evidence that instruction of this nature has promise, and as a result of that promise, calls to continue to be made for more classroom based studies of this nature (Cross, 2009; Graham & Macaro, 2008; Vandergrift, 2003a; 2004; 2007; Vandergrift & Tafaghodtari, 2010).
Purpose of this Study

The present study, therefore, sought to answer those calls for an empirical investigation of a pedagogical model designed to increase learners’ metacognitive awareness and metacognitive strategy use when engaging in L2 listening practice. The model which was investigated provided repeated exposure to and practice with metacognitive strategies in order to raise learners’ awareness of the processes underlying successful listening comprehension, while simultaneously incorporating a select number of cognitive and socio-affective strategies. This cycle was based on one originally proposed by Field (2000), utilized by Vandergrift (2003b), and then elaborated on by Vandergrift (2004), but it was modified in order to incorporate research into the efficacy of increased opportunities for reflection (Goh, 1997; Graham & Macaro, 2008), the critical importance of an integration of bottom-up (word and phrase-level recognition) and top-down (activation of background knowledge and use of inference) processing practice (Buck, 2001; Carrier, 2003; Dunkel, 1991; Fender, 2003; Goh & Taib, 2006; Macaro, 2006; Vandergrift & Tafaghodtari, 2010), and the integration of cognitive, socio-affective and metacognitive strategy instruction (Brown & Palincsar, 1982; Graham & Macaro, 2008; Macaro, 2006; O’Malley, 1987; Purpura, 1997; Thompson & Rubin, 1996) delivered according to the guidelines specified by Chamot et al. (1999) and Wenden (1991). This cycle was the end-product of years of research aimed at identifying those elements most critical to the development of successful L2 listening comprehension abilities, the history of which will be discussed in the upcoming literature review.
CHAPTER 2: LITERATURE REVIEW

Historical Background

Listening has historically been the least researched of the four language skills (Dunkel, 1991; Morley, 1991; Vandergrift, 2007). This neglect may stem from its intangible (Morley, 1991) and “inside-the-head” (Dunkel, 1991, p. 438) nature.

However, listening plays a critical role in both the acquisition of a second language and the development of communicative competence in that language, as several models of SLA have shown (Dunkel, 1991; Feyten, 1991; Krashen, 1981). For example, Krashen’s (1981) “Input Hypothesis” (p. 9) posited that language learning is dependent upon the presence of what he termed “comprehensible input” (p. 9). This is especially relevant at the earliest stages of language acquisition, and the receptive skills are crucial for providing this input. This historic neglect led Dunkel (1991) to call for more empirical research into the processes underlying listening comprehension, with the aim of improving L2 listening instruction. Several researchers have heeded that call, and the last 20 years have seen a growing interest in and understanding of the listening process (Buck, 2001; Chamot, 2004; 2005; Field, 2000; Goh, 1997; 2002; 2008; Vandergrift, 1997; 2003a; 2003b; 2007).

Learners’ Perceptions of L2 Listening

Additionally, there has been a small body of research examining learners’ perceptions of L2 listening, which has yielded informative results regarding an apparent disconnect between the importance of listening and its inaccessibility in the eyes of learners. Arnold (2000) found that learners rated listening as the most difficult of the
four skills. She hypothesized that this was due to the inability of novice listeners to parse the rapid speech stream and comprehend the input. Moreover, she explained that increases in anxiety put more pressure on working memory resources, leave less energy available for attending to the input, and place the learner in an unproductive downward spiral of poor performance creating anxiety which in turn creates even poorer performance. Graham (2006) also found that listening can create anxiety in some learners. One-fourth of the 595 adolescent learners she surveyed rated listening as their least successful skill, and the 15% who rated it as their most successful area attributed this success to such factors as “luck” (p. 177) and innate ability. None of these learners perceived a connection between strategic listening and successful listening, and many expressed uncertainty regarding ways to improve their listening skills. Hasan (2000) surveyed 81 listeners and found that the majority rated listening tasks to be difficult. He hypothesized that they were not utilizing listening strategies in appropriate and efficacious ways. Vogely (1995) also observed that less successful listeners perceived themselves as less able to listen strategically. Therefore, instruction aimed at raising learners’ awareness of the processes underlying successful listening outcomes has been advocated by several researchers active in the field (Chamot et al., 1999; Goh, 2008; Vandergrift, 2004).

Importance of a Process-Based Approach

This gap between the importance of listening and learners’ impressions of its difficulty needs to be addressed, as does the lack of understanding on listeners’ part about how to gain control of the listening process. Goh (2008) promotes a process-based
approach to listening instruction, in order to “demystify the skills” (p. 192) involved in listening comprehension. While crediting the communicative language movement of the 1980s with improvements in L2 listening instruction (such as the widespread recognition of the importance of prediction in the pre-listening phase), she points out that many techniques popular for teaching listening in the 1960s have been resistant to change, such as the use of multiple-choice comprehension questions, due to wash-back from high-stakes assessments. “Comprehension-based techniques” (p. 189) such as these focus on the product of listening rather than the underlying processes and turn listening practice activities into miniature listening tests. Goh argues that teaching listening by testing listening has the potential to create anxiety, brings fear of negative evaluation into the listening classroom, and does little to equip learners with the tools necessary to gain control over the listening process and work towards improvement, whereas a process-based approach, such as the strategic, metacognitive models proposed by Wenden (1991), Chamot et al. (1999), and Vandergrift (2004), would bring increased control, confidence, and eventually proficiency to L2 listeners.

These three instructional models are not the only strategic, metacognitive models in existence, but they are the only ones that informed the present study. They share a focus on the importance of metacognition in learning and are all grounded in the theory that learners can be trained in strategy use in order to gain control of metacognition, but they also share important differences. Both the Wenden (1991) and the Chamot et al. (1999) models provide useful information regarding the most efficacious method of introducing learning strategies to learners, but the Chamot et al. and Vandergrift (2004)
models encourage the orchestration of multiple strategies simultaneously due to repeated exposure to the metacognitive processes underlying listening comprehension. Finally, only Vandergrift’s (2004) model focuses exclusively on the use of these techniques for the improvement of L2 listening abilities. Each of these models will be examined in more detail in the upcoming section.

Instructional Models

Wenden’s (1991) book *Learner Strategies for Learner Autonomy* provides instructors with a series of guidelines and action plans for implementing strategy instruction in the L2 classroom. Her primary objective is to equip teachers with the skills necessary to train their students to become strategic, autonomous learners. The concepts of autonomy and metacognitive awareness are thus inextricably linked, for if learners have awareness of and control over their metacognition, they will simultaneously be self-directed, autonomous learners. Her guidelines are firmly grounded in cognitive psychology and an understanding of the cognitive and metacognitive processes at work within the mind of the learner. Her guidelines are also based on the aforementioned descriptive studies of the successful language learner and Flavell’s (1979) model of metacognition, and they rest on the premise that novice or unsuccessful language learners can be taught how to use the strategies, or “mental steps or operations that learners use to learn a new language and to regulate their efforts to do so” (p. 18), of their more successful peers.

She suggests that each instructor select those cognitive strategies which are best suited to the particular skill, proficiency level, and group of learners to be taught, and this
way advocates an individualized rather than one-size-fits-all approach to strategy instruction. She refers the reader to influential taxonomies of learner strategies for assistance in selecting an appropriate group of cognitive strategies (O’Malley & Chamot, 1990; Rubin, 1989) but warns that the inclusion of too many strategies in too short a time is likely to be counterproductive. Metacognitive strategies, however, are intended to be incorporated throughout the intervention, due to their ability to help learners gain control over the entire learning process and their applicability across the entire range of potential learning tasks.

Wenden’s (1991) guidelines state that the following factors have the power to determine the effectiveness of a strategic intervention:

- the degree to which learners are informed of the purpose and value of the training
- the degree to which teacher scaffolding (support) allows them to first use strategies under guidance but ultimately leads them to autonomous self-selection and self-regulation of strategy use
- the degree to which the training is contextualized for both the skill and the specific group of learners in question
- the degree to which learners’ inherent beliefs about their roles and capabilities as language learners are addressed and revised

Chamot et al. (1999) also devised a metacognitive instructional model called the Cognitive Academic Language Learning Approach (CALLA). The CALLA model is a recursive, five-part framework based on four metacognitive processes: planning,
monitoring, evaluating, and problem-solving. The model leads the learner through each of these first four stages, but the authors make it clear that this process is not linear, and at any time the learner may have to loop back to a prior stage in order to refine and improve the comprehension process. The fifth stage, expansion, is an attempt to assist the learner in transferring the skills obtained in the first four stages to novel learning contexts.

Chamot et al. (1999) provide a taxonomy in which each metacognitive process is divided into lists of more behavior-specific metacognitive strategies which assist the learner in achieving the overarching metacognitive goal. For example, the metacognitive process of planning is subdivided into six attendant strategies such as “setting goals,” “directed attention,” “activating background knowledge,” and “predicting” (p. 18-19). For each strategy the authors also provide a clear definition, a concrete example, reasons for selecting this strategy, and potential contexts in which this strategy can be employed. In order for strategy use to be internalized and transferred to contexts beyond the immediate practice situation, learners must be informed of each of these elements, in addition to being provided with ample opportunities to practice each strategy in moderately challenging activities (Chamot et al., 1999; Wenden, 1991; Winne; 1995).

Vandergrift (2004) devised his own model of metacognitive instruction geared specifically at improving listening comprehension through a process-oriented approach to the skill. This model is based on a cycle initially proposed by Field (2000), to which Vandergrift made slight modifications. The most important of these is arguably Vandergrift’s suggestion that the cycle be implemented with every single listening
practice activity, rather than the occasional inclusion of this type of listening practice advocated by Field, in order to fully sensitize listeners to the metacognitive processes underlying listening comprehension. The cycle involves five stages through which the listener progresses linearly. The first is a planning or predicting stage, and guides the learner through the metacognitive process of planning and the related strategies of directed attention, goal setting, activating background knowledge, and predicting.

Learners are informed of the topic and selection type (interview, telephone conversation, etc) and must then predict what content they are likely to hear as well as what specific lexical items they might encounter. These predictions are written down, and learners are instructed to make plans for directed attention. Next comes the “first verification stage” (p. 11), in which the learners listen to the selection for the first time. During the listening, they are to check their predictions and note any other information they have understood. This stage is related to the metacognitive process of monitoring and the related strategies of selective attention and comprehension monitoring, or asking oneself if the input is being understood. The second half of this stage involves collaboration with a peer, in which comparisons are made, hypotheses are verified or rejected, differences are discussed, problems are identified, and plans are made for the subsequent listening.

This phase brings the learner back to the metacognitive process of planning, in order to prepare for the “second verification stage” (p. 11), during which the learners listen again to the selection to verify any differences noticed with their partners during the previous stage, make any necessary corrections to what they had written, and note additional information that they have understood. This is related to the metacognitive processes of
monitoring and problem-solving, and is followed by a whole group discussion in which learners work together to orally reconstruct the text and to reflect on strategies which were useful for solving particular listening problems. This part of the second verification stage involves the metacognitive processes of monitoring and evaluating. This discussion is followed by a “final verification stage” (p. 11) in which learners listen for important points which were determined as problem-spots by the entire group. This stage requires monitoring and incorporates the strategy of selective attention. The final stage of this model is called the “reflection stage” (p. 11) and involves the metacognitive process of evaluation. During this stage learners write personal reflections about their experience as listeners. Themes for this writing can include such topics as successful strategies employed as well as unsuccessful attempts at solving listening problems and personal goals for the next listening activity.

When Field (2000) initially proposed his version of this instructional model, he explained that pertinent cognitive strategies (such as listening for particular discourse markers or stressed words) could be interwoven into the cycle. However, he explained that cognitive strategies, when taught in isolation, have been shown to improve L2 listeners’ abilities with the specific listening tasks at hand, but have not been demonstrated to lead to their overall improvement as listeners. In other words, it is both the cyclical, repetitive nature of the cycle and the combination of instruction in behavior-specific cognitive strategies with regulatory metacognitive strategies which is hypothesized to lead to the development of strategic, successful listeners. Vandergrift (2004) proposes that listening instruction be delivered in a repetitive cycle such as this so
that the metacognitive strategies become familiar. He explains that repeated exposure to the same set of strategies can help learners to internalize those strategies, leading to autonomous learning and strategy transfer. When Vandergrift published this cycle he explained that there was as of yet little empirical evidence in support of such an approach and called for carefully designed and controlled classroom-based examinations into its effect on listening proficiency.

Empirical Investigations

There have been a few promising, though largely inconclusive, empirical investigations of such models for the teaching of listening comprehension. Due to the challenges inherent in classroom-based, action research (Dörnyei, 2007) each of these studies was limited by one or more methodological constraints. Additionally, it will be noticed that the majority of these investigations were only able to achieve statistically significant findings on researcher-developed alternative assessments designed to mirror instructional techniques, and either did not attempt or failed to do so on institutional, standardized listening assessments. Testing research has shown that task familiarity can have a considerable effect on students’ performance (Peña & Quinn, 1997), so there is a possibility that those investigations biased their experimental participants’ results by employing assessment types which mirrored instructional methodology and content. Of additional concern is the fact that today’s learners are expected to demonstrate their performance on high-stakes assessments while at the same time being expected to achieve a working level of communicative competence in the target language. The models reviewed in the previous section and investigated in the following studies are
hypothesized to improve learners’ abilities to comprehend real-time, authentic L2 input, but they must also improve learners’ abilities to succeed on multiple-choice format standardized assessments in order for the experimental methodologies to be incorporated into textbooks or to have an impact on the policies and procedures of academic institutions.

Listening research has shown that multiple-choice assessments have the potential to create enough anxiety in learners to negatively affect performance (Arnold, 2000) and that the act of problem-solving while comprehending input leads to such heavy demands on working memory that information held in working memory can be lost, negatively impacting performance (Yi’an, 1998). While multiple-choice assessments may not provide the most accurate assessment of a listener’s proficiency, standardized assessment practices are not likely to be altered in the near future. On the other hand, Purpura’s (1997) investigation into the relationship between cognitive and metacognitive strategy use and test-taking performance provides some evidence that instruction which empowers learners with control over their metacognitive and cognitive processes can lead to improved results on standardized assessments. He found metacognitive strategy use to have an indirect effect on test-taking performance, and cognitive strategy use to have a direct impact. This is due to metacognitive strategies’ ability to regulate cognitive strategies and allow learners greater control over those cognitive processes. Therefore, he recommended that strategic interventions combine both types of strategies, because they do not function in isolation but are utilized in complex and interwoven clusters, or groups.
One of the earliest strategic interventions aimed at examining the effect of cognitive and metacognitive strategy training on L2 listening was conducted by O’Malley (1987). Participants (n=75) were divided into a metacognitive group, which received training in one metacognitive, one cognitive, and one socio-affective strategy per listening activity, one cognitive group, for whom the metacognitive strategy was removed, and one comparison group, which did not receive strategy training. O’Malley’s results did not reach statistical significance, but the reasons underlying that fact were very informative for future researchers. First, it was hypothesized that the short treatment time (50 minutes per day for two weeks) limited participants’ abilities to practice and become proficient in the use of the presented strategies. Second, the learners in O’Malley’s study were not given the opportunity to reflect on or evaluate their strategy use. Finally, the strategies were not presented to learners in a cyclical format. The isolation of a single metacognitive strategy may not be as effective as instruction which empowers learners with control over all four of the metacognitive processes of planning, monitoring, evaluation, and problem-solving.

The first longitudinal investigation of the influence of process-based cognitive and metacognitive strategy training on listening proficiency was undertaken by Thompson and Rubin (1996). The experimental treatment was successful in showing that strategy training of this nature can improve listening comprehension on an alternative assessment, but was unable to achieve statistical significance on a standardized “Comprehensive Russian Proficiency Test” (p. 334).
The experimental group \((n=24)\) was instructed by one of the researchers for two years according to a process-based approach in which metacognitive strategies relating to planning, monitoring, and evaluating (such as goal setting, predicting, and selective attention) were interwoven with the cognitive strategies of listening for cognates and familiar words and grammatical structures, listening for redundancies, and listening to tone and intonation. The comparison group \((n=12)\) was exposed to the same listening selections and spent the same amount of time engaged in listening activities, but they did not receive strategy training. Additionally, the comparison group’s instruction was product-based, meaning that it was the content of the listening selections which was the basis for classroom discussions and writing activities, rather than the listening process itself. This important study marks the first empirical investigation in which an approach of this nature led to statistically significant results in end-of-treatment listening proficiency, though only on the alternative assessment, termed “the video test” (p. 336), which was designed by the researchers to mirror the method of instruction.

Another smaller-scale strategic intervention was undertaken by Carrier (2003). This study also showed that process-based, explicit strategy instruction can improve listening comprehension abilities, but it was beset by methodological deficiencies. The Carrier study was unique because at the time it was the only such study to examine the effects of such an approach on English as a second language (ESL) learners, whom the author argued have a much more pressing need to develop listening competency than do those learning foreign languages, due to their immersion in the target language and the related demands to comprehend input for day-to-day survival.
Carrier’s (2003) adolescent participants (n=7) were exposed to explicit strategy training for 30 minutes per lesson over the course of a 15 week semester. Explicit strategy training involves the naming of each strategy as well as inclusion of an explanation of its benefits, immediate applications, and potential applications beyond the immediate context as advocated by Wenden (1991) and Chamot et al. (1999). Instruction also included teacher think-alouds to model strategy use, daily reminders of those strategies which had already been presented, and a focus on both bottom-up and top-down auditory processing.

Every single student in the training showed improvement in their listening comprehension abilities on both top-down and bottom-up processing assessments. Additionally, their gains reached statistical significance according to the Wilcoxon Signed-Rank Test for small sample sizes. However, the generalizability of this study is limited not only by small sample size, but also by the fact that all participants were volunteers. They might have, therefore, had higher levels of motivation than a randomly sampled participant pool (Dörnyei, 2007). Also, the lack of a comparison group prevents the important step of making comparisons between groups, but the improvement of the participants does lend evidence to arguments in favor of an approach of this nature.

Goh and Taib (2006) engaged in a very similar, small-scale study (n=10) of the effects of metacognitive strategy instruction on elementary-aged EFL learners. Interestingly, the participants in this study were being instructed in strategies for the express purpose of improving their performance on a primary school exit exam, and so the both the instruction and the assessment included traditional “listen and answer” (p.
31 activities and excluded pre-listening activities in order to mirror a testing environment. As was the case in the Carrier (2003) investigation, every student improved his or her listening ability as a result of the treatment, but the weakest students exhibited the greatest gains, providing preliminary evidence that less-successful listeners have more to gain from such interventions than their more-successful peers, provided they are above the linguistic threshold necessary for comprehending the strategy training (Cross, 2009; Imhof, 2001; Macaro, 2006; Vandergrift, 1997; Winne, 1995).

Additionally, all of the students reported increased levels of confidence with L2 listening tasks, all believed that they had improved their L2 listening abilities, and all were observed to have improved control of their metacognitive processes and to have increased levels of metacognitive awareness. However, this study’s generalizability was limited due to small sample size, the absence of statistical tests, and the lack of a comparison group with which to compare the findings.

Another important longitudinal investigation was carried out by Graham and Macaro (2008). The researchers examined the effects of a strategic intervention on the listening comprehension of lower-intermediate level French learners, and examined teacher scaffolding (support) as an independent variable. As was the case in the Thompson and Rubin (1996) investigation, statistically significant results were only achieved on an alternative assessment. Though Graham and Macaro did not examine the effects of the treatment on a standardized assessment, they were careful to insure that the method of assessment employed was sufficiently different from those used to practice listening during the treatment to prevent bias on the part of the experimental participants.
Two experimental groups (n=68) were exposed to cognitive and metacognitive strategy training over the course of six months. The cognitive strategies included both bottom-up strategies (such as segmenting the speech stream) and top-down strategies (such as making predictions and inferences). These were embedded within a metacognitive framework of monitoring and evaluating in order to foster learner autonomy, which was hypothesized to lead to gains in long-term listening achievement. One of the experimental groups (n=29) was given a higher degree of teacher scaffolding (support) than the other, which came in the form of listening diaries in which the learners were asked to reflect on strategy use and in which they received teacher feedback on their reflections. This group also participated in activities designed to raise their awareness about learner autonomy. The other experimental group (n=39) received strategy training but did not use listening diaries or participate in awareness-raising activities. The comparison group (n=39) did not receive strategy training, did not participate in awareness-raising activities, and was not provided with opportunities for reflection.

The construct of listening comprehension was measured via free-recall listening tests, on which the participants restated everything that they could remember from a given listening passage. The experimental groups’ gains over the comparison group were statistically significant on both the immediate and the delayed listening post-tests, but the level of scaffolding only resulted in significant gains on the immediate post-test. The fact that the high-scaffolding group made statistically significant gains in listening comprehension over the low-scaffolding group on the immediate post-test provides
evidence that feedback on reflections can have a positive effect on learners’ progress in the short term.

Another interesting strategic intervention was undertaken by Cross (2009). In this study, both the experimental (n=7) and the comparison (n=8) groups were exposed to a repetitive instructional methodology which cycled through the metacognitive processes of planning, monitoring, and evaluating, but the experimental groups also received explicit cognitive strategy training in addition to listening practice. The construct of listening comprehension was measured with a special test designed to mirror the method of instruction, and no method of standardized assessment was utilized. Both groups made statistically significant gains on the post-test, but there was no significant difference between the two groups. Cross (2009) hypothesized that it may have been the metacognitive cycle itself which was the key variable influencing the learners’ success, rather than the explicit strategy training.

The most recent strategic intervention to be discussed is that of Vandergrift and Tafaghodtari (2010). The researchers utilized Vandergrift’s (2004) model of listening instruction and investigated not only the effect of this cycle on listening comprehension but also on the metacognitive awareness of high-beginner to low-intermediate level learners of French. Listening comprehension was measured with the University of Ottawa’s French as a second language (FSL) placement test, making this one of the few studies to measure the effects of a strategic listening intervention solely with a standardized assessment. Metacognitive awareness was measured with a retrospective
self-report instrument called the Metacognitive Awareness Listening Questionnaire (MALQ), designed by Vandergrift et al. (2006).

Following the findings of the Goh and Taib (2006) investigation, the researchers hypothesized that training of this nature would be more beneficial to less-skilled listeners, and sub-divided the experimental (n=59) and comparison (n=49) groups into more- and less-skilled listeners on the basis of the listening pre-test. The primary variable under consideration was the effect of formal, regularly scheduled opportunities for collaboration, reflection, and discussion about the metacognitive processes underlying listening comprehension. To that end, the experimental groups followed Vandergrift’s (2004) model in every listening lesson, whereas the comparison groups were not given opportunities to predict, to discuss comprehension or strategy use with a partner, or to reconstruct the text. Additionally, they were neither instructed in how to monitor their comprehension nor invited to reflect upon the listening process. They were, however, informed of the topic prior to listening and demonstrated understanding by writing down what they had heard.

Gains in metacognitive awareness, as measured by the MALQ, were inconclusive. However, on the final measure of listening comprehension, the experimental groups outperformed the comparison groups. While the gains of the more-skilled listeners in the experimental groups over those in the comparison groups failed to reach statistical significance, the less-skilled listeners in the experimental groups made statistically significant gains over the less-skilled listeners in the comparison groups. The gains of
the less-skilled listeners in the experimental groups mark the first time that treatment of this nature has reached statistical significance on a standardized listening assessment.

Conclusion

In conclusion, prior investigations have provided some evidence that strategy instruction has the ability to improve learners’ L2 listening comprehension, when delivered in a cyclical and repetitive format which allows learners ample exposure to and practice with the metacognitive processes of planning, monitoring, and evaluating comprehension. However, the vast majority of these studies failed to reach statistical significance on standardized listening assessments, and those variables with the most direct influence on increases in listening comprehension have yet to be unequivocally identified, though progress is being made. Moreover, classroom-based studies are prone to limits of generalizability, requiring that many replications be undertaken with different groups of learners in novel contexts before a consensus can be reached. For these reasons, there is a need for additional investigations into the effects of a cyclical, cognitive and metacognitive strategy-based process-approach on L2 listening comprehension.

The present study sought to fill that need, and was a quasi-replication of Vandergrift and Tafaghodtari’s (2010) experimental design. Modifications to that design were intended to incorporate relevant elements of previous investigations in order to create a maximally effective pedagogical cycle. To that end, Vandergrift’s (2004) model was used, but the third verification stage was replaced by bottom-up processing practice as utilized by Carrier (2003), Goh and Taib (2006), and Graham and Macaro (2008) and
advocated, but not utilized, by Vandergrift and Tafaghodtari (2010). Additionally, the present study incorporated reflection diaries as advocated by Goh (1997) and as used with the high scaffolding group in Graham and Macaro’s (2008) investigation. Furthermore, individual strategies were introduced to students following the guidelines detailed by Chamot et al. (1999) and Wenden (1991). Finally, the instruments used to assess both listening and metacognitive awareness were very similar to those used by Vandergrift and Tafaghodtari (2010). This is especially important in terms of the test of listening comprehension, because the Vandergrift and Tafaghodtari study was the only strategic intervention reviewed here in which the learners’ gains reached statistical significance on a standardized, institutional test of listening comprehension, and until listeners’ gains can be successfully measured by standardized tests, the value of training of this nature will remain questionable.
CHAPTER 3: METHODOLOGY

Research Questions

Classroom-based, empirical investigations into the influence of instructional models need to be replicated with different groups of learners, learning different languages in different contexts in order to shed light on which approaches are the most effective for the greatest number of learners. The present study sought to examine the effect of a cyclical, cognitive and metacognitive strategy-based, process-approach to listening instruction on both the self-reported levels of metacognitive awareness and listening comprehension of multiple groups of university-aged learners of intensive, full-time English. One group, consisting of two classes, or cohorts, received this experimental treatment, and the other group, consisting of one class (cohort), received an alternate form of instruction which was not controlled by the researcher.

The research questions to be examined during this investigation are as follows:

1. Are there significant differences between the experimental and comparison groups in terms of either listening comprehension or metacognitive awareness after the treatment?

2. Are there significant differences between the experimental and comparison groups’ pre-test and post-test scores of either listening comprehension or metacognitive awareness?

3. Is there a correlation between self-reported metacognitive awareness and listening comprehension?
4. What are the participants’ attitudes and opinions regarding each method of instruction, and what are the similarities and differences of the opinions of the members of each group?

Participants

Selection

All of the students enrolled in intermediate level Listening & Speaking classes (n=28) in an Intensive English Program (IEP) participated in this research. Placement in the intermediate level was determined by the IEP and based upon previous teachers’ recommendations, Michigan Placement Test (MPT) total scores, and a timed composition test given at the end of each term and scored anonymously by the instructors in the IEP. All of these participants planned to enroll in graduate or undergraduate level university classes upon completion of their full-time English studies. As will be discussed in the upcoming Procedures section, these participants were divided into three separate classes, or cohorts. Two of the cohorts received the experimental treatment and one did not. In Table 2, the participants are categorized according to their cohort and experimental status. Basic demographic information regarding their genders, native languages, and ages at the time of the research is also presented.
Table 2

Participants’ Demographic Information and Experimental Status

<table>
<thead>
<tr>
<th>EXPERIMENTAL GROUP PARTICIPANTS</th>
<th>Cohort 1</th>
<th>Cohort 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant ID</td>
<td>L1 Gender Age</td>
<td>Participant ID</td>
</tr>
<tr>
<td>P1A</td>
<td>Chinese F 19</td>
<td>P2A Chinese F 22</td>
</tr>
<tr>
<td>P1B</td>
<td>Russian F 19</td>
<td>P2B Chinese M 19</td>
</tr>
<tr>
<td>P1C</td>
<td>Chinese M 19</td>
<td>P2C Arabic M 26</td>
</tr>
<tr>
<td>P1D</td>
<td>Chinese F 20</td>
<td>P2D Chinese F 19</td>
</tr>
<tr>
<td>P1E</td>
<td>Chinese M 18</td>
<td>P2E Chinese F 20</td>
</tr>
<tr>
<td>P1F</td>
<td>Chinese F 19</td>
<td>P2F Chinese F 18</td>
</tr>
<tr>
<td>P1G</td>
<td>Chinese M 20</td>
<td>P2G Chinese F 20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P2H Chinese M 18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P2I Chinese F 20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P2J Chinese M 21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P2K Arabic F 23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COMPARISON GROUP PARTICIPANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant ID</td>
</tr>
<tr>
<td>P3A</td>
</tr>
<tr>
<td>P3B</td>
</tr>
<tr>
<td>P3C</td>
</tr>
<tr>
<td>P3D</td>
</tr>
<tr>
<td>P3E</td>
</tr>
<tr>
<td>P3F</td>
</tr>
<tr>
<td>P3G</td>
</tr>
<tr>
<td>P3H</td>
</tr>
<tr>
<td>P3I</td>
</tr>
<tr>
<td>P3J</td>
</tr>
</tbody>
</table>

As Table 2 displays, there were almost twice as many experimental participants (n=18) as comparison participants (n=10). 86% of the participants were between the ages of 18 and 22. The other 14% were between the ages of 23 and 27 and were all native
Arabic speakers. In fact, only one of the Arabic speakers who participated in this research was in the 18 to 22 age range; the rest of them were slightly older than their classmates. Table 2 also illustrates the fact that the vast majority of the participants in this research were Chinese native speakers (75%). The next largest native language group was Arabic (18%). One native Russian speaker and one native Korean speaker also took part. The comparison group, with three Arabic speakers and the Korean participant, was somewhat more diverse than the either of the experimental cohorts, which contained only one or two non-Chinese speakers. Finally, there were slightly fewer male (39%) than female (61%) participants.

Intermediate level learners were selected for two reasons. First, they were at the appropriate linguistic level for training of this nature. There is preliminary evidence that process-oriented training of this type is most beneficial to listeners in the early stages of second language acquisition (Chamot, 2005; Goh & Taib, 2006; Rost & Ross, 1991; Vandergrift, 1997; 2007; Vandergrift & Tafaghodtari, 2010), and Vandergrift (1997) found that the first two years of language learning are a “pivotal” (p. 406) time for the acquisition of effective learning strategies. However, the effects of strategy training in the L2 for beginning level learners has been insufficiently researched (Chamot, 2004, 2005). As metacognitive strategy use places demands on the same attentional resources used to comprehend L2 input, intermediate level learners’ abilities to process large chunks of information, thus freeing those resources for strategy implementation, place them in the best position to benefit from strategy training of this nature (Vandergrift, 1997).
Second, the intermediate level Listening & Speaking classes in this IEP devote 65% of their class time to listening practice. In fact, one of the primary objectives at this level is to vastly improve the learners’ listening skills in order to prepare them for the challenges of the subsequent advanced level Listening, Speaking, & Note-taking class, which is intended to prepare them for the rigors of academic study in the university (M. Hunter, personal communication, January 4, 2011). Conversely, at the beginning level, listening practice is shared with an equal amount of speaking practice, and in the advanced Listening, Speaking, & Note-taking classes learners only spend 40% of class time on the listening skill. Therefore, for an investigation of a methodology intended to improve listening comprehension, the intermediate level listening classes have the most listening time built into the curriculum (OPIE Curriculum Guide).

Placement

Full-time students in this IEP are placed in particular levels through a combination of teacher recommendations (for continuing students), their results on a timed composition test scored anonymously by instructors in the IEP, and their results on the MPT (for new and continuing students). The MPT assesses three language areas: listening, reading, and receptive grammar knowledge, and it is the cumulative score which determines placement. As listening scores are not considered separately, the possibility of outliers, or participants with high or low enough listening abilities to pull their classes out of range, was present at the beginning of the treatment. To insure that there was no statistically significant difference in entry-level listening comprehension between the two groups and to check for outliers, the MPT listening scores of all
potential participants were extracted and analyzed. For each group (experimental and comparison) and for the participant pool as a whole, both the mean and standard deviation (SD) of the MPT were calculated and examined for any scores falling two or more SDs from the mean. All of the potential participants’ scores were within the acceptable range of less than two SDs from the mean. Additionally, this data was examined for outliers with the generation of a box plot by the statistics program R, which illustrated that there were no outliers among the potential participants. As all of the students enrolled in the intermediate level had statistically similar listening proficiency scores, all were included as participants in the research.

Instructors

Two classes, or cohorts, received the experimental treatment (n=18), and one class (cohort) at the same level acted as a comparison group (n=10) and received non-experimental instruction. The comparison group was taught by an instructor with a Master of Arts in Teaching English to Speakers of Other Languages (MA TESOL) and experience in intensive English instruction (Ohio University website, 2010). The two experimental cohorts were instructed by the researcher, who also had experience teaching intensive English and was pursuing an MA TESOL at the time of the research. The researcher-as-teacher model has been shown to contribute to the overall success of strategic interventions, due to both increased levels of understanding about the nature of learning strategies and researchers’ increased levels of commitment to strategy instruction (Chamot, et al., 1999; Thompson & Rubin, 1996). The comparison group’s instructor was selected by the administration of the IEP. His selection was based on the
administrative and organizational criteria of the program and can be considered random for the purposes of this research. The researcher did not instruct the comparison classes in order to prevent overlap between the two methodologies. This design allowed each teacher to provide instruction “to the best of his or her ability” (S. Jarvis, personal communication, September 27, 2010), so that the experimental methodology could be compared to instruction typical at this level and in this context.

Materials

*Measuring Listening Comprehension*

At the beginning and end of every quarter, all students whose Test of English as a Foreign Language (TOEFL) scores have placed them into full-time (non-academic) intensive English classes must sit for the Michigan Placement Test (MPT), as this test is considered by the IEP to be better able to discriminate between lower levels than the TOEFL exam (J. Bagnole, personal communication, September 29, 2010). In order to minimize the level of inconvenience to the students as well as to employ an instrument already in use within the institution, this test was used as the basis for determining the participants’ pre- and post-treatment listening comprehension levels.

*Measuring Metacognitive Awareness*

The participants also completed a retrospective self-report questionnaire designed to measure metacognitive awareness at both the beginning and end of the treatment. The instrument is called the Metacognitive Awareness Listening Questionnaire (MALQ) (Appendix A). It was developed by Vandergrift, et al. (2006) and underwent rigorous validity testing, including exploratory and confirmatory factor analyses and multiple field
tests with large samples (n=966) in a variety of language learning contexts, several of which mirror the context of use in the present study. For example, 65% of the field test participants were university students, as were the participants in the present investigation. Additionally, 22% of the field test participants were studying English as a foreign language; the participants in the present study were studying English as a second language. Finally, 193 of the field test participants were Asian (from Singapore), and 75% of the participants in the present study were Chinese.

The MALQ was found to have robust psychometric properties and high internal validity (Goh, 2008; Vandergrift et al., 2006). The instrument published in Vandergrift et al. (2006) was designed for native English speakers learning French as a second language, but was intended for broader use and intentionally written in simple language (Goh, 2008). In order to insure comprehensibility for this group of participants, it was piloted with a group of intermediate level learners in the Fall Quarter, 2010. The details of this piloting will be discussed in the upcoming section.

Pilot of the Metacognitive Awareness Listening Questionnaire (MALQ)

During the Fall Quarter of 2010, the MALQ was piloted with a group of learners (n=10) at the intermediate level. This pilot was intended to insure that the instrument be understandable to learners at this level and in this context. These learners were given copies of the questionnaire as it appeared in Vandergrift et al. (2006), and were asked to identify (by circling) any words, phrases, or entire questions which they did not completely understand. The researcher was present during this activity, and circulated amongst the learners, questioning them about their responses and asking meaning check
questions to insure that those areas identified as comprehensible were actually being comprehended. All of the sections of the questionnaire which were identified as potentially problematic by these learners were rewritten into simplified English.

Measuring Learners’ Perceptions of the Instruction They Received

At the end of the treatment, participants in both groups were asked to complete an on-line survey (Appendix B) regarding their perceptions of the instruction they received. This survey targeted both the participants’ beliefs about their overall improvement in listening proficiency as well as their opinions regarding exactly which elements of the instruction they received (whether it be experimental or comparison) were the most (and least) beneficial. Finally, each cohort was invited to participate in a separate focus group interview. These interviews were intended to give learners an opportunity to expand upon answers given in the survey as well as to reflect upon their instruction. Both of these measures served the function of “member checking” (Dörnyei, 2007, p. 60), and allowed the researcher access to the thoughts and opinions of the students regarding the methodology and instruction they received. The interviews were conducted by the researcher, and the threat of an insider-outsider conflict was diminished by the use of carefully designed questions which targeted specific activities and instructional methodologies rather than overarching perceptions of the course or instructor (D. Bikowski, personal communication, September 30, 2010).
Procedure

*Experimental Group Treatment*

Both the experimental and comparison groups met the IEP’s curricular objectives for the quarter (Appendix C), but the experimental group did so via a specific pedagogical cycle intended to increase their metacognitive awareness, develop their understanding of the processes underlying listening comprehension, and train them in the effective use of strategies for improving listening performance (Field, 2000; Vandergrift, 2004; Vandergrift & Tafaghodtari, 2010). The primary focus of this cycle was on explicit instruction in the metacognitive processes and related strategy clusters (groups of strategies) which underlie successful L2 listening, based on Graham’s (2006) findings that when learners lack an “overarching metacognitive strategy to monitor the effectiveness of [individual strategies’] use” (p. 178) they have a tendency to perceive listening as difficult and to perceive themselves as incapable listeners, as well as findings that instruction in individual strategies can improve listening performance in immediate contexts but not in the long term (Field, 2000). The cycle included five stages, each of which was directly related to a specific metacognitive process or pair of processes, as explained in Table 3. This cycle was originally developed by Field (2000), expanded upon by Vandergrift (2004), and examined in an empirical investigation by Vandergrift and Tafaghodtari (2010). The model used in the present research was a modification of the one investigated by Vandergrift and Tafaghodtari with the following alterations:

- The incorporation of a bottom-up processing component as advocated by Vandergrift and Tafaghodtari (2010).
• Addition of increased opportunities for reflection as advocated by Graham and Macaro (2008).

• Utilization of video-based aural/oral listening diaries which were modeled after Goh’s (1997) listening diary research, intended to provide learners with regular opportunities for reflection.

• Slight modifications to the ways in which the cycle was presented week after week, in order to prevent it from appearing “generic” and “bor[ing]” (p. 22) to the learners, as had been the case in Vandergrift and Tafaghodtari’s (2010) investigation.
### Table 3

**Stages of Instruction and Related Metacognitive Processes and Strategies**

<table>
<thead>
<tr>
<th>Stages of Instruction</th>
<th>Metacognitive Processes</th>
<th>Individual Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Planning and Predicting Stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Learners are informed of the topic and text type (interview, radio program, etc.) and predict what types of information and specific words might be heard. Learners set personal listening goals.</td>
<td>Planning</td>
<td>Metacognitive: Prediction&lt;br&gt;Activation of background knowledge&lt;br&gt;Directed attention&lt;br&gt;Goal setting&lt;br&gt;Socio-Affective: Cooperation</td>
</tr>
<tr>
<td><strong>First Verification Stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Learners listen to the text to check predictions and make notes about any other information they understand during the listening. Learners identify personal problem spots in need of selective attention in the subsequent listening and check personal listening goals.</td>
<td>Monitoring</td>
<td>Metacognitive: Selective attention&lt;br&gt;Comprehension monitoring&lt;br&gt;Double-check monitoring&lt;br&gt;Prediction verification&lt;br&gt;Cognitive: Linguistic inferencing&lt;br&gt;Between-parts inferencing&lt;br&gt;Personal elaboration&lt;br&gt;World elaboration&lt;br&gt;Creative elaboration&lt;br&gt;Evaluating</td>
</tr>
<tr>
<td><strong>Second Verification Stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Learners compare notes with a listening partner or group of partners, to make comparisons, discuss discrepancies, and identify additional problems areas requiring selective attention during the next exposure.</td>
<td>Evaluating</td>
<td>Metacognitive: Problem identification&lt;br&gt;Socio-affective: Cooperation</td>
</tr>
<tr>
<td>4. Learners set new personal listening goals and listen to the text to determine discrepancies between their original notes, what their partner(s) had understood, and information now</td>
<td>Planning</td>
<td>Metacognitive: Goal setting&lt;br&gt;Selective attention&lt;br&gt;Directed attention</td>
</tr>
</tbody>
</table>
understood. They make corrections as necessary and write down any additional information.

5. Learners check personal listening goals, and engage in small or whole group discussions aimed at reconstructing the main points and relevant details. These discussions also involve learners’ reflections of how they were able to understand the text, how they understood the meaning of particular words, and how they solved problems while listening.

*Twice during the treatment, learners created written summaries (for assessment), in which they included everything they could understand in a logical summary.

6. Learners receive transcript of text for bottom-up processing practice. Learners compare spoken speech to written form and perform tasks such as marking stressed words, identifying intonation contours, or distinguishing reduced forms to develop auditory discrimination and word recognition skills.

7. Learners discuss reflections of the listening experience in pairs, small groups, and individually in video-diaries. These reflections include problems, successes, experiences, and strategies used in and out of class.

* These video-diary entries were used for process-based discussions in which learners shared strategies and suggestions with one another, as per Goh’s (1997) suggestions.

Adapted from Vandergrift (2004)
Learners in the experimental classes participated in this cycle seven times during treatment. Each cycle focused on a different topic; all topics were taken from the course book (Mackey, Blass, & Kisslinger, 2007). As Table 3 shows, each stage of the cycle is accompanied by a metacognitive process or pair of metacognitive processes. These metacognitive processes are actualized via the incorporation of specific metacognitive strategies in conjunction with particular cognitive or socio-affective strategies. For example, during the Planning and Predicting Stage, the metacognitive process of planning was executed through the use of the metacognitive strategies of prediction, activating background knowledge, goal setting, and directed attention, with the simultaneous utilization of the socio-affective strategy of cooperation.

The individual metacognitive, cognitive, and socio-affective strategies selected for inclusion were chosen based on both their usefulness in guiding the learners through the metacognitive cycle under investigation (Vandergrift, 2004; Vandergrift & Tafaghodtari, 2010) and the specific needs of the group of learners involved in this training (Goh, 2008; Grahan & Macaro, 2008; Wenden, 1991). These strategies were introduced in five out of the 10 weeks of the treatment, in order to allow the participants ample opportunities to practice each strategy or set of strategies (Chamot, 2005; Chamot et al., 1999; Field, 2000; Mendelsohn, 1998). A schedule detailing the order in which they were introduced is presented in Table 4.
Table 4  
*Strategy Presentation Schedule*

**WEEK TWO**

<table>
<thead>
<tr>
<th>Metacognitive Strategies</th>
<th>Socio-affective Strategies</th>
<th>Cognitive Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prediction</td>
<td>Cooperation</td>
<td></td>
</tr>
<tr>
<td>Activating Background Knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selective Attention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prediction Verification</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**WEEK THREE**

<table>
<thead>
<tr>
<th>Metacognitive Strategies</th>
<th>Socio-affective Strategies</th>
<th>Cognitive Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directed Attention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problem-identification</td>
<td></td>
<td>Summarization</td>
</tr>
<tr>
<td>Performance Evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strategy Evaluation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**WEEK FOUR**

<table>
<thead>
<tr>
<th>Metacognitive Strategies</th>
<th>Socio-affective Strategies</th>
<th>Cognitive Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal setting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension Monitoring</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-check monitoring</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal checking</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**WEEK SEVEN**

<table>
<thead>
<tr>
<th>Metacognitive Strategies</th>
<th>Socio-affective Strategies</th>
<th>Cognitive Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**WEEK NINE**

<table>
<thead>
<tr>
<th>Metacognitive Strategies</th>
<th>Socio-affective Strategies</th>
<th>Cognitive Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As Table 4 shows, students were explicitly introduced to a total of 19 strategies during the course of the treatment. 12 of these were metacognitive strategies necessary for proper implementation of the metacognitive pedagogical cycle under examination. For this reason, learners were introduced to the pertinent metacognitive strategies prior to
being presented with the cognitive strategies selected for inclusion. Additionally, those strategies introduced first were the ones most likely to be familiar, already in use, or considered easy by the learners, as advocated by Chamot et al. (1999). As a result, inside-the-head strategies such as comprehension monitoring were the last of the metacognitive strategies to be incorporated into the cycle. After the entire metacognitive cycle had been explicitly presented, learners were given two weeks to practice using these strategies before any new strategies were introduced. Of the six cognitive strategies, five were selected based on the specific needs of this group of learners in this context, as advocated by Wenden (1991). The seventh, summarization, was included in order to be able to insert assessment into a non-evaluative framework. The single socio-affective strategy, cooperation, was included because it was necessary for proper implementation of the cycle. After being introduced to each new set of cognitive strategies, the learners spent a week practicing using those strategies before being introduced to any more (Chamot, 2005; Chamot et al., 1999; Field, 2000; Mendelsohn, 1994). Students were given multiple opportunities to work with and practice each strategy (Field, 2000) and were encouraged to use the strategies in effective and self-selected combinations, primarily through the use of regular self- and strategy-evaluation (Chamot et al., 1999; Graham & Macaro, 2008; Macaro, 2006; Vandergrift, 2004).

Each strategy was introduced during class time, and learners were given a handout explaining each strategy in simplified, easily understandable language (see Appendix D for an example). These handouts provided learners with each strategy’s name, a graphic representation of the strategy, a simple definition, an example of the
strategy in use, information regarding when the strategy could be employed, and an explanation of the strategy’s potential value. These handouts were created by the researcher and based on information taken from both Chamot et al. (1999) and Vandergrift (1997). Learners were required to keep all of these handouts together in their class binders, and were invited to refer to them periodically throughout the treatment. Likewise, all predictions, notes, identified problems, and personal listening goals were recorded on listening pages (see Appendix E) which learners also kept in their class binders. These pages were designed by the researcher and intended to provide learners a cohesive and organized space for all of the predictions and notes which they produced during the course of each cycle. These listening pages were an adaptation of the “carnet d’écoute” (p. 9) utilized by Vandergrift and Tafaghodtari (2010), which achieved the similar goal of providing each learner with a single location for the recording of predictions, notes from the first and second exposure, and reflections during listening practice.

**Documentation of Comparison Group Treatment**

In order to document any overlap between the experimental and comparison treatments, the researcher observed the comparison class twice during the course of the treatment, collected documents (course syllabus, homework assignments, and tests), and interviewed the instructor once the treatment had ended. Field notes collected during the classroom observations were transferred onto an extensive event sampling checklist (Dörnyei, 2007), designed by the researcher and intended to quantify instances of elements integral to the experimental methodology (Appendix F). All data obtained by
these measures was used to provide a rich description of the nature of the instruction taking place in the comparison classes, with the aim of isolating both those variables which were present in the two methodologies and those which were distinct and therefore under examination. In addition to the aforementioned procedures, pilot observations were also conducted in the Fall Quarter of 2010, with the goal of determining which elements can be considered typical in the context of this IEP at the intermediate level. The event sampling checklist used to quantify the instruction taking place in the comparison class was based on data collected during these pilot observations, the details of which will be discussed in the following subsection.

**Pilot Observations**

In the quarter prior to the investigation, observations were conducted in every Listening & Speaking class at both the intermediate and high-intermediate levels in this IEP. These were intended to determine whether or not elements of the metacognitive pedagogical cycle under examination are typically present at this level and to identify those elements which can be considered typical of listening instruction at this level. Six separate classes were observed, for a total of seven hours of observation. One class was observed on two occasions because the listening activity under observation spanned two class periods. Four of these classes were at the high-intermediate level and two were at the intermediate level, though it should be noted that there is no difference between these two levels in terms of curricular objectives (OPIE Curriculum Guide, AE 45 Listening & Speaking).
On the basis of these observations, typical intermediate level listening instruction in this IEP can be distinguished from the metacognitive pedagogical cycle under investigation in the following ways:

- Individual listening strategies are sometimes presented, but they are not presented within a semi-standardized cyclical format intended to guide learners through the metacognitive processes of planning, monitoring of comprehension, self-evaluation, or reflection on the listening process.

- Individual elements of the metacognitive cycle under investigation are sometimes present (learners may be asked to utilize selective attention, for example), but these elements are separated from the rest of the cycle, and learners are rarely informed that they are utilizing a strategy.

- Individual strategies are sometimes practiced without the students being told the name, usefulness, or potential applications of those strategies.

- Class discussions focus on the content of the listening selections rather than the listening process or the effective or ineffective use of strategies to achieve success.

- Learners do not make predictions regarding specific lexical items that they might hear, and learners are not always asked to actively participate in the prediction process.

- Learners do not set personal listening goals and are not given opportunities to reflect on their progress as L2 listeners.

- Bottom-up processing practice is not generally incorporated.
Comparison Group Treatment

The instructor of the comparison group was aware that his class was involved in a research project regarding listening instruction, but he was not given the details of the experimental treatment. He instructed his class in the manner and style which he deemed most appropriate for achieving the IEP’s curricular objectives (Appendix C). The purpose of this investigation was not to restrain him from teaching to the best of his ability. There was, therefore, some overlap between the experimental treatment and what occurred in the comparison classes.

In order to document this overlap and to provide a rich description of the comparison group’s instruction, this class was observed twice during the term; documents such as the course syllabus, homework assignments, and tests were collected; and the instructor was interviewed at the end of the term. The comparison group’s instruction was defined on the basis of these three methods of data collection, and can be summarized as a practice-based, product-approach to listening in which a small number of strategies were introduced “casually” and “sporadically,” (Comparison Group Instructor, personal communication, March 15, 2011) based on the assumption that the majority of learners had already acquired these strategic listening behaviors on their own. The instructor’s primary goal was to provide the learners with “rich listening experiences of various kinds,” later followed by “talking in a minor way about metacognitive issues” because the metacognitive processes underlying the listening process were “not the focus of how I teach listening” (Comparison Group Instructor, personal communication, March 15, 2011).
The comparison group practiced listening in some form or another during every lesson. When working with the course book (Mackey et al., 2007) learners would first join with a designated partner, and then discuss a either a set of images or a series of content-based discussion questions designed to help them activate background knowledge about the topic. These questions incorporated vocabulary and topics related to the upcoming listening text, and the students were informed that they should think about what they already knew about the topic. The remainder of the pre-listening activities involved whole group discussions of both the pronunciation and meaning of lexical items integral to the upcoming selection. Learners were then asked to scan the list of multiple-choice comprehension questions included in the book, one of which targeted the main idea of the text, seven of which focused on details, and two which required the learners to make inferences about the selection (Mackey et al., 2007). The learners were then given the opportunity, though not required, to make predictions about the content of the upcoming selection prior to listening. Learners were then exposed to the selection twice, with the goal of correctly answering the comprehension questions. Therefore, these practice activities can be defined as product-based, as the learners’ comprehension was measured by their ability to correctly answer multiple-choice questions (Goh, 2008). In between the two exposures, the learners discussed their responses with a partner. They were given explicit information regarding how to cooperate effectively, such as discuss discrepancies, help each other, try and convince the other if you are sure of your selection, etc. After the second exposure, one student, called a Guest Teacher, was invited to the center of the room to go through the correct answers to the questions.
Comprehension breakdowns or other evidence of misunderstanding were either discussed by the entire group or ignored, depending on individual learners’ willingness to make their incorrect answer known. Misunderstandings were dealt with by placing the transcript onto the course management system (CMS) for those learners who desired further, independent clarification.

In the first half of the term, there was no mention of the concept of strategies. During some of those lessons which took place in the second half of the term, learners were informed that using what you already know, guessing, and predicting are all activities which can improve listening comprehension. They were also informed that these activities could be engaged in outside of the classroom and in a variety of L2 listening contexts for improved comprehension (Chamot et al., 1999).

When practicing listening with supplemental materials, very different approaches were utilized. Learners in the comparison classes engaged in two additional types of listening practice, one of which was termed Fast Listening and the other Conversation Starters by the instructor. Fast Listening activities involved “quirky” (Comparison Group Instructor, personal communication, March 15, 2011) headlines taken from NPR’s Morning Edition program. The pre-listening phase was eliminated from Fast Listening exercises, in order to more closely mimic real-world listening situations. Learners were asked to listen to the short (thirty second to one minute) headlines, which were delivered in rapid, authentic native English speech. While listening, they were instructed to note any words or phrases which they could understand. After the headline had finished, learners would share what they had gleaned from the selection, and the instructor would
write these words and phrases on the board. They would then listen again to clarify what was already written on the board and to catch additional words and phrases. This additional information would then be added to the board, and the learners would break into small groups to try and develop oral summaries of what they had understood. Each group’s interpretation of the text was then shared with the rest of the class. The instructor remained neutral during this sharing phase and provided no clue as the accuracy of these summaries. After each group had shared its summary, learners were provided with a transcript and a brief, simplified summary from the instructor. These Fast Listening activities were similar to the experimental treatment in that comprehension was determined by the learners’ abilities to reconstruct the text. However, they can be distinguished from the experimental treatment by the lack of pre-listening activities and the lack of pre-, during, or post-listening strategy instruction or use.

The Conversation Starters were intended to provide the learners with authentic opportunities for “talking with each other and negotiating meaning” (Comparison Group Instructor, personal communication, March 15, 2011). Students were given Conversation Starter strips containing questions which lent themselves to short conversations. After engaging in one short conversation with one partner, the students would rotate, so that in the course of a single Conversation Starter activity, each learner would converse with a large number of partners on a variety of related topics. Each Conversation Starter contained vocabulary from and was connected to the current unit in the course book (Mackey et al., 2007), which means they supported content-based (as opposed to process-based) discussions. Further comparisons and contrasts between the two approaches will
be described in the upcoming subsections, which break these comparisons down according to elements integral to the experimental treatment.

Comparison and Contrast of the Experimental and Comparison Group Treatments

Strategy Instruction

Learners in both classes were explicitly taught about the concept of strategies. However, in the experimental classes instruction was focused on those metacognitive strategies that regulate and control the listening process, and in the comparison classes the focus was on “giving them lots of [listening] experiences” (Comparison Group Instructor, personal communication, March 15, 2011). Thus it can be said that the experimental treatment was strategy-based, whereas the comparison group’s instruction was experience-based.

Learners in the experimental classes were explicitly presented with all of the strategies in Table 4. They learned their names, values, and potential applications (Wenden, 1991) and engaged in discussions regarding when each strategy could be used, when it could not be used, if it had been used previously, and how effective it had been (Chamot et al., 1999). Experimental group participants were also exposed to additional cognitive strategies, such as identifying stressed syllables and reduced forms, during the bottom-up processing stage of the cycle (see Third Verification Stage, in Table 3, for a complete list). These were not referred to as strategies by the instructor or presented in the same way as the other 19 strategies. These were treated as receptive pronunciation activities and were always followed by practice in which the learners were asked to produce these features in their own speech, based on the understanding that listening and
pronunciation are “inseparable” (Mendelsohn, 1998, p. 88) and are best approached in an integrated fashion.

Learners in the comparison group were introduced to two metacognitive strategies (activating background knowledge and making predictions) and one cognitive strategy (inferencing) during the course of the term. These strategies were given simplified names, practiced, and discussed in terms of other potential applications beyond the classroom context. These three strategies were introduced mid-way through the course, as the first half of the term was devoted exclusively to listening practice. These strategies were introduced “sporadically” and “casually” (Comparison Group Instructor, personal communication, March 15, 2011), based on the assumption that most learners had already adopted these strategies on their own. Additionally, the cognitive strategies of listening for prominent words, listening for intonation cues, and summarization were also regularly employed. These strategies were given simplified names and practiced extensively. Finally, the socio-affective strategy of cooperation was heavily encouraged and employed, though it was not given a simplified name as the others were, on the assumption that this would have been unnecessary.

The role of prediction

The role of prediction in the two classes merits special mention. The importance of pre-listening predictions has long been acknowledged as an important contributor to improved comprehension (Goh, 2008). Therefore, prediction is a widely utilized and oft-practiced strategy in L2 listening classrooms. Not surprisingly, pre-listening prediction took place in both classes; however, it was conceptualized and utilized in markedly
different ways. In the experimental classes, each individual learner was required to actively take part in the prediction-making process at the beginning of every listening cycle. Sometimes, this prediction stage was engaged in as a whole group activity, in which each learner was expected to contribute; at other times prediction was treated as a small group activity; and the remainder of the time learners were asked to view images or videos on the CMS related to the upcoming topic and make individual predictions in anticipation of the upcoming lesson, which they would then share with one another prior to listening. These pre-listening predictions focused not only on content but also on specific lexical items which might be encountered, and were always recorded in the same location on the learners’ listening pages (Appendix E).

In the comparison classes, individual learners were invited, though not required, to take part in pre-listening predictions. Prior to course book listening activities, the whole group engaged in discussions designed to activate background knowledge on the topic, and questions regarding possible upcoming topics were addressed to the group. Additionally, learners were invited to predict the answers to the comprehension questions in the book during some, but not all, of the course book listening activities. Prediction was not utilized prior to Fast Listening activities, in order to more closely mimic real-world listening situations. As such, learners were asked to rely upon linguistic knowledge to pull words and phrases from the uninterrupted speech stream.

**Bottom-up Processing Practice**

In the comparison classes, bottom-up processing practice was incorporated as an optional homework activity, in which they were instructed to listen and read while
focusing on intonation cues, stressed syllables and reduced forms. Additionally, learners in the comparison group were reminded on a regular basis of the presence of reduction in native speech, were given one explicit and focused lesson on the phenomenon of linking, and provided with periodic reminders to listen for these elements in both in-class activities and homework assignments. Finally, learners in the comparison group were always given the transcript following Fast Listening exercises, which they would read this silently to themselves. In the experimental classes, bottom-up processing practice consisted of activities which focused on either reduced forms, prominent syllables, or other specific elements present in natural speech. These activities were a mandatory part of in-class listening practice and always employed the same texts that had previously been used for top-down, strategic listening practice.

Additional materials and activities used to practice listening in the classroom.

In the experimental classes, the metacognitive pedagogical cycle was always practiced with selections taken from the course book (Mackey et al., 2007), as per various researchers’ suggestions that strategic, process-based methods can be incorporated into any pre-existing listening course (Chamot et al., 1999; Goh, 2008; Vandergrift, 2004). Learners were exposed to each of these course book texts three times. These repeated exposures were intended to increase the listeners’ awareness of the metacognitive processes underlying listening comprehension, to allow them to repeatedly make predictions, monitor comprehension, evaluate comprehension, and make plans for dealing with problems encountered during previous exposures, as well as to provide them with an
opportunity for bottom-up processing practice with every listening text (Vandergrift & Tafaghodtari, 2010).

On the other hand, comparison group participants engaged in three very different types of in-class listening practice: course book work, Fast Listening, and Conversation Starters. During both course book practice and Fast Listening exercises, learners were exposed to each selection two times. As was previously stated, bottom-up processing was available as an optional homework activity following course book listening practice, and learners were given the transcript to read silently to themselves following Fast Listening practice. During Conversation Starters, learners engaged in interactional listening practice and were invited to ask clarification and meaning check questions in order to negotiate meaning with one another.

**Underlying purpose of listening practice activities in the classroom**

In the experimental classes, the purpose of listening was intended to mirror the purposes involved in real-world listening. To that end, the focus was on the comprehension of main idea and details and the ability to make inferences. Learners demonstrated comprehension by checking their predictions, noting what they had understood while listening, comparing those notes with listening partners at designated intervals, and engaging in both oral and written reconstructions of the text (Vandergrift & Tafaghodtari, 2010). Each listening activity was presented as a series of three well-defined stages (pre-, during-, and post-listening), each of which is connected to a set of strategies which, if used in effective combinations, are theorized to have the potential to improve listening performance.
In the comparison classes, the purpose for listening was connected to practice type: course book, *Fast Listening*, or *Conversation Starters*. When working with the course book, the goal of the listening was to correctly answer the multiple-choice questions presented in the text (Mackey et al., 2007). However, when participating in *Fast Listening* exercises, the purpose for listening was very similar to the one in the experimental cycle, in that it focused on comprehension of the main idea and details and the ability to make inferences. Learners demonstrated comprehension by noting what they understood while listening, comparing these notes with other members of the class, and through oral reconstructions of the text. The primary differences between the *Fast Listening* and the metacognitive cycle under investigation were related to pre-listening activities, strategy instruction, and post-listening evaluation and reflection. Finally, during *Conversation Starters* activities, the purpose was to interact with peers and negotiate meaning.

*Materials and activities used to practice listening in homework assignments.*

Learners in both the experimental and comparison groups had access to the same independent listening practice activities linked through identical course management systems (CMS). From the CMS, learners had access to two different web-sites providing listening practice to English language learners, as well as a variety of short videos on such topics as American culture, idioms, and slang. The CMS was also linked to a series of sites providing assistance with aspects of pronunciation, from work with individual sounds, to discrimination between minimal pairs, to suprasegmental features such as intonation and rhythm.
Learners in the comparison group used one of these listening practice sites as the basis for audio journals which were due at the end of each week. They were required to listen to any selection designated as medium or difficult on the site and then produce recorded oral summaries of what they had understood. At the beginning of the term, they were expected to create two of these each week. This homework requirement was increased to three a week during the third week, and again increased to four in the seventh week. Each week the instructor selected one of these journals at random for assessment and feedback purposes.

Learners in the experimental group were also required to produce journals twice a week, and this amount was never increased. Experimental participants recorded these journals with video-based technology inside the CMS. The instructor video-recorded herself giving each assignment, and the students had to listen to this video and then respond appropriately. One of these weekly journals was content-based. These assignments involved the viewing of authentic, real-time internet videos related to class topics and either summarizing the content or answering a set of comprehension questions provided in the instructor’s video. The other set of video journals were process-based and covered topics such as “Problems I encountered in today’s listening lesson” or “How, when and why I have used a particular strategy.” These journals encouraged learners to reflect on strategy use and on their strengths, weaknesses, and progress as L2 listeners. They received immediate feedback on these reflections, in the form of handwritten, personal notes from the instructor which they received in class the day after the assignment had been due (Graham & Macaro, 2008).
The role of discussion in the classroom

Experimental group participants took part in process-based discussions throughout the treatment, in which they shared information about strategy use with one another. The instructor mediated these discussions, and the goal was for learners to share strategies and have an opportunity to learn from one another (Goh, 1997). Content-based discussions were built into the metacognitive pedagogical cycle under examination and took place immediately after both the First and Second Verification Stages. At these times learners would break into small groups and share information which they had understood with each other, note discrepancies, and work together to solve listening problems or reconstruct the text. During some cycles, content-based discussions also took place prior to the Planning and Predicting Stage, as learners conversed with one another about the topics and themes of the upcoming listening, or worked together to make predictions about the topic.

Learners in the comparison classes engaged in content-based discussions which were handled in two separate ways. First, learners shared what they had understood with a partner with the goal of reconstructing the text during Fast Listening exercises. This was similar to the content-based discussions engaged in by the experimental classes. Second, Conversation Starters activities provided comparison group participants with regular, content-based discussion opportunities.

Assessment

One of the primary theoretical tenants of the metacognitive pedagogical cycle under examination is that listening practice be separated from listening performance,
based on research showing that the threat of evaluation can increase learners’ anxiety. When both the instructor and the listeners focus on the process of listening, rather than correct (or incorrect) answers on comprehension questions, learners can instead focus on increasing their control over the listening process (Goh, 2008; Graham & Macaro, 2008; Vandergrift, 2002; 2007; Vandergrift & Tafaghodtari, 2010). However, the IEP required that learners’ listening comprehension be assessed periodically throughout the term. In the experimental classes, this requirement was met in two ways, both of which came at the end of the experimental pedagogical cycle, and as such were not treated very differently from the remainder of listening practice activities. The learners were informed that a grade was being taken, but these assessments were not called tests.

The first method of assessment involved the production of written summaries. Twice during the term the class discussion at the end of the Second Verification Stage (see Table 3) was replaced by individual summary writing, in which learners were requested to include as much logically connected information as they could from the listening. They were permitted to utilize the notes they had made during both the First and Second Verification Stages, to insure that this was an assessment of comprehension rather than memory. The second form of assessment also occurred twice during the term, and involved instruction in strategic listening for the purpose of answering the comprehension questions provided in the course book (Mackey et al., 2007). This was done both to meet curricular requirements that learners be able to respond to questions based on a given discourse (Appendix C), and on learners’ needs to be capable of answering comprehension questions on standardized tests.
This activity was incorporated into the metacognitive pedagogical cycle by inserting an additional prediction phase into the Planning and Predicting Stage (as presented in Table 3) that focused completely on the comprehension questions in the text. Learners first predicted the answers to these questions individually, and then discussed these predictions with a partner. This was followed by a whole group examination of the questions in which the instructor detailed strategic approaches to dealing with certain types of questions, such as those which contain an “all of the above” or “none of the above” option or those containing negative wording in the question stem. During this time, learners were free to alter or adjust their predictions. During the First Verification Stage, learners focused solely on these questions and verified their predictions; they did not note any additional information which they had understood. At the end of the First Verification Stage, participants engaged in a brief discussion of those questions about which they remained unsure (problem identification) and their plans for tackling these problem questions in the Second Verification Stage. In the Second Verification Stage, the learners again focused solely on the comprehension questions, utilizing self-selected strategies such as selective attention or double check monitoring to confirm the answers to the questions. This was followed by a whole group discussion of the correct answers and strategies used to accomplish this task. Misunderstandings were clarified in two ways. Either the learners provided explanations to each other or the instructor would read the transcript at a slow pace after giving instructions to use selective attention to focus on the place where the answer could be found. No change was made to the Third Verification Stage during these lessons.
In the comparison classes, learners’ listening was formally assessed three times using listening tests provided in the teacher’s manual (Mackey et al., 2007). These tests followed the same format as the listening practice activities in the course book, and consisted of 10 multiple-choice questions: one main idea, seven related to details, and two requiring inference. These listening assessments were treated differently from listening practice activities in the following ways:

- Elimination of both pre-listening and post-listening activities.
- Elimination of the Guest Teacher and whole group discussion of correct and incorrect answers (correct answers were written on the test sheet and returned; learners were invited to visit the instructor’s office with any questions or concerns that they might have).

Role of teacher scaffolding and student autonomy

The instructor of the experimental classes slowly removed scaffolding (support) throughout the treatment, guiding students from a high level of teacher support towards increasing autonomy (Graham & Macaro, 2008; Wenden, 1991). In the early stages, the instructor explained which strategies could be used in the various stages of the listening process (before, during, and after listening), modeled those strategies in use, and then invited students to attempt these strategies on their own. However, towards the end of the treatment, the listeners were asked to brainstorm which strategies could be employed at each stage of the process. The instructor proffered questions such as “What should we do now?” and the learners provided the strategies, and therefore the direction, to the final lessons. This was intended to encourage autonomous strategy use (Graham & Macaro,
In the comparison classes, autonomy was fostered in a different way. Rather than slowly removing support, the instructor periodically invited the learners to choose between two or three potential activities. The instructor’s goal in providing autonomy was to foster an atmosphere of freedom and shared responsibility so that the students would feel that they were a part of a learner-centered community of practice.

Data Analysis

Description of the Data

For the purposes of this examination, both experimental cohorts were treated as a single entity, or group, to be compared with the single comparison cohort, or group. The listening comprehension, self-reported metacognitive awareness, and opinions of each of these groups were compared. The construct of listening comprehension was measured with the listening section of the Michigan Placement Test (MPT), which consists of 20 multiple-choice items on a scale from zero to 20. The construct of metacognitive awareness was measured with the Metacognitive Awareness Listening Questionnaire (MALQ), which consists of 21 items on a six-point semantic differential Likert scale (Appendix A). A choice of “1” on this scale corresponds to “strongly agree,” and a choice of “6” represents “strongly disagree.” As a result, lower scores represent higher degrees of self-reported metacognitive awareness. The items on this questionnaire are loaded onto five separate factors, or categories, each of which represents a different feature of metacognition. These categories are: Mental Translation, Person Knowledge, Problem-Solving, Planning and Evaluation, and Directed Attention. The final category, Mental Translation, correlates negatively with metacognitive awareness, so each item in
this category (numbers four, 11, and 18) requires reverse coding to be scored. Learners’ perceptions of their instruction were determined using two separate data collection methods: focus group interviews and an on-line opinion-survey created by the researcher (Appendix B). Each of the aforementioned data collection instruments was employed in the examination of the four research questions which informed this investigation. In the following sections each of those questions and the methods of analyses used with each will be discussed.

*Research Question 1*

*Are there significant differences between the experimental and comparison groups in terms of either listening comprehension or metacognitive awareness after the treatment?*

To answer this question and determine whether there was a significant difference between the two groups’ end-of-treatment listening comprehension abilities, the MPT listening post-tests of the experimental and comparison groups were subjected to an independent samples t-test using the statistical program R. Additionally, in order to ensure that there were no significant differences between the two groups at the beginning of the treatment, the means of the two groups’ MPT pre-test scores were also analyzed with an independent samples t-test.

In order to determine whether there was a significant difference between the two groups’ self-reported metacognitive awareness at the end-of-treatment, respondents’ MALQ total scores were calculated by reversing the scales on those items which had been negatively worded (item numbers three, four, eight, 11, 15, 16, and 18) then adding up each of the individual Likert scale (1-6) responses. The means of the totals of the
experimental group and the comparison group were then analyzed using an independent samples t-test with R. Additionally, each of the individual categories assessed by the MALQ (Mental Translation, Person Knowledge, Problem-Solving, Planning and Evaluation, and Directed Attention) were also examined. Mean scores for each category were calculated, and an independent samples t-test was run on each mean. This was done to more accurately ascertain which elements of metacognitive awareness had been most strongly affected by the treatment. Each of these methods of analysis was also conducted with the two groups’ MALQ pre-test scores to examine the possibility of significant differences between the total mean scores or the category mean scores at the beginning of the treatment.

Research Question 2

Are there significant differences between the experimental and comparison groups’ pre-test and post-test scores of either listening comprehension or metacognitive awareness?

To determine whether or not there was a significant amount of difference between beginning and end of treatment listening comprehension scores, the MPT pre-test and post test mean scores of both the experimental and comparison groups were compared with paired-samples t-tests with R. Additionally, the experimental group’s MPT listening tests were separated from those of the comparison group, and separate paired samples t-tests were run on the pre- and post-tests of each, to examine whether or not there was a statistically significant difference for either group. Finally, gain scores were calculated by subtracting the MPT pre-tests from the post-tests, and the means of these were analyzed using an independent samples t-test with R.
To determine whether the amount of difference in self-reported metacognitive awareness had reached statistical significance for either group, gain scores were calculated by subtracting the pre-test total MALQ scores from the post-test total MALQ scores. The means of the gain scores of the two groups were then compared with independent-samples t-tests using R. Additionally, gain scores for each category assessed by the MALQ were also calculated for each group, and the means of these were also examined with independent sample t-tests.

**Research Question 3**

*Is there a correlation between self-reported metacognitive awareness and listening comprehension?*

Increases in metacognitive awareness are hypothesized to lead to increases in L2 listening improvement (Chamot, 2005; Goh, 2008; Vandergrift, 2002; 2003b; 2004; Vandergrift et al., 2006), and this hypothesis has informed this entire investigation. However, the presence or absence of a statistically significant correlation between listening comprehension as measured by the MPT and metacognitive awareness as measured by the MALQ needs to be empirically examined. In order to investigate whether or not a correlation exists between these two instruments and the constructs they represent, the pre- and post-test MALQ total scores and the pre- and post-test MPT listening scores were analyzed with a two-tailed non-parametric Spearman’s bivariate correlation test using the statistical program R. The scores of the two groups were not separated for the purposes of this particular investigation, as this analysis was intended to
shed light on the presence (or absence) of a statistically significant correlation between the two constructs.

The non-parametric Spearman’s bivariate correlation was utilized due to the relatively small number of participants involved in this aspect of the investigation (n=26). With such a sample size, use of the parametric Pearson’s correlation might have created results which were influenced by unequal variances, non-linearity, outliers, or other elements indicative of an abnormal distribution of data. Spearman’s bivariate correlation, on the other hand, measures the strengths of the associations between the two variables without making any assumptions about a normal distribution of data, and was therefore the more appropriate test to use with a sample of this size. The two-tailed model was employed due to its more conservative nature. Though there was a hypothesis guiding the correlation (that lower MALQ total scores would correlate with higher MPT listening scores), the greater requirements for achieving significance with the two-tailed model would lend more credibility to the findings, should a correlation in fact exist.

**Research Question 4**

What are the participants’ attitudes and opinions regarding each method of instruction, and what are the similarities and differences of the opinions of the members of each group?

All data used in the analysis of this question was analyzed qualitatively by the researcher. Learners in both groups were invited, but not required, to attend focus group interviews. These interviews took place immediately after the learners’ last morning class; the interview of each cohort took place on separate days. These interviews were
audio-recorded and partially transcribed by the researcher, and their purpose was to provide learners with an opportunity to express their thoughts and opinions regarding the instruction they received as well as to determine which elements of their instruction were perceived as salient and memorable for either positive or negative reasons. An on-line survey, designed to gather information about learners’ opinions of both the usefulness of their listening instruction in general and those specific elements deemed to be most helpful, was also employed. This survey was created using the website www.surveygizmo.com (Appendix B) and administered to each of the three cohorts during class at the end of the term. Each question required a response; all responses were tallied and percentages were calculated based upon the total number of respondents or responses in each group.

The survey was designed to insure that each learner was given an opportunity to voice his or her opinions, due to the fact that the focus group interviews might potentially be dominated by those learners with either the strongest personalities or the most proficient English speaking skills. This turned out to be the case, and while the focus group interviews yielded some information regarding learners’ opinions, too many individuals were silent in each group for an accurate picture of what the majority of students perceived to emerge. Therefore, the data generated during the focus group interviews was not used to answer this question, as it appeared to only present the viewpoints of the most proficient and outgoing L2 speakers, though some quotations and general information from those interviews has been included.
CHAPTER 4: RESULTS

The purpose of this study was to compare an experimental method of listening instruction which involved cyclical exposure to and practice with a selection of listening strategies intended to increase learners’ metacognitive awareness of the processes underlying listening comprehension to a method which was practice-based and focused on providing learners with “rich listening experiences of various kinds” (Comparison Group Instructor, March 15, 2011). These two methods were compared in terms of their effect on both listening comprehension abilities as measured by the listening section of the Michigan Placement Test (MPT), self-reported metacognitive awareness as measured by the Metacognitive Awareness Listening Questionnaire (MALQ), and learners’ perceptions of their usefulness in improving listening comprehension, as measured by an on-line survey (Appendix B). Additionally, participants’ self-reported metacognitive awareness scores and listening scores were examined for the presence of a correlation between the two constructs. The results of each of these analyses will be explored in the forthcoming sections, each of which is devoted to one of the constructs under investigation.

Listening Comprehension

As was mentioned previously, the participants’ listening comprehension was measured with the 20 item listening portion of the Michigan Placement Test (MPT) at the beginning and end of the treatment. Table 5 shows the experimental participants’ MPT listening pre- and post-test scores.
Table 5

Experimental Participants’ Pre- and Post-Test MPT Listening Scores

<table>
<thead>
<tr>
<th>Participants</th>
<th>Listening Pre-Test</th>
<th>Listening Post-Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>aP1A</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>P1B</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>P1C</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>P1D</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>P1E</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>P1F</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>P1G</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>bP2A</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>P2B</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>P2C</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>P2D</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>P2E</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>P2F</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>P2G</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>P2H</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>P2I</td>
<td>NA</td>
<td>13</td>
</tr>
<tr>
<td>P2J</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>P2K</td>
<td>12</td>
<td>14</td>
</tr>
</tbody>
</table>

Note. MPT = Michigan Placement Test.

a P1A-P1G=Cohort 1 (n=7). b P2A-P2K=Cohort 2 (n=11).

Table 5 shows that 10 of the experimental participants experienced small (one to three point) improvements in their overall MPT listening scores. Three learners demonstrated no change as measured by the MPT, and three additional learners’ scores dropped by one to three points. Participant P2I, for whom there is no pre-test data, took part in the treatment for the entire 10 weeks and has data included in both the investigation of metacognitive awareness and participants’ perceptions.

Table 6 shows the MPT listening pre- and post-test scores for the comparison group participants.
Table 6

Comparison Participants’ Pre- and Post-Test MPT Listening Scores

<table>
<thead>
<tr>
<th>Participants</th>
<th>Listening Pre-test</th>
<th>Listening Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3A</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>P3B</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>P3C</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>P3D</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>P3E</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>P3F</td>
<td>NA</td>
<td>7</td>
</tr>
<tr>
<td>P3G</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>P3H</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>P3I</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>P3J</td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>

Note. MPT = Michigan Placement Test

Table 6 shows that six out of 10 comparison group participants made improvements of two to three points. As was the case with the experimental participants, three comparison group participants also experienced losses of between one and five points. No learner in the comparison group earned the same score during both the pre-test and the post-test. Participant P3F, it should be noted, took part in the comparison instruction for the entire 10 weeks, but like Participant P2I did not take the MPT at the beginning of the term and could not, therefore, be included in the analysis of listening comprehension scores.

The data in Tables 5 and 6 was examined in four ways. First, the post-test means of the experimental group and the comparison group were compared with an independent-samples t-test to determine if a statistically significant level of difference existed between the post-tests of the two groups. Second, the pre- and post-tests of the two groups were compared with a paired-samples t-test to investigate the presence of a
statistically significant amount of difference between the beginning and end of the treatment. Third, gain scores were calculated and the means of these were compared using an independent samples t-test. Finally, the scores of the experimental and comparison groups were separated from each other and the pre- and post-test scores of each group were analyzed using paired-samples t-tests in order to ascertain whether or not the levels of improvement of the two groups were significantly different from one another. Figure 1 shows the mean scores for the two groups at both the beginning and end of treatment.

![Figure 1. MPT Listening pre-and post-test mean scores.](image)

*Note.* MPT = Michigan Placement Test.
As Figure 1 shows, the two groups began the treatment with very similar mean listening scores. An independent samples t-test run on these pre-test listening scores showed that this was a statistically insignificant amount of difference ($t = -0.5652$, $df = 13.203$, $p = .5814$), indicating that the two groups began the experiment with statistically similar listening comprehension scores. At the end of treatment, an independent samples t-test run on the post-test means also showed a statistically insignificant amount of difference between the two groups ($t = -1.283$, $df = 11.817$, $p = .2241$). Therefore, the experimental treatment did not result in a statistically significant difference between the two groups at the end of the treatment.

Figure 1 illustrates the fact that the experimental group ended the treatment with a larger amount of improvement than the comparison group. This was analyzed in a number of ways to examine its significance. First, a paired samples t-test was run to compare the pre- and post-tests of the two groups combined. This analysis showed that the amount of improvement of the entire participant pool was insignificant ($t = -2.026$, $df = 25$, $p = .05357$). Next, gain scores were calculated by subtracting the pre-tests from the post-tests, and the means of these were analyzed with an independent samples t-test. This analysis showed that the difference in improvement between the two groups was also not great enough to reach statistical significance ($t = -.1013$, $df = 11.415$, $p = .921$). Finally, to examine the fact that the experimental group’s gain was larger than that of the comparison group, the MPT listening tests of the two groups were separated from each other, and each set of pre- and post-test means was examined using paired samples t-tests. Table 7 shows the results of this analysis.
Table 7

*Results of Paired-Samples T-Tests on MPT Listening Pre- and Post-Test Mean Scores*

<table>
<thead>
<tr>
<th>Experimental Status</th>
<th>t-statistic</th>
<th>degree of freedom</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>2.0628</td>
<td>16</td>
<td>.05575</td>
</tr>
<tr>
<td>Comparison</td>
<td>-.8278</td>
<td>8</td>
<td>.4318</td>
</tr>
</tbody>
</table>

*Note.* MPT = Michigan Placement Test.

As Table 7 shows, neither the experimental nor the comparison group’s gains on the MPT listening test reached statistical significance. However, the p-values of the two analyses are much different from one another, with the p-value of the experimental group being much closer to .05 than that of the comparison group. This analysis illustrates a trend that may have reached statistical significance had the sample sizes been larger.

The box-and-whisker plots in Figure 2 provide an alternate look at the distribution of the results on the MPT listening post-tests of the two groups.
Figure 2. Distribution of MPT listening post-test scores.

Note. MPT = Michigan Placement Test.

Figure 2 is interpreted as follows:

- 50% of the participants’ scores are contained within the boxes.
- The lines that cut across the boxes represent the median scores.
- The whiskers represent extreme values that are within 1.5 inter-quartile ranges, meaning that they are extreme, but not extreme enough to be considered outliers.

Figure 2 illustrates the fact that the median score of the experimental group was slightly higher than that of the comparison group, and that 50% of the experimental participants’ scores fell into a range occupied by the top 25% of the comparison group’s scores. The most important information that emerges from Figure 2, however, relates to the whiskers. The comparison group had more extreme scores. In fact, the standard
deviation (SD) of the experimental group was 1.6, but the comparison group had a much larger SD at 3.03. However, this SD was not great enough to signify the presence of outliers, which is also illustrated by the absence of small circles in the space beyond the whiskers. The large standard deviations and extremity of scores present in the data of the comparison group explain, in part, why the improvements of the experimental group were not great enough to reach statistical significance.

Metacognitive Awareness

As has been previously discussed, participants’ metacognitive awareness was measured with a 21 item, six-point semantic differential Likert scale questionnaire called the Metacognitive Awareness Listening Questionnaire (MALQ) (Vandergrift et al., 2006) (Appendix A). Lower total MALQ scores represent higher self-reported metacognitive awareness, as a choice of “1” on the Likert scale corresponds with “strongly agree” whereas a choice of “6” indicates “strongly disagree.” The highest numerical score on this instrument is 126. This would indicate a very low level of self-reported metacognitive awareness. Conversely, the lowest numerical score, 21, would indicate a very high level of self-reported metacognitive awareness. Table 8 displays the experimental participants’ total MALQ scores at the beginning and end of treatment, after reversing the codes for all negatively worded items and all items in the Mental Translation category, which corresponds negatively with metacognitive awareness (Vandergrift et al., 2006).
Table 8

Experimental Participants’ Pre- and Post-Test MALQ Total Scores

<table>
<thead>
<tr>
<th>Participants</th>
<th>MALQ Pre-Test</th>
<th>MALQ Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1A</td>
<td>64</td>
<td>59</td>
</tr>
<tr>
<td>P1B</td>
<td>54</td>
<td>41</td>
</tr>
<tr>
<td>P1C</td>
<td>69</td>
<td>58</td>
</tr>
<tr>
<td>P1D</td>
<td>50</td>
<td>49</td>
</tr>
<tr>
<td>P1E</td>
<td>54</td>
<td>55</td>
</tr>
<tr>
<td>P1F</td>
<td>71</td>
<td>48</td>
</tr>
<tr>
<td>P1G</td>
<td>60</td>
<td>47</td>
</tr>
<tr>
<td>P2A</td>
<td>61</td>
<td>49</td>
</tr>
<tr>
<td>P2B</td>
<td>65</td>
<td>51</td>
</tr>
<tr>
<td>P2C</td>
<td>68</td>
<td>68</td>
</tr>
<tr>
<td>P2D</td>
<td>54</td>
<td>56</td>
</tr>
<tr>
<td>P2E</td>
<td>75</td>
<td>54</td>
</tr>
<tr>
<td>P2F</td>
<td>75</td>
<td>59</td>
</tr>
<tr>
<td>P2G</td>
<td>63</td>
<td>47</td>
</tr>
<tr>
<td>P2H</td>
<td>68</td>
<td>60</td>
</tr>
<tr>
<td>P2I</td>
<td>63</td>
<td>55</td>
</tr>
<tr>
<td>P2J</td>
<td>54</td>
<td>53</td>
</tr>
<tr>
<td>P2K</td>
<td>60</td>
<td>45</td>
</tr>
</tbody>
</table>

Note. MALQ = Metacognitive Awareness Listening Questionnaire.

Table 8 shows that at the beginning of treatment, the experimental participants’ self-reported metacognitive awareness scores existed in a range of 25 points, with the highest score at 75 and the lowest at 50. At the end of the treatment, there was a similar range of 24 points, but overall the scores had dropped (meaning that levels of self-reported metacognitive awareness had risen); the end of treatment high was 68 points and the end of treatment low score was 41. A total of 15 experimental participants (83%) exhibited a growth in metacognitive awareness as measured by the MALQ. Of these, 10 (56% of the total) had decreases in their total MALQ scores of 10 points or more, and two (11% of total) had very large decreases of over 20 points.
Table 9 shows the pre- and post-treatment total MALQ scores (after reverse coding) of the comparison group participants.

Table 9

<table>
<thead>
<tr>
<th>Participants</th>
<th>MALQ Pre-test</th>
<th>MALQ Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3A</td>
<td>64</td>
<td>70</td>
</tr>
<tr>
<td>P3B</td>
<td>67</td>
<td>60</td>
</tr>
<tr>
<td>P3C</td>
<td>67</td>
<td>58</td>
</tr>
<tr>
<td>P3D</td>
<td>70</td>
<td>68</td>
</tr>
<tr>
<td>P3E</td>
<td>56</td>
<td>55</td>
</tr>
<tr>
<td>P3F</td>
<td>67</td>
<td>74</td>
</tr>
<tr>
<td>P3G</td>
<td>70</td>
<td>58</td>
</tr>
<tr>
<td>P3H</td>
<td>51</td>
<td>62</td>
</tr>
<tr>
<td>P3I</td>
<td>67</td>
<td>59</td>
</tr>
<tr>
<td>P3J</td>
<td>65</td>
<td>61</td>
</tr>
</tbody>
</table>

*Note. MALQ = Metacognitive Awareness Listening Questionnaire.*

Table 9 shows that at the beginning of the term, the comparison participants’ self-reported metacognitive awareness scores were in a 19 point range, with the highest score at 70 and the lowest at 51. At the end of the term, there was a slightly narrower range of 16 points, and though individual learners exhibited changes, in the aggregate scores did not change much; the end of term high was still 70 points and the end of term low score was 55. A total of seven comparison participants (70%) exhibited a growth in metacognitive awareness as measured by the MALQ. Of these, only one (10% of the total) had a decrease in the total MALQ score of 10 points or more, and none had decreases of over 20 points.

This data was examined in two ways. First, the post-test means of the experimental group and the comparison group were compared using an independent-
samples t-test to discover whether or not there was a statistically significant level of difference between the post-tests of the two groups. Then the means of both the pre-test and post-test total scores of each group were compared using a paired-samples t-test to investigate whether the differences on the post-test had reached statistical significance. Figure 3 compares the means of two groups’ pre-test and post-test total MALQ scores.

![MALQ pre-and post-test total mean scores](image)

Figure 3. MALQ pre-and post-test total mean scores.

Note. MALQ = Metacognitive Awareness Listening Questionnaire.

Figure 3 clearly demonstrates the fact that the two groups began the treatment with very similar MALQ total mean scores, though the experimental group did self-report slightly higher metacognitive awareness at the beginning of the term. An independent
samples t-test run on these pre-test MALQ total mean scores showed that this was a statistically insignificant amount of difference (t = .669, df = 21.381, p = .5107), meaning that the two groups began the experiment with statistically similar levels of self-reported metacognitive awareness.

As Figure 3 also shows, the mean of the comparison group had dropped slightly at the end of the term (indicating a slight rise in self-reported metacognitive awareness), but the experimental group’s mean had dropped much farther (indicating a much higher increase in self-reported metacognitive awareness). In fact, it is interesting to note that the pre-test means of both groups and the post-test mean of the comparison group all exist within a range of just 1.9 points, whereas the post-test mean of the experimental group is 9.5 points lower than the pre-test mean.

An independent samples t-test run on these post-test mean scores confirmed that the post-test mean of self-reported metacognitive awareness of the experimental group is significantly lower than that of the comparison group (t = 3.8387, df = 19.867, p = .001036*). The effect size (eta-squared = .43) is moderate, meaning that the significance of the difference between the two groups’ post-test scores, or the strength of the relationship between these two scores, is of moderate strength (Cohen, 1988).

Additionally, gain scores for each group were calculated by subtracting the MALQ total pre-test scores from the post-test scores. An independent-samples t-test used to compare the means of these gain scores confirmed that the amount of gain exhibited by the experimental group was also great enough to reach statistical significance (t = 2.5757, df = 18.596, p = .01872*). The effect size, or strength of the
significance, related to this examination (eta-squared = .26) is smaller than the one generated from the analysis of the two groups’ post-test means; however, it can still be considered moderate (Cohen, 1988). The box-and-whisker plots in Figure 4 provide an illustration of the distribution of the two groups’ post-test MALQ total scores.

![Box-and-whisker plots showing distribution of MALQ post-test total scores](image)

**Figure 4.** Distribution of MALQ post-test total scores.

*Note.* MALQ = Metacognitive Awareness Listening Questionnaire.

As Figure 4 shows, the median score of the experimental group was far lower than that of the comparison group. Moreover, there was almost no overlap in the scores of the two groups, with 50% of the experimental participants’ scores falling just below the range of 50% of the scores of the comparison participants.
As the individual items on the MALQ were loaded onto five separate factors, or categories, individual category scores were calculated and analyzed to determine exactly which elements of metacognition were the most powerfully influenced by the treatment. Figure 5 compares the pre-test and post-test mean scores of both groups for each category represented by the MALQ: Mental Translation, Person Knowledge, Problem-Solving, Planning and Evaluation, and Directed Attention (Vandergrift et al., 2006). As was previously mentioned, the Mental Translation category is negatively correlated with metacognitive awareness, so each of the item numbers in this category were reverse coded, and the results displayed in Figure 5 represent the learners’ scores after reverse coding.

![Figure 5. MALQ pre- and post-test category mean scores.](image)

*Note.* MALQ = Metacognitive Awareness Listening Questionnaire.

Horizontal Axis:  1 = Pre-Test; 2 = Post-Test.
An examination of Figure 5 reveals that the comparison participants had higher scores (lower levels of metacognitive awareness) than the experimental participants in the areas of Mental Translation and Person Knowledge at both the beginning and end of term. This means that the comparison group participants were self-reporting stronger beliefs about use of Mental Translation than the experimental participants, for example, selecting “I strongly agree” to such statements as “I translate in my head as I listen” (item number four), and self-reporting stronger beliefs about the difficulty of English listening, such as selecting “I agree” to items like number eight: “It is a challenge for me to understand when I listen in English.”

The experimental group also showed little to no change in these two areas over the course of the 10 weeks. Furthermore, the method of instruction had little to no measurable effect on the comparison group’s levels of self-reported metacognitive awareness in terms of Person Knowledge; however, their levels of self-reported awareness of Mental Translation actually increased from the beginning to the end of term (indicating an increase in either the amount of translation or beliefs about translation while listening.)

The experimental group had higher levels of self-reported metacognitive awareness on those questions targeting Problem-Solving than the comparison group at both the beginning and end of treatment, but both groups’ levels of awareness about Problem-Solving were raised over the course of the 10 weeks. The same is true of the Directed Attention category, but the degree to which the experimental group’s scores
dropped (indicating higher levels of awareness) is more pronounced than that of the comparison group. Finally, in terms of Planning and Evaluation, at the beginning of the term, the comparison participants had higher levels of self-reported metacognitive awareness than the experimental participants, but at the end of the term, the comparison group’s scores had raised slightly (indicating lower levels of metacognitive awareness), but the experimental group’s scores had dropped considerably. In order to determine if any of these changes were statistically significant, independent samples t-tests were run on each category.

First, independent samples t-tests were run on the pre-test mean scores in each category, to determine if there were any significant differences between the two groups at the beginning of the treatment. As was the case with the total MALQ scores, no statistically significant differences were found. The results of these pre-treatment independent samples t-tests are presented in Table 10.

Table 10

*Results of Independent-Samples T-Tests on the Experimental and Comparison Groups’ Pre-Test MALQ Category Mean Scores*

<table>
<thead>
<tr>
<th>MALQ Category</th>
<th>t-statistic</th>
<th>degree of freedom</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental Translation</td>
<td>.818</td>
<td>18.99</td>
<td>.4235</td>
</tr>
<tr>
<td>Person Knowledge</td>
<td>.3015</td>
<td>21.641</td>
<td>.766</td>
</tr>
<tr>
<td>Problem-Solving</td>
<td>.9146</td>
<td>17.912</td>
<td>.3725</td>
</tr>
<tr>
<td>Planning &amp; Evaluation</td>
<td>.9415</td>
<td>21.843</td>
<td>.3568</td>
</tr>
<tr>
<td>Directed Attention</td>
<td>.4957</td>
<td>22.135</td>
<td>.625</td>
</tr>
</tbody>
</table>

*Note. MALQ = Metacognitive Awareness Listening Questionnaire.*
The data in Table 10 confirms that the two groups began the experiment with equal amounts of self-reported metacognitive awareness in the five categories measured by the MALQ. Independent samples t-tests were also run on the post-test mean scores in each category, and statistically significant differences were found between the scores of the two groups in terms of both Directed Attention and Planning and Evaluation. The results of these tests are displayed in Table 11.

Table 11

<table>
<thead>
<tr>
<th>MALQ Category</th>
<th>t-statistic</th>
<th>degree of freedom</th>
<th>p-value</th>
<th>eta-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental Translation</td>
<td>1.8307</td>
<td>23.016</td>
<td>.08012</td>
<td></td>
</tr>
<tr>
<td>Person Knowledge</td>
<td>.4138</td>
<td>19.088</td>
<td>.6837</td>
<td></td>
</tr>
<tr>
<td>Problem-Solving</td>
<td>1.2846</td>
<td>14.606</td>
<td>.2189</td>
<td></td>
</tr>
<tr>
<td>Planning &amp; Evaluation</td>
<td>2.7264</td>
<td>17.499</td>
<td>.01410*</td>
<td>.30</td>
</tr>
<tr>
<td>Directed Attention</td>
<td>2.6094</td>
<td>20.181</td>
<td>.01670*</td>
<td>.25</td>
</tr>
</tbody>
</table>

*Note. MALQ= Metacognitive Awareness Listening Questionnaire. * p < .05.

Table 11 shows that p < .05 in the Planning and Evaluation category and the Directed Attention category, and in both categories the effect sizes are moderate (Cohen, 1988). In the Planning and Evaluation category, the effect size is .30, and in the Directed Attention category, the effect size is .25, meaning that the strength of the relationship between the two variables is of moderate strength. Additional independent-samples t-
tests were run on the two groups’ gain score means in each category (calculated by subtracting the pre-test scores from the post-test scores) in order to determine whether there was a significant amount of difference between the beginning and the end of treatment in any of the five categories. The results of these independent samples t-tests are displayed in Table 12.

Table 12

*Results of Independent-Samples T-Tests on MALQ Category Gain Mean Scores of Both Groups*

<table>
<thead>
<tr>
<th>MALQ Category</th>
<th>t-statistic</th>
<th>degree of freedom</th>
<th>p-value</th>
<th>eta-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental Translation</td>
<td>.8726</td>
<td>15.457</td>
<td>.3962</td>
<td></td>
</tr>
<tr>
<td>Person Knowledge</td>
<td>.1969</td>
<td>19.02</td>
<td>.846</td>
<td></td>
</tr>
<tr>
<td>Problem-Solving</td>
<td>.5441</td>
<td>17.806</td>
<td>.5931</td>
<td></td>
</tr>
<tr>
<td><strong>Planning &amp; Evaluation</strong></td>
<td><strong>3.4386</strong></td>
<td><strong>21.941</strong></td>
<td><strong>.0024</strong>*</td>
<td><strong>.35</strong></td>
</tr>
<tr>
<td>Directed Attention</td>
<td>.9182</td>
<td>21.192</td>
<td>.3689</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* MALQ = Metacognitive Awareness Listening Questionnaire.

* p = < .05.

Table 12 displays the results of the independent samples t-tests run on the gain score means in each category. As Table 12 shows, Planning and Evaluation was the only category in which the amount of difference between the gain score means of the two groups was large enough to reach statistical significance. A look back at Figure 5 reveals that the comparison group also self-reported higher levels of awareness regarding Directed Attention on the post-test than they had on the pre-test, eliminating the
possibility of a significant amount of difference (in awareness) in this category. However, the comparison participants’ self-reported levels of awareness in the Planning and Evaluation category rose slightly on the post-test (indicating a decrease in metacognitive awareness), creating a greater contrast with drop in the scores of the experimental group (indicating an increase in metacognitive awareness).

Correlation between Self-Reported Metacognitive Awareness and Listening Comprehension

In order to investigate whether there is a correlation between self-reported metacognitive awareness as measured by the MALQ and listening comprehension as measured by the MPT, a two-tailed non-parametric Spearman’s bivariate correlation test was run on both sets of data. This test demonstrated that there was no correlation between the MALQ and the MPT pre-tests given at the beginning of the treatment (rho = .027, p = .8919). Figure 6 illustrates the distribution of the scores and the degree of correlation between the two instruments.
Figure 6. Scatter plot of all participants’ MPT Listening and MALQ pre-test scores.

Note. MPT = Michigan Placement Test; MALQ = Metacognitive Awareness Listening Questionnaire.

Each of the dots on the scatter plot in Figure 6 represents one participant. The horizontal axes represent MPT listening scores, and the vertical axes represent MALQ total scores. The line bisecting the chart is the degree to which the two sets of scores are correlated with one another. As Figure 6 clearly shows, the participants’ scores are scattered all over the chart, because some learners had very low MPT scores and low
MALQ scores (indicating high levels of metacognitive awareness), others had high MPT scores and high MALQ scores (indicating low levels of metacognitive awareness), and every possible combination in between. Therefore, the correlation line is nearly horizontal. A perfectly horizontal line would indicate a zero (or nonexistent) correlation, and the line on the scatter plot in Figure 6 is nearly horizontal.

However, there was a significant (though weak) correlation between the post-test scores on the two instruments (rho = .4, p = .03518*); at the end of treatment, lower scores on the MALQ (indicating higher levels of metacognitive awareness) had become statistically significantly correlated with higher scores on the MPT listening test. A scatter plot showing the distribution of the participants’ post-test scores and the strength of the correlation between the two instruments is presented in Figure 7.
Figure 7. Scatter plot of all participants’ MPT Listening and MALQ post-test scores.

Note. MPT = Michigan Placement Test; MALQ = Metacognitive Awareness Listening Questionnaire.

As was the case with Figure 6, each dot represents the scores of one individual; however, participants AO2C and A02G had exactly the same scores on both post-test instruments, so they are represented by a single dot. The most striking difference between the scatter plot of the pre-test scores (Figure 6) and the post-test scores (Figure 7) is angle of the correlation line. Figure 7 clearly demonstrates that the strength of the
correlation was much stronger at the end of the treatment. An examination of the dots reveals that while there are still a few individuals with either high MPT listening scores and relatively high total scores on the MALQ (representing relatively low levels of self-reported metacognitive awareness) or the reverse, the majority of the participants’ scores fall into a more predictable pattern from the upper left hand corner of the box (high MPT and low MALQ) to the lower right hand corner of the box (low MPT and high MALQ). This means that at the end of the treatment, higher levels of self-reported metacognitive awareness had become correlated with higher scores on the MPT listening test. A more in-depth analysis of the individual dots reveals that of the four which do not follow the direction of the correlation line, three belong to comparison group participants. The only experimental group participant whose scores on the two instruments did not correlate was participant P1B, who evidenced very high levels of metacognitive awareness (this participant had a lower MALQ score than anyone else at the end of the treatment) but did not perform particularly well on the MPT listening post-test.

Participants’ Perceptions of Method of Instruction Received

Perceptions of L2 Listening Comprehension Improvement

In order to determine the participant’s perceptions of the instructional methodology to which they were exposed, focus group interviews were conducted with each cohort during the last week of the term, and learners were also asked to complete opinion surveys during the final week (Appendix B). The focus group interviews were largely dominated by those members of each class with the strongest L2 speaking skills, and while they yielded some information about learner preferences, too many students
remained silent in each interview to draw real conclusions about what was or was not considered beneficial by the majority of learners in each group. Therefore, the majority of the data in the present section was taken from the online survey, which was taken (and completed) by 27 participants: experimental (n=18) and comparison (n=9).

The first question on the survey asked the learners to rate their perceived levels of L2 listening improvement. Learners were asked to choose from the following options in answering this question:

- I don’t know.
- It has not improved.
- It has improved a little bit.
- It has improved somewhat (so-so).
- It has improved a good deal.
- It has really improved a lot!

For purposes of comparison, the It has improved a little bit and the It has improved somewhat (so-so) options have been combined into a single category, and the It has improved a good deal and It has really improved a lot! selections have also been combined. Figure 8 compares the percentage of experimental and comparison group participants who selected each option to this question, after this re-categorization.
As Figure 8 shows, none of the participants in either group selected "It has not improved" in answer to this question, indicating that all participants perceived themselves to have made some progress as L2 listeners during the course of the term. One experimental participant selected "I don't know," and the remainder of the students selected one of the four options indicating degrees of improvement. As the chart shows, slightly more of the experimental participants (50%) gave themselves high improvement ratings ("a good deal" or "a lot") than did comparison group participants (44%). Likewise, 56% of the comparison group participants rated their improvement with lower descriptors, "a little bit" or "somewhat." This is a slightly higher percentage than the number of experimental participants (44%) who rated their improvement similarly. Due to the similarity of these numbers, it can be inferred that the experimental pedagogical cycle did not have a considerably effect on the participants’ perceptions of their
improvement than the method of instruction used in the comparison class. These grouped improvement ratings will be returned to throughout the Participants’ Perceptions of Method of Instruction Received section, as interesting information sometimes emerges when the participants are separated in this way.

Independent Listening Practice

Another set of items on the questionnaire asked the learners to self-report on the amount of independent listening practice which they engaged in at both the beginning and the end of the term. The learners’ estimates of their behaviors at the beginning of the term were, therefore, retrospective accounts. Independent practice was defined according to the following parameters: watching movies, watching TV, listening to the radio, talking to Americans, listening to Americans talk to each other, practicing listening on the Internet, etc. Independent practice was investigated due to its importance as a variable in an investigation of L2 listening comprehension abilities. As ESL students living in an English speaking environment, these participants had the opportunity to immerse themselves in the target language during the term, and such decisions might have an impact on the results. However, 74% of the participants (20 students) who took the survey were Chinese speakers and 19% (5 students) were Arabic speakers. Both of these communities are fairly sizeable at the university where this research took place. Therefore, members of both of these native language groups can elect to immerse themselves in their native languages outside of class, and it cannot be assumed that they are immersed in the target language simply by virtue of their presence in an English-speaking country.
Learners in both groups were asked to select from the following set of options for both the beginning and the end of the 10-week term:

- 0—30 minutes a week
- 30 minutes—1 hour a week
- 1 hour—2 hours a week
- 2 or more hours a week

For the purposes of presenting this data, the questionnaire items have been combined, so that the 0—30 minutes a week and 30 minutes—1 hour a week options are now represented by a single 0—1 hour category, and the 1 hour—2 hours a week and 2 or more hours a week options have been combined into a 1—2+ hours category. Table 13 displays the self-reported amounts of independent listening practice at both the beginning and end of the term for both groups of learners, displayed in percent form to equalize the fact that there were twice as many experimental as comparison participants.

Table 13

*Participants’ Self-Reported Amounts of Independent Listening Practice*

<table>
<thead>
<tr>
<th>Experimental Participants</th>
<th>Comparison Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beginning of Term</td>
</tr>
<tr>
<td>0—1 hour</td>
<td>50%</td>
</tr>
<tr>
<td>1—2+ hours</td>
<td>50%</td>
</tr>
</tbody>
</table>
As Table 13 reveals, there was not a very large difference between the two groups’ self-reported amounts of independent listening practice, though at the beginning of the term a somewhat larger percentage of the comparison group participants retrospectively reported to be spending more time on independent listening practice than their experimental counterparts, and by the end of the term a similarly sized proportion of the experimental participants were reporting more independent practice time. This is due to the fact that more experimental participants than comparison participants claimed to have increased the time they spent on independent practice activities. Additionally, when the comparison group participants’ responses are examined in the aggregate there appears to have been no change over the course of the quarter (though individuals did self-report both increases and decreases). Due to the similarities between the two groups’ self-reported amounts of independent practice, it can be inferred that differences in independent practice time were not great enough to have influenced the results.

**Participants’ Perceptions of Instructional Elements**

The primary purpose of the questionnaire was to determine which elements of instruction influenced the learners’ perceived levels of improvement. Three separate questions were used to acquire this information, each of which had a slightly different aim. The first such question asked those learners who had rated themselves as having improved their L2 listening comprehension skills to select those elements of listening practice which had had the greatest impact on this improvement from a predetermined list. This item was followed by an open-ended question inviting them to elucidate upon their answer choices. As this question was only available to those learners who rated
their listening as having improved, another question was included which asked all learners taking the questionnaire to select those elements of instruction which had been the most beneficial for improving their L2 listening skills, again from a predetermined list. Finally, the participants were presented with an open-ended question asking them to discuss their favorite and least favorite instructional activities. The participants’ responses to each of these questions will now be examined in separate subsections.

Why do you think your listening ability improved?

As there was only one participant who did not perceive him/herself to have made improvement, (selecting “I don’t know”), this question was completed by the remaining 26 participants taking the survey. These learners were given the following four options to choose from, and were asked to select all of those which they believed had impacted their level of L2 listening improvement:

- I practiced listening on my own outside of class.
- I practiced listening in homework activities outside of class.
- I practiced listening during in-class activities.
- I learned useful things about how to listen more effectively (better).

Figure 9 shows the distribution of responses of the participants in each group. Due to the fact that the participants could select as many of the aforementioned options as they believed to be applicable, the bars in Figure 9 are presented in percent form, according to the number of individual learners within the group who selected a given option. For example, *I practiced listening on my own outside of class* was selected by
eight out of 17 (47%) of the experimental participants, and three out of nine (33%) of the comparison participants.

Figure 9. Elements participants perceived to have influenced their L2 listening improvement.

Perhaps the most striking feature of Figure 9 is the data regarding the option *I practiced listening on my own outside of class*. As the graph displays, this item was less frequently selected than any other. 53% of the experimental participants and 67% of the comparison participants did not perceive independent listening practice to have been instrumental in their L2 listening improvement. This lends support to the previous postulation that independent listening practice was not a variable which could influence the results; it is a negative trend nonetheless.
Figure 9 also provides some illuminating information regarding learners’ beliefs about in-class activities. The only category selected by approximately the same proportion of learners was *I practiced listening during in-class activities*, suggesting that a similar proportion of learners in both groups felt that the activities they engaged in during their Listening & Speaking classes were beneficial. However, a comparison of the number of learners who selected *I learned useful things about how to listen more effectively (better)* is telling. 14 out of 17 (82%) of the experimental participants chose this option, whereas five out of nine (56%) of the comparison participants made the same selection. This suggests that while the learners were appreciative of the activities in both classes, the explicit strategy instruction and cyclical process-based approach of the experimental classes was perceived by more learners to be more beneficial than the “sporadic” and “casual” (Comparison Group Instructor, personal communication, March 15, 2011) approach to strategy instruction which took place in the comparison class. Additionally, this was the most frequently selected item by experimental group participants, suggesting that direct strategy instruction made this element more salient than the others to the learners in the experimental group. Finally, separation of the two groups of learners into those who perceived themselves to have made “a good deal” or “a lot” of improvement and those who perceived themselves to have made “somewhat” or “a little bit” of improvement yields another interesting perspective. 100% of the experimental participants who rated their level improvement with the descriptors “a good deal” or “a lot” selected this option, compared with 67% of the comparison participants who chose the same descriptors.
It is worth noting that a slightly greater percentage of comparison participants, whose class was practice-based, selected *I practiced listening during in-class activities* (67%) than *I learned useful things about how to listen more effectively (better)* (56%), whereas the experimental participants, who took part in a process-based class which focused on listening strategically to gain control of the listening process, selected *I learned useful things about how to listen more effectively (better)* more often (82%) than the in-class practice option (65%). This may simply be a reflection of the fact that the learners were selecting those instructional elements which were given weighted importance during the term, or it may be due to the fact that instruction which empowers learners with information about how to approach listening tasks is appreciated. Considering the fact that 100% of those experimental participants who rated their improvement as being “a good deal” or “a lot” selected this option, and the fact that it was selected by more experimental participants than any other choice lends credence to the second possibility.

*Why was this/were these helpful?*

As has been mentioned, the previous item was followed by an open-ended question in which learners were asked to explain why the aforementioned elements had helped them to improve. These responses were coded by the researcher, and the following categories emerged (presented here along with a selection of the words and phrases used to code these categories):

- *Practice*: “practice,” “practiced,” & “speak and speak everyday”
- **Homework**: “speaking journal,” “listening and speaking journal,” “homework,” & “audio journal”
- **Vocabulary**: “more vocabulary,” “many words,” & “some words”
- **Strategies**: “strategy,” “strategies,” “the information…help [me] to know how I can improve,” & “listening skills”
- **Fast Listening**: “fast listening”
- **None**: “It’s helpful my English better,” “I have a better listen after I study in this class,” & “The listening activities help me improve my English more”
- **Negative Response**: “I believe the English class doesn’t help more than 20-30%”

Each category represents a component of instruction specifically mentioned in an open-ended response. The majority of learner comments focused on one item, and therefore resulted in a single code. However, a few learners wrote lengthy answers containing information that fell into multiple categories, and the *None* category was created to encompass those responses which were too vague to be properly categorized. Each time a response fell into one of the categories, it was tallied by the researcher, and then the totals for each category were divided by the total number of responses in order to determine percents. The categorized responses of both the experimental and comparison group participants are presented in Figure 10.
Figure 10. Self-reported reasons why listening instruction had a positive effect on L2 listening improvement.

As Figure 10 shows, Practice and Strategies were the most commonly mentioned responses to this question made by experimental learners, and were mentioned the same number of times (seven times each). Conversely, Practice and Fast Listening were the most commonly mentioned items by comparison group learners (three times each). This is interesting, because in-class listening practice in the experimental classes invariably involved work with listening strategies, but in-class listening practice in the comparison class was made up of several components, of which Fast Listening was only one. Therefore it can be inferred that this was perceived to be more helpful than the other types of listening practice which the class engaged in.
It is also interesting to note that *Fast Listening* was an activity engaged in exclusively in the comparison group classes (explaining why no experimental participants mentioned it), but strategy instruction was included in both classes. The nature of that instruction was very different, however. In the experimental classes strategy instruction was explicit, cyclical, and the primary focus of listening practice. In the comparison classes, strategy instruction was not the primary focus of listening practice. Instead, the goal of the instruction was to provide practice with a variety of different types of input. However, three strategies were explicitly discussed, though these discussions were intermittently scattered throughout the course of the term, and several other strategies were practiced implicitly (without mention). The responses of the learners would suggest that the approach toward strategy instruction utilized in the experimental classes makes strategies appear to be more beneficial to learners. Finally, the single *Negative Comment* made by a comparison group participant warrants examination. This learner said “I believe the English class doesn’t help more than 20-30%.” While this is only one student’s opinion, and goes against the grain of the remainder of his classmates’ opinions about the helpfulness of the class (in the focus group interview, these learners made it clear that they had enjoyed their class and found many things about it beneficial), it remains informative. Perhaps this learner was one of the 33% of comparison group participants who both practiced listening independently and found it helpful for the improvement of his/her L2 listening comprehension skills. Such an individual might have felt that a practice-based listening class only offered more of the same types of listening practice that s/he could access independently.
Which of the following in-class or homework activities was the most helpful for improving your listening skills?

The next questionnaire item was similar to the one which was previously discussed in that it asked the participants to think critically about which elements of instruction had been the most beneficial for improving L2 listening skills. The two primary differences between this question and the previous are the fact that this question was required of all participants taking the survey (n=27) rather than just those who had indicated improvement on the first question, and the items from which the learners could select on the present item focused on much more specific components of instruction.

Participants were given a list of eight activities engaged in by both classes (as well as an open-ended “other” option), and were asked to select all of those which they had done during the term. The purpose of this item was to provide the participants with a finite set of choices to consider in the subsequent item, which asked them to pinpoint which of those elements had been the most useful in the improvement of their L2 listening skills. The items from which learners could select were:

- Practiced listening to audio CDs / audio recordings on the internet
- Practiced listening to videos
- Practiced listening to my classmates
- Practiced listening to audio/video and answered questions about what I heard
- Practiced listening to audio/video and wrote summaries about what I heard
- Practiced listening to audio/video and talked about what I heard
• Practiced listening skills (strategies) like guessing (inferencing), making predictions, using what I already know (activating background knowledge), etc.

• Other

Each of these items was carefully designed in order to avoid bias. First, every major activity used in both classes is represented on this list, and all but one (summary writing) was engaged in by both classes, albeit in different ways. For example, in the experimental classes, learners often watched videos for homework and had to either summarize the main ideas or answer open-ended questions about them, whereas in the comparison class, the learners once watched an entire full-length movie as a part of their in-class listening practice. Second, each item includes terminology intended to be recognizable to both groups of learners. This is most relevant on the final item, because the experimental learners were exposed to a much larger amount of meta-language than the comparison group learners. Therefore, each of the strategies selected for inclusion on this item were introduced and practiced in the comparison class, and the terminology used by both instructors was provided. Learners were asked to select a single item that was the most beneficial, but the question was open-ended, allowing them the freedom to make multiple selections if they so wished. 21 of the 27 participants taking this questionnaire pin-pointed one thing as being the most beneficial element of the class, three of the participants selected two items, one participant chose six out of the seven, and two typed in the words “all things” and “all of them.” Participants’ responses were converted to percents by dividing the number of times a given option was selected by the
total number of participants in each group. This information is graphically depicted in Figure 11.

\[\text{Figure 11.} \text{ Specific instructional components perceived to be the most beneficial for improving L2 listening.}\]

Figure 11 yields some interesting information regarding the ways in which instructional activities were perceived by the learners. First, as the two pie charts show, learners in each group had a wide variety of opinions about that which was the most helpful element of the class. In the comparison group, only one individual element was selected by three individuals (\textit{Practiced listening to audio/video and talked about what I heard}, represented by the label “Talked about Selections” in Figure 11), and of the seven options, four were selected by only one or two learners. Learners in the experimental
group also had a wide variety of opinions, with six of the seven elements having been selected. It can thus be inferred that different learners respond differently to different types of tasks and have very different perceptions from one another regarding best methods. Therefore, the incorporation of a variety of different types of activities is important to insure that learners perceive to be benefitting from instruction.

Second, despite the variety of preferences evidenced in Figure 11, support of the seventh option, *Practiced listening skills (strategies) like guessing (inferencing), making predictions, using what I already know (activating background knowledge), etc.* by learners in the experimental group remains apparent. This choice was selected by 10 of the experimental participants. It was selected five more times than the next most popular choice of the experimental cohorts (*Practiced listening to audio/video and talked about what I heard*), and was the most frequently selected item overall. Support of this item is even more apparent amongst those learners who rated their improvement with the descriptors “a good deal” or “a lot.” Eight out of nine (89%) of those learners selected this item as having been the most helpful element of the class. In contrast, this item was not selected by any members of the comparison class.

Finally, a look at the learners’ selections reveals that learners in each group tended to select those items which were most commonly utilized in their classes. The *Practiced Listening to Videos* option, for example, was selected by 15% of the experimental participants, but by none of the comparison group participants. As was previously discussed, internet-based videos were a regular component of the experimental group’s required homework assignments, whereas the comparison class
watched a full-length film one time as a method of in-class listening practice. Considering this, it is not at all surprising that the learners’ selections of this option are distributed as they are. The same is true of the Practiced Listening to my Classmates option, which was selected by 11% of the comparison group participants, for whom structured interactional conversation practice was a regular component of instruction, but by none of the experimental participants, for whom listening to classmates was rarely isolated as an activity unto itself. This being the case, it is notable that Practiced Listening to Audio/Video and Talked about What I Heard was the most frequently selected option by comparison group participants (34%) and the second most frequently made selection by experimental participants (19%). The fact that more learners (in both groups) professed to prefer discussion of listening selections over answering questions or writing summaries about listening selections is informative, suggesting that learners enjoy collaborating and working together to make sense of L2 input.

*What were your favorite in-class (or homework) listening practice activities? Why?*

The final set of questionnaire items asked participants to discuss their favorite and least favorite instructional activities. The word *favorite* was selected so that they would feel free to write about those activities which they had really enjoyed or not enjoyed rather than only discussing those elements which they considered beneficial. Naturally, some learners enjoy those activities which are useful for improving their skills, but others may enjoy those things which are entertaining, easy, or informative, and this is reflected in the learners’ responses. Additionally, these questions were open-ended, so participants
were free to write about anything they wished instead of being required to select from a predetermined set of options.

Responses were coded, tallied, and categorized by the researcher. A large number of categories emerged; therefore, each of the categories containing only one learner’s response were combined into a single Other category. Additionally, a None category was created for those responses which did not answer the question or were too vague to be categorized. These categories, along with a selection of the words and phrases used by the researcher in the coding process, are as follows:

- **Video-Homework:** “video,” “listening and speaking homework,” & “listening homework”
- **Fast Listening:** “fast listen,” “fast listening,” & “listening fast”
- **Strategy Instruction:** “strategy,” “strategies,” “making predictions,” & “check my prediction”
- **In-class Listening Practice:** “in class listening practice,” “listening in the class,” & “making practice listening in the class”
- **Talking to Classmates:** “whit my classmates talk!” & “conversation”
- **Audio Journal Homework:** “listen for audio and toke about what do you hear…,” & “audio journal”
- **Other:** “Play game in class,” “practice vocabulary,” “distinguishing between similar pronunciation,” “listen and check answer,” & “speaking journal”
- **None:** “[j]” & “speaking in some topic”
The number of learner comments in each category was converted to percent by dividing the number of responses in a given category by the total number of categorized responses. Therefore, if a single learner mentioned two specific activities in his/her response, this was counted as two responses. This information is represented graphically in Figure 12.

![Figure 12](image-url)

*Figure 12. Participants’ favorite instructional activities.*

The most important conclusion to be drawn from the data in Figure 12 relates to those two categories which contained the largest number of learner comments. As was mentioned in the Procedures section of the Methodology Chapter, learners in the experimental classes watched short, authentic videos as a part of their weekly homework requirements. These videos were generally two to four minutes long, and were always
related to the content of the in-class listening selections. After watching the videos, the learners were tasked with either summarizing what they had understood or answering a set of comprehension questions which had been written and delivered orally (via a video on the CMS) by the instructor. The experimental participants responded very favorably to this element of the class, and their positive responses can be interpreted in one of three ways. Perhaps they enjoyed utilizing the visual channel to aid them in the processing of rapid speech. Alternatively, they may have appreciated the fact that these were authentic selections, as was likely the case with the learner who said “Listen to the new video. It makes me know more information.” It is also possible that they liked having the ability to watch and re-watch the videos until they had comprehended the main ideas, as one participant clearly explained “let me listened it again and again, because I must finish my homework, so I must understand what the people said in the video.” Unfortunately, the majority of learner comments do not reveal the reasons underlying the selection, so no conclusions can be drawn as to why the learners enjoyed this aspect of their class so much.

Also of interest is the category that received the largest number of learner comments in the comparison group: Fast Listening. As was explained in the Procedures section of the Methodology Chapter, the Fast Listening technique was utilized by the comparison group instructor periodically throughout the term. The learners listened to very short (30 second to one minute) authentic, real-time selections (NPR Morning Edition headlines) and were asked to make notes regarding any words or phrases which they could parse from the rapid speech stream. Then they would share these bits of
comprehended information with the group while the instructor recorded these words and phrases on the board. The class would then repeat this procedure. After they had shared what they could understand a second time, the learners broke into small groups and were tasked with developing oral summaries explaining what they had just heard.

This procedure shares the following important characteristics with the procedure utilized in the experimental classes:

- Noting what could be understood during the first exposure
- Utilizing a second exposure to verify what had been understood the first time and note additional information
- Collaborating with partners to reconstruct the text
- The purpose of listening is comprehension of main ideas, details and practice making inferences, rather than the correct or incorrect answering of multiple-choice comprehension questions
- Listening practice is separated from listening evaluation.

A more in-depth discussion of the similarities and differences between the experimental cycle and *Fast Listening* will take place in the upcoming Discussion Chapter, but the most important point to keep in consideration is that *Fast Listening* is similar to the experimental cycle in several important ways. Therefore, comparison group support for this activity can be considered as support for these elements of the experimental cycle as well.

The fact that *Fast Listening* was also the most frequently mentioned activity mentioned by comparison group participants in response to the open-ended question
regarding why the class had helped them improve their L2 listening skills (see Figure 10) bears repeating, because it evidences the fact that this was not only a fun activity but also perceived to be very beneficial. It can thus be inferred that exposure to authentic materials, working together to compare ideas, and creating oral reconstructions of listening texts are enjoyable and valuable ways to approach L2 listening instruction.

The final aspect of Figure 12 which will be discussed relates to the second and third most commonly mentioned categories of the experimental group participants: Strategies (24%) and In-Class Listening Practice (14%), because both of these categories are related to the experimental pedagogical cycle under investigation. All in-class listening practice activities in the experimental classes were carried out using the steps described in the Procedures section of the Methodology Chapter (see Table 3). Additionally, all in-class listening practice involved strategy use. Therefore these two categories are actually related to each other and could be combined into a greater category such as Experimental Cycle, which would then contain 38% of the total comments. Therefore, it can be said that 38% of the experimental participants really enjoyed the experimental pedagogical cycle, but 47% enjoyed other elements of the instruction more. It is interesting to contrast this with Fast Listening, which was considered the most enjoyable component of the class by 43% of the comparison participants, whereas 57% enjoyed other elements more.

*What were your least favorite in-class (or homework) listening practice activities? Why?*

The final questionnaire item to be analyzed asked the participants to discuss their least favorite instructional activities. As was the case with the favorite activities
question, learners were free to write anything they wished and were not asked to think critically about whether or not the activity had been beneficial or not, so their responses only indicate the fact that they did not enjoy a particular activity. Every written response fell into a single category, meaning that each response focused on an individual element of instruction. However, a None category was created to encompass those responses which lacked clarity or evidenced misunderstanding of the question. The categories, along with a selection of the words and phrases used in the coding process, are as follows:

- **Video Homework**: “the listening and speaking journal,” “video homework,” “speaking homework,” & “record video”
- **No Least Favorite**: “everything is good,” “i liked everything,” “i think no one is least,” & “there are not what I were least favorite”
- **Tests**: “test,” & “quiz listening”
- **Summary Writing**: “take the summary,” & “to write”
- **Strategy Instruction**: “strategy”
- **Vocabulary (too easy)**: “vocabulary, because, sometime I think it is easy”
- **Talking with Classmates**: “coumumate with my classmate”
- **Oral Reading**: “read the summary”
- **Too Much Homework**: “the homework was more 4 audio every week that made me confuse”
- **Listening Practice**: “listening practice…is not enough to improve me”
• *None*: “I think homework very important,” & “fast listening! very helpful!”

Each of these categories was converted into a percent by dividing the number of comments in the category by the total number of comments made by the group. This information is displayed in Table 14.

Table 14

*Participants’ Least Favorite Instructional Activities*

<table>
<thead>
<tr>
<th>Category</th>
<th>% of Comments</th>
<th>Category</th>
<th>% of Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Least Favorite</td>
<td>35%</td>
<td>No Least Favorite</td>
<td>11%</td>
</tr>
<tr>
<td>Video Homework</td>
<td>35%</td>
<td>Homework (too much)</td>
<td>11%</td>
</tr>
<tr>
<td>Summary Writing</td>
<td>12%</td>
<td>Tests</td>
<td>33%</td>
</tr>
<tr>
<td>None</td>
<td>12%</td>
<td>None</td>
<td>11%</td>
</tr>
<tr>
<td>Vocabulary (too easy)</td>
<td>6%</td>
<td>Oral Reading</td>
<td>11%</td>
</tr>
<tr>
<td>Strategy Instruction</td>
<td>6%</td>
<td>Listening Practice</td>
<td>11%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Talking to Classmates</td>
<td>11%</td>
</tr>
</tbody>
</table>

As Table 14 illustrates, five categories emerged from the responses of the experimental participants regarding their least favorite activities, and seven emerged from the responses of the comparison group participants. Two of the experimental participants’ least favorite elements, *Strategy Instruction* and *Summary Writing*, relate directly to the experimental methodology under examination and as such will be first to
be discussed. One learner indicated that *Strategy Instruction* was his/her least favorite element of the class by writing “strategy” in response to this question. It is interesting to note that this is the same learner who selected “I don’t know” in response to the first question on the survey regarding perceived level of improvement. There may be a link between this learner’s unfavorable attitude toward the metacognitive strategy training, which placed a heavy focus on reflection and self-evaluation, and his/her inability or unwillingness to make a self-evaluation of his/her progress as an L2 listener. Regardless, this learner’s response is telling, because it evidences the fact that there are students who do not appreciate instruction of this nature. However, it is also interesting to note that one of the learners in the comparison class claimed that *Listening Practice* was his/her least favorite aspect of the class, stating “listening practice is least favorite because sometime it is not enough to improve me.” This was not the same learner who claimed that “the English class doesn't help more than 20—30%,” evidencing that two comparison group participants (22%) believed that this practice-based instruction did not have a considerable impact on their L2 listening improvement.

Two experimental learners indicated that *Summary Writing* was their least favorite instructional activity. Summarization was presented as a cognitive strategy intended to help learners logically connect comprehended details into a cohesive narrative, and was practiced both orally and in written form. The primary purpose of the written summaries was to assess the learners’ listening abilities without straying too far from the non-evaluative instructional methodology under examination, as was stated in the Procedures section of the Methodology Chapter (see Table 3). However, the learners
practiced summarizing orally at the end of every cycle, as a means of reconstructing the selections. The learners did not express negative opinions about the oral summarization, which is understandable considering the fact that this was a Listening & Speaking class. It can be assumed that it was the writing element which these participants did not like, and one of the two comments, “I don’t want to write too much on listening class,” supports this interpretation. It is worth noting that while these two learners did not enjoy the implementation of the summarization strategy, two other learners identified summary writing as the most helpful element of instruction (Figure 11).

It is interesting to contrast the experimental participants’ perceptions of this form of assessment with the opinions of the comparison participants. As Table 14 shows, Tests were the most frequently cited least favorite element of the comparison group. As was stated in the Procedures section of the Methodology Chapter, in the comparison class listening was formally assessed three times during the term with multiple choice listening assessments taken from the teacher’s manual of the course book (Mackey et al., 2007). One of the learners who indicated that these tests were his/her least favorite element of the class wrote “I don’t like test. Because test makes me nervous.” In contrast with the mixed reviews of summary writing present amongst the experimental learners, no one in the comparison classes indicated that the listening tests were a positive feature of instruction.

The other features of Table 14 which will be discussed are the two categories into which the majority of experimental participants’ comments fell: No Least Favorite and
Video Homework. The experimental participants’ reasons for selecting Video Homework were varied enough to warrant sub-categorization into the following:

- Internet Speed: “the slowly internet!!” (one learner)
- Speaker Accent: “…the have different accent” (one learner)
- Self-Recording of Videos: “record video” & “do some speaking homework” (two learners)
- General: “…practiced to listening to the video and make a summary or talk about what I heard” & “I don’t like do the video homework” (two learners)

Considering the diverse array of reasons behind the learners’ selection of Video Homework, the No Least Favorite category is actually the most common response to this question. However, both of these are unrelated to the experimental methodology. In fact, examination of the experimental participants’ responses in the aggregate reveals that 83% either fell into categories which were unrelated to the experimental methodology.

When the experimental learners’ responses to all of the questions on the survey are considered together, the following information becomes apparent:

- Elements of the experimental methodology were the most frequently mentioned reasons cited for making L2 listening improvement by those learners who perceived themselves to have improved.
- 66% of the responses regarding the most beneficial element of instruction were directly related to the experimental methodology.
• 38% of these learners’ favorite activities were related to the experimental methodology; 33% involved authentic listening selections (videos).

• 83% of responses about the least favorite element of instruction were unrelated to the experimental methodology.

When the comparison group participants’ responses to all of these questions are examined in the aggregate, the following facts emerge:

• Practice and Fast Listening were the most frequently cited reasons for making L2 listening improvement by those learners who perceived themselves to have improved.

• 34% of the responses regarding the most beneficial elements of instruction related to discussion of listening selections.

• 43% of these learners claimed that Fast Listening was their favorite activity.

• 33% of these learners cited listening tests as being their least favorite activity.

• 22% of these learners made comments which indicated that their instruction was not helpful for improving L2 listening comprehension.

As positive references to strategy-based listening instruction appeared in several of the experimental participants’ comments, and Fast Listening was commonly featured in the comparison participants’ responses, a final method of qualitative analysis was conducted to discern the exact percentage of learners in each group who advocated for these two approaches. The entire questionnaire was examined for instances of positive
comments regarding either strategy instruction or Fast Listening, regardless of whether the comment was in reference to the question about most beneficial or favorite instructional activity. The results of this analysis showed that 10 out of 18 (56%) of the experimental participants expressed strong support of strategy instruction, while the remaining eight students (44%) claimed that other instructional elements were either more enjoyable or more beneficial for the improvement of their L2 listening skills. Four out of these eight students made comments which indicated that other elements of the pedagogical cycle (such as summarizing or listening and then discussing the content with partners) were either the most beneficial or most enjoyable aspects of the class. The remaining four experimental participants did not indicate that strategy instruction or other elements of the cycle were enjoyable or beneficial, and one of these stated that strategy instruction was his/her least favorite thing about the class. Therefore, it can be concluded that 56% of the experimental participants had positive attitudes about the strategy training, and 78% of the experimental participants had positive attitudes about various aspects of the experimental pedagogical cycle. In the comparison group, Fast Listening was specifically mentioned in the comments of seven out of the nine (78%) comparison participants taking the survey. Of the two students who did not specifically mention Fast Listening, one perceived Conversation Starters to be more beneficial/enjoyable, and one indicated that audio journal homework was the most enjoyable/helpful element of the class.
Each of the results enumerated in this chapter will be considered in more depth in the upcoming Discussion Chapter. Pedagogical implications and suggestions for future research will also be discussed.
CHAPTER 5: DISCUSSION

This study investigated the effects of a pedagogical cycle intended to raise learners’ awareness of the metacognitive processes underlying successful listening comprehension. Instruction involved a combination of explicit strategy instruction, in which the learners (n=18) were informed of the names, values, and potential applications of each metacognitive, cognitive, and socio-affective strategy selected for inclusion (Chamot et al., 1999; Wenden, 1991). Learners were given repeated exposure to and practice with these strategies in a cyclical and semi-repetitive format hypothesized to heighten their sensitivity to the phases of the listening process (pre-listening, during listening, and post-listening) (Field, 2000; Goh, 2008; Vandergrift 2002; 2003b; 2004; Vandergrift & Tafaghodtari, 2010) and were given ample practice with clusters of strategies employable during each stage (Brown & Palincsar, 1982; Graham & Macaro, 2008; Macaro, 2006; Purpura, 1997; Thompson & Rubin, 1996). Additionally, bottom-up processing practice (Buck, 2001; Carrier, 2003; Dunkel, 1991; Fender, 2003; Goh & Taib, 2006; Macaro, 2006) and ample opportunities to reflect upon the listening process, L2 listening progress, and strategy use were incorporated (Goh, 1997; Graham & Macaro, 2008; Vandergrift & Tafaghodtari, 2010).

An alternate methodology was also examined for the purpose of comparison. It provided learners (n=10) with large amounts of listening practice and utilized three separate types of listening activities (one of which was similar to the methodology under investigation). The alternate methodology, rather than focusing on metacognition or explicit strategy training, was intended to provide learners with a wide variety of listening
experiences with various types of L2 input, some of which were designed to mirror real-world listening experiences by removing the pre-listening phase. These two pedagogical approaches were analyzed to determine their effect on three constructs: listening comprehension, self-reported metacognitive awareness, and learner’s perceptions of value and effectiveness. The construct of listening comprehension was measured with the 20-item, multiple-choice listening portion of the Michigan Placement Test (MPT), which is used by the IEP where the study took place at the beginning and end of every term for placement purposes. The construct of self-reported metacognitive awareness was measured with the Metacognitive Awareness Listening Questionnaire (MALQ) (Vandergrift et al., 2006), a 21-item, six-point semantic differential Likert scale instrument with answer choices ranging from “1” (strongly agree) to “6” (strongly disagree) (Appendix A). Additionally, the presence of a correlation between the two instruments was investigated. Finally, learners’ opinions were gauged through the use of an online survey, created by the researcher specifically for this group of participants in this context (Appendix B) and focus group interviews. The results of each of these investigations and the resultant pedagogical implications and avenues for future research will now be discussed in separate sections devoted to each investigation.

Listening Comprehension

Of the three constructs under examination, the effect of the experimental pedagogical cycle on the learners’ listening comprehension is without a doubt the most important. The primary purpose of the investigation was to investigate whether or not this cycle is a more effective way to raise learners’ listening comprehension abilities than
other methods, as measured by the listening portion of the Michigan Placement Test (MPT). The MPT, rather than a free-recall or other form of specialized assessment, was selected for a number of reasons. First, though the majority of similar strategic investigations either did not attempt to measure the construct of listening with standardized tests (Carrier, 2003; Cross, 2009; Graham & Macaro, 2008), incorporated them but failed to reach statistical significance on the standardized instrument (Thompson & Rubin, 1996), or did not analyze the results on the standardized assessment with statistical tests (Goh & Taib, 2006), the most recent investigation of this nature by Vandergrift and Tafaghodtari (2010) used a standardized listening assessment and succeeded in reaching statistical significance. The present study was a quasi-replication of Vandergrift and Tafaghodtari’s experimental design. Therefore, in addition to replicating the essential features of their experimental treatment, their methods of assessment were also replicated in an attempt to replicate their findings. Second, though standardized listening assessments have been found to be imperfect tools for measuring L2 listening comprehension (Yi’an, 1998), they are relied upon by educational institutions and influence the development of instructional materials. Furthermore, significant results on standardized instruments carry much more weight than those achieved by other means, so an empirical investigation achieving significant results on a standardized assessment will have a greater impact on the field than one measuring listening comprehension with a test designed by the researcher for the purpose of the study. Third, as was already stated, the MPT is given at the beginning and end of every term by the IEP where this study took place. Therefore, utilization of the MPT
minimized inconvenience to the participants and afforded them two additional days of instruction, as any other type of test would have to have been given during class time.

The results of this aspect of the investigation were informative. There are several possible explanations for the fact that the experimental participants’ MPT listening scores did not reach statistical significance. First, it is entirely possible that this methodology was not superior to the alternate methodology utilized in the comparison classes, at least in terms of the methodologies’ abilities to help learners on the MPT. Moreover, the slightly larger amount of gain evidenced by the experimental participants could have been due to chance. Vandergrift and Tafaghodtari (2010) controlled the instruction of their comparison classes in a number of ways, including not allowing the comparison participants to discuss comprehension with classmates or to predict the content of the upcoming selection. The comparison group in the present study, on the other hand, was not controlled by the researcher, and the participants in that group not only were invited to predict but also regularly engaged in collaborative discussions with listening partners. While the experimental methodology may be superior to one in which its key elements are removed, it may not be superior to alternative methods which are typically employed in ESL/EFL contexts, but do not focus on metacognitive awareness raising or explicit strategy training.

Second, it is also possible that the methodology has merit, but that 10 weeks of exposure, and seven cycles, was not enough to truly sensitize learners to the metacognitive processes underlying listening comprehension. In Vandergrift and Tafaghodtari’s (2010) investigation, the learners were involved in the training for 13
weeks, and went through the pedagogical cycle 13 times. Those five extra exposures may have been pivotal. Additionally, Vandergrift and Tafaghodtari’s participants were Canadian nationals, and the participants in the present study were Chinese, Saudi Arabian, South Korean, and Russian. Culture may have an influence on metacognitive strategy training and metacognitive awareness raising. It may take longer to impact the strategic listening behaviors of those from cultures in which self-directed learning, reflection, and self-evaluation are not a part of standard K-12 educational practices. In fact, one learner’s comment, taken from the reflective video journals, was telling. In response to a question regarding the usefulness of self-evaluation, Participant P1E said “self-evaluation is not helpful because it’s not real.” This response indicates an underlying belief that only the evaluation of the instructor has merit, and a belief of this nature could have a considerable impact on the effect of the training. It would be interesting for future researchers to isolate the effect of the training on a single, homogenous population or to utilize a sample large enough to separate the results of those of different cultural backgrounds. Moreover, when implementing a treatment such as the one investigated in this study with non-Western participants, a much longer treatment period or a treatment which spans the entire curriculum might be required to truly effect change and to help the participants move from simply having declarative knowledge of positive listening behaviors (as measured by the MALQ) to being able to implement procedural knowledge (as measured on listening assessments).

Third, there is also a chance that the results of this investigation on listening comprehension were unduly influenced by the instrument selected for measuring
listening. As was explained in detail in the Listening Comprehension section of the Results Chapter, the paired samples t-test which compared the MPT pre- and post-tests of the entire participant pool failed to reach significance. This is relevant, because when examined from the perspective of MPT scores alone, it would appear that only 62% of the entire participant pool made any improvements at all. Moreover, 100% of those improvements were very small, registering as only one to three points of gain on the post-test. However, 96% of the total participant pool perceived themselves to have made some degree of listening comprehension improvement on the opinion survey given at the end of the term, and 48% of the participants in this study selected the descriptors “a good deal” or “a lot” to characterize their improvement. Furthermore, when examined from the perspective of MPT scores alone, it would seem that 15% of the total participant pool did not improve their L2 listening skills at all (MPT scores did not change) and 23% of them actually became less proficient listeners over the course of the 10-week term (MPT post-test scores dropped).

The contrast between perceived improvement and MPT listening results is also apparent when examined exclusively from the perspective of the experimental participants. As was displayed in Figure 8, 50% of the experimental participants chose the descriptors “a good deal” or “a lot” to describe their improvement. The following positive comments regarding L2 listening improvement, taken from the experimental participants’ reflective video journals, evidence that some learners perceived themselves to be greatly benefitting from the metacognitive strategy instruction:
• “I can say that I can understand more information than the beginning of this quarter. I can feel my change. I can feel my improvement. I think that is not only because of these two [comprehension monitoring & double check monitoring] strategies, but also all of strategies help me go ahead” (P2J—no change in pre- and post-test MPT listening scores).

• “[Today I] got the main idea and 80% of the details. [I] think my level went up and I’m so happy about it and I think that the 11 strategies help me a lot” (P2H—no change in pre- and post-test MPT listening scores).

• “[My listening level] has improved a lot because previously I listen without any strategies, so listening in that time so hard but now these different strategies help me to think, plan, decide how I will listen and how I can catch the information quickly and easily” (P2K—two point increase on the MPT post-test)

• “I got the main idea [today] with strategies inferencing, comprehension monitoring and attention. I’m happy with my level and listening is becoming easier and easier. Now I can understand much better than my first time” (P1B—one point increase on MPT post-test).

While it is possible that all of the learners who perceived themselves to have made a good deal or a lot of improvement, those who had no increase in MPT scores, and those whose post-test scores dropped were inaccurate in their self-assessments and over-estimating their abilities, it is also likely that the MPT, which was designed for use as a placement test, is not sensitive enough to register the incremental changes in ability
characteristic of achievements made during the course of a 10-week term. These results, therefore, call into question its use as a measure of achievement at the end of every term, though it should be stated that the IEP does not isolate listening scores when determining placement and does not require statistically significant results in order to move students between levels.

While none of the participants in this study (experimental or comparison) made much improvement on the MPT listening section, of secondary concern is the question of whether any multiple-choice standardized assessments are capable of measuring L2 listening comprehension effectively. Confounding variables, such as the reading skills required to decode the questions, the possibility of testing alternate constructs (such as grammatical knowledge), and the fact that multiple meanings inherent in texts can lead a learner to interpret main idea and inference-type questions in a different manner from the author of the question (Perkins, 1998) all make the testing of receptive skills in general, and listening in particular, challenging and complex. Those listening researchers who elected to use alternative assessments in lieu of standardized assessments (Carrier, 2003; Cross, 2009; Graham & Macaro, 2008) to measure learners’ listening comprehension abilities were likely doing so for sound pedagogical reasons, such as a desire to measure actual listening comprehension abilities. Unfortunately, alternative assessments, such as the free-recall technique, involve problematic validation procedures and are time-consuming to administer and evaluate. They are not, therefore, practical replacements for standardized assessment in many classrooms or by institutions needing to test large numbers of people in valid and reliable ways.
Finally, the small samples in the present study deserve special mention. A total of only 26 participants took both the MPT pre-test and post-test. Vandergrift and Tafaghodtari’s (2010) investigation, in which the experimental participants did reach statistical significance on the standardized University of Ottawa FSL placement test, had a considerably larger participant pool (n=106). In the present study, a larger sample might have created the conditions necessary for statistically significant results despite the fact that only small amounts of gain appear to be possible on the MPT listening assessment after a 10-week course of study. This conjecture is based upon the results of the two paired samples t-tests examining the means of the pre- and post-tests of each group separately. This analysis showed that the experimental participants’ amount of gain came considerably closer to significance than that of the comparison group. The amount of gain and the size of the sample were both too small to achieve a significant result, but had the number of participants been larger, this trend may have been realized as statistical significance.

In conclusion, the results on the listening comprehension portion of this investigation suggest that the experimental pedagogical cycle was ineffective at improving the experimental participants’ listening comprehension abilities. This may have been due to the fact that it was not superior to the methodology with which it was compared, may have been the result of too short a treatment period, may have been related to instrument used to measure the construct of listening, or the results may have been negatively impacted by small sample size or the cultural backgrounds of the participants. Due to the number of variables which may have had an effect on the results,
the cycle needs to be empirically investigated again with a larger (and possibly heterogeneous) participant pool, possibly for a longer period of time, and measured with either a more discriminate standardized listening assessment, or a standardized assessment in conjunction with a holistic instrument.

Metacognitive Awareness

*Metacognitive Awareness Listening Questionnaire Total Scores*

In contrast with the results on the MPT listening test, the participants’ self-reported metacognitive awareness total mean score as measured by the Metacognitive Awareness Listening Questionnaire (MALQ) did reach statistical significance, both in terms of its relationship with the comparison group’s post-test total mean score and in an analysis of the differences in the pre-test and post-test means of the two groups. Therefore, the experimental participants’ levels of self-reported metacognitive awareness were significantly raised as a result of having been instructed with the experimental pedagogical cycle.

The fact that the MALQ is a measure of self-reported metacognitive awareness is important to consider. Because the participants’ actual metacognitive strategy use was not measured through the use of think-aloud protocols or stimulated-recall interviews, there is no way to verify that the listening behaviors which the learners’ agreed with on the questionnaire were also actually being utilized during listening tasks. The experimental participants spent 10 weeks learning about, thinking about, practicing, and discussing both the listening process and strategic listening behaviors. It is entirely possible that this training led these students to beliefs about the desirability of certain
listening behaviors, regardless of whether or not these behaviors were actually in use. Therefore, there is a chance that the experimental participants’ significant drop in scores on the MALQ could be due to social desirability bias (Dörnyei, 2007), or an understanding that the statements on the MALQ, such as item number two, “I pay more attention to the text when I have trouble understanding,” are the correct thing to do, whether or not this is what is actually being done.

This is especially likely to be the case with those items on the MALQ which are related to monitoring and attention. Item number 13, “As I listen, my interpretation quickly changes once I realize that it was not correct,” for example, describes a monitoring strategy. Effective monitoring requires metacognitive strategy use in conjunction with the processing of L2 input, which is much more challenging than using strategies before or after the actual listening has taken place. One participant explained this phenomenon particularly well during a focus group interview, saying “some strategies are very physical—in the brain always takes—energy in your mind” (P2D). Attention strategies are challenging for the same reasons. One learner, in a reflective video journal, had this to say about the selective attention strategy (deliberately focusing one’s attention on particular, predetermined elements in the selection) and directed attention (telling oneself that distractions will be ignored and 100% of one’s attentional resources will be focused on the text), “When I listen I don’t have enough time to think these two points” (P2E). However, no two learners are alike, and the strategies that some learners find difficult are considered in a very different light by others. For example, another learner had this to say about directed attention, “It helps us get complete
information and good details from listening and speaking,” and made this comment about selective attention “It helps us get what we need from listening and reading easily” (P2K).

Whether or not a particular strategy or group of strategies was considered feasible or actually utilized by a particular learner is dependent upon a number of factors, including prior use and the perceived level of difficulty of particular strategies. Keeping in mind that taxonomies of learner strategies were based on studies which described and analyzed the strategies and behaviors of successful language learners (Murphy, 1985; O’Malley & Chamot, 1990; Rubin, 1975; Vandergrift, 1997; 2003a), it is logical to conclude that some of the participants in the present study would already have been using certain strategies as a natural part of the information processing and decoding process. There was some evidence of this in the reflective video journals, in such comments as: “I have used [prediction and activation of background knowledge], but I didn’t know the strategies’ names” (P2E); “I [always] think about its bad or good [self-evaluation] after every listening” (P1D); and “I have always used comprehension monitoring, but I never knew the name” (P2G). It stands to reason that someone who has “always used” comprehension monitoring is going to have a very different opinion of it than someone who has never listened in this way and is attempting to for the first time. Therefore, it is possible that the experimental participants’ responses on the MALQ represented a combination of actually utilized strategies and strategies which they believed they should be trying to utilize.
The following comments, also made in the reflective video-journals, evidence beliefs that the more difficult strategies represent goals to work toward when practicing listening: “I think some strategies are a little difficult and some are easy but the difficult strategies I must use in the listening” (P2F); and “I think [some strategies] are very useful but sometimes I couldn’t use it very well, but after class I will try to study and master” (P2I). The next comment evidences the fact that this participant had yet to gain control over his ability to remain focused, but his awareness of this issue had potentially been raised as a result of the metacognitive strategy training: “I’m not good enough today because after one or two minutes my attention wandered” (P1C). It is possible that this participant was already aware of his issues with attention and the training had no effect on his level of awareness. However, it is equally likely that the same learner would not have been as aware of the effects of wandering attention on listening comprehension prior to the training. The statistically significant change in MALQ post-test total scores suggests that the latter interpretation is closer to the truth.

In fact, learners whose comments (either in class or in reflective video-journals) revealed that they already used certain strategies prior to the treatment were asked during the focus group interviews about their opinions regarding the value of explicit instruction in strategies with which they were already familiar. The response was positive, as is evidenced by the following comments:

- “I think if I know the strategies’ names and I will know it better and more clearly and next time or in the future I will pay more attention to it” (P2G).
• “Yes [the instruction was beneficial, because] before we always do this, but is faster –[the instruction] made it easier to use” (P2E).

In sum, the statistically significant drop in MALQ total scores of the experimental participants is likely a reflection of both an increase in metacognitive strategy use and an increase in awareness of positive listening behaviors to strive for during listening activities. Whether or not an individual’s responses were a reflection of the former or the latter is dependent on both the ways in which particular strategies are deployed and the individual learner’s prior use of particular strategies. It may take some learners longer than others to implement some of the more cognitively challenging metacognitive strategies, and different strategies are perceived as being cognitively challenging by different learners. Nevertheless, it can be said that a statistically significant increase in an awareness of the strategies utilized by successful listeners, or a significant increase in the understanding of positive listening behaviors, was achieved as a result of the experimental treatment.

Conversely, the practice-based listening instruction taking place in the comparison class did not have a significant effect on the comparison participants’ post-test MALQ total mean scores. This can be interpreted to mean that the participants’ awareness and understanding of their listening behaviors was unchanged as a result of their instruction. In other words, a class rich in listening practice but with very little explicit strategy instruction or focus on the metacognitive processes underlying listening acts will have very little impact on learners’ self-reported metacognitive awareness,
which is not surprising since the focus of the instruction was not on raising the learners’ metacognitive awareness.

Though it is listening comprehension improvement that is the goal of such training, and increases in metacognitive awareness are only of secondary importance, a methodology which is clearly capable of the latter remains relevant. It stands to reason that increased awareness of positive listening behaviors will transfer to increased use of positive listening behaviors immediately in some learners, eventually in others, and not at all in only a percentage of those who are exposed to such training. However, learners who are not exposed to such training may have less of a chance to alter their negative listening behaviors, and may have less awareness of the effect of their positive listening behaviors on listening comprehension. Studies have shown that increases in language awareness can have a positive impact on learning outcomes (Bolitho et al., 2003; Chamot, 2004; Jessner, 2006; Wenden, 1987; Vandergrift et al., 2006), suggesting that the MPT listening post-test scores may not be the end of the story as far as the experimental participants’ listening comprehension development is concerned. Future researchers investigating the effects of metacognitive strategy training should consider incorporating delayed post-tests of metacognitive awareness, and more importantly, of listening comprehension. Inclusion of such an analysis would provide the field with great insight into whether or not training of this nature has the potential to increase learners’ abilities as their language proficiencies also increase, or if the effects of the training wane as time passes.
As was reported in the Metacognitive Awareness section of the Results Chapter, the individual category mean scores on the MALQ were also analyzed in order to determine which areas of metacognitive awareness were the most powerfully impacted by the experimental treatment. The results of each of these analyses will now be discussed in separate subsections devoted to each category.

_Mental Translation Category Scores_

The experimental participants’ self-reported awareness regarding Mental Translation was not affected by the treatment. Conversely, the comparison participants’ Mental Translation scores rose during the course of the term, indicating lower levels of metacognitive awareness. Vandergrift and Tafaghodtari (2010) defined a low score (or a high degree of awareness) in this category as “the ability to use mental translation parsimoniously” (p. 8). Therefore, a rise in scores such as that of the comparison participants would either indicate the use of liberal amounts of mental translation in the course of listening tasks or increased awareness regarding that use. However, the pre- and post-tests of the comparison participants were not examined for the presence or absence of statistical significance, so it is not clear if this increase was simply due to chance or not.

Of more importance is the fact that the experimental participants’ awareness in this category was unchanged. This can be interpreted to mean that the treatment had hardly any effect on either use of or beliefs about Mental Translation. As mental translation is a negative listening behavior which novice learners must move beyond in order to be fluent listeners (Vandergrift & Tafaghodtari, 2010), this result needs to be
taken into consideration. Future researchers interested in this cycle would be advised to incorporate more tasks which raise learners’ awareness of the mental processes they engage in during listening, such as the peer-to-peer or teacher-student think-alouds as advocated by Chamot et al. (1999).

**Person Knowledge Category Scores**

Person Knowledge is the other category in which the experimental participants experienced very little growth, and it is noteworthy that Vandergrift and Tafaghodtari’s (2010) investigation produced a similar result. Vandergrift and Tafaghodtari defined this category as “learner perceptions concerning how they learn best, the difficulty presented by L2 listening, and their self-efficacy in L2 listening” (p. 8). The results demonstrate that the experimental participants’ attitudes towards L2 listening and beliefs regarding its level of difficulty were not greatly impacted as a result of the treatment.

Several researchers have discussed the relationship between low self-efficacy beliefs (or negative beliefs about one’s capability as an L2 learner) and decreased willingness to utilize strategies (Chamot et al., 1999; Macaro, 2006; Palmer & Goetz, 1988; Wenden, 1991; Winne, 1995). Wenden (1991) clearly explained that a strategic intervention which does not address learners’ self-efficacy beliefs runs the risk of being less successful because those who perceive themselves to be less able learners will be less likely to attempt the new strategies.

Analysis of the reflective video-journals reveals evidence that some individuals’ self-efficacy beliefs were improved as a result of the treatment, through comments such as “[Prediction and activation of background knowledge] can improve confidence,” and
“if I used these strategies all the time I think my listening skill will be the easiest for me.” However, the results make it clear that the experimental pedagogical cycle did not have a great enough effect on the participants’ attitudes and beliefs about L2 listening as a group, and this may have had a negative effect on their willingness to experiment with strategies both during the treatment and in future listening tasks. As Vandergrift and Tafaghodtari’s (2010) investigation produced a similar result in the Person Knowledge category, the pedagogical cycle may be improved upon by the inclusion of specific activities which target learners’ self-efficacy beliefs, such as those suggested by Wenden (1991) and Chamot et al. (1999), or inclusion of instruction in and practice with socio-affective strategies such as “lowering anxiety,” “self-encouragement,” and “taking emotional temperature” (Vandergrift, 1997, p. 395).

**Problem Solving Category Scores**

Problem Solving, defined by Vandergrift and Tafaghodtari (2010) as “inferencing on what is not understood and monitoring those inferences” (p. 8) is the only category in which the experimental and comparison participants’ mean scores dropped to the same degree, suggesting that both the experimental and comparison group methodologies were equally capable of improving the learners’ problem-solving abilities. However, since the experimental group’s decrease in scores was not great enough to reach statistical significance, it can be inferred the comparison group’s was not either, and neither approach was as effective at raising awareness of problem solving as they might have been.
Interestingly, in Vandergrift and Tafaghodtari’s (2010) study, Problem Solving was the only category (other than Mental Translation) to result in a significant decrease (a significant increase in awareness) at the end of the treatment. This indicates that something about the way that they implemented the experimental pedagogical cycle was particularly effective for increasing their participants’ awareness of monitoring and inferencing strategies. Vandergrift and Tafaghodtari hypothesized that their participants’ post-test scores in this category may have been a result of repetitive and cyclical exposure to metacognitive processes over the course of the treatment, leading the learners to “[acquire] implicit knowledge about L2 listening through task performance” (p. 19). However, the participants in the present study were also exposed to the metacognitive processes underlying successful listening comprehension in a repetitive and cyclical format, but their scores in the Problem Solving category of the MALQ were significantly different.

One possible explanation might be the treatment time. As was explained in the previous section, Vandergrift and Tafaghodhari’s (2010) experimental participants worked through the pedagogical cycle 13 times in 13 weeks. In the present study, the experimental participants worked through the cycle seven times in 10 weeks. Perhaps more exposure is necessary to effect changes in awareness of Problem Solving. Another possible explanation might be related to differences in the types of listening selections utilized in both experiments. Vandergrift and Tafaghodtari (2010) employed “authentic-type texts” (p. 9). Though they do not describe this term, it can be inferred that these selections were similar to real-time, authentic texts, as Dr. Vandergrift has often
advocated for the inclusion of authentic materials for practice in L2 listening instruction (Vandergrift, 2002; 2003a; 2003b; 2004; 2007). Conversely, the selections utilized in the present study came from the course book (Mackey et al., 2007) and were specifically produced for intermediate-level English language learners. It is possible that the degree of challenge inherent in decoding “authentic-type texts” (p.9) was greater than that necessary to decode the selection-types utilized in the present study. This difference in degree of challenge may have caused the participants in Vandergrift and Tafaghodtari’s investigation to rely more heavily on monitoring and inferencing strategies, leading to greater levels of awareness of these strategic processes.

If this is the case, this might also explain the similarity between the comparison and experimental participants’ scores in this category. The Fast Listening activity, which was regarded so favorably by the comparison group participants, always employed authentic materials (NPR’s Morning Edition headlines), and required large amounts of inferencing and problem-solving on the part of the learners. One comparison group participant, in the opinion survey administered at the end of the treatment, made a comment which evidences his/her use and appreciation of inferencing during Fast Listening, “Fast listening is my favorite activity because this activity let me listen and try to gauss.” Therefore, it may well be that the experimental pedagogical cycle minus authentic listening materials, and practice with authentic materials minus strategy instruction, are equally effective means of improving listeners’ problem solving skills. However, neither of these methods is as effective as an approach which combines the two. As a result, it is recommended that future researchers and educators interested in a

Planning and Evaluation Category Scores

The Planning and Evaluation category, comprised of questions aimed at uncovering “how listeners prepare themselves for listening and evaluate the results of their listening efforts” (Vandergrift & Tafaghodtari, 2010, p. 8) was the only category in which the experimental learners’ mean scores reached statistical significance both in comparison of the two groups’ post-test mean scores and in analysis of the means of the gain scores of the two groups. This is because the comparison participants’ mean scores actually rose (indicating a decrease in metacognitive awareness) over the course of the treatment, whereas the experimental participants’ scores significantly dropped (signifying a rise in metacognitive awareness). These results would suggest that something being done in the experimental classes had a much different effect on the learners’ attitudes towards the pre- and post-listening phases than was happening in the comparison class.

The most likely explanation involves the fact that in the experimental classes, the experimental pedagogical cycle was used with every single in-class listening activity. It exploited the pre- and post-listening phases, involved explicit instruction in the metacognitive processes that accompany those phases, and provided the learners with ample amounts of controlled practice with specific metacognitive strategies necessary for the actualization of those processes. In other words, before every listening act, the learners were actively engaged in activating background knowledge, were personally responsible for making logical predictions based on that background knowledge, were
tasked with setting personal listening goals, and were required to make plans regarding which strategies they would like to use while listening. Several learners made comments in their reflective video-journals which evidence their raised awareness regarding the pre-listening phase:

- “I didn’t catch the information, but I used my background knowledge to understand why and how [manga comics are] popular” (P1D).
- “When I know information for listening it makes it easier to understand the new information coming in while listening” (P1B).
- “When the listening begins, we have to prepare to touch the words, and we have to think about the main idea so it’s easy to listen. Making predictions is also important because you should always use your brain to think” (P1C).
- “[Today] I understood most of the main idea and almost all of the details because I’m a big fan of [this topic]. I’m so happy about my level, but I wish to have this kind of success when I don’t have this much background knowledge” (P2J).
- “I’m very happy that I got my goal because of lots of background knowledge and good predictions. I thought about my background knowledge while I listened, and that was really helpful” (P2G).

In addition, every listening practice activity was followed by formal opportunities to reflect on individual problem spots, check goals, and evaluate performance, and evaluate strategic choices and the strategies’ effectiveness. The number and quality of
learner comments in the reflective-video journals regarding these aspects of the experimental pedagogical cycle were much fewer than those regarding the pre-listening phase, suggesting that perhaps these activities were not perceived to have as salient and immediate effects on their L2 comprehension abilities. Nevertheless, the following comments offer some insight into the raised awareness of the learners as a result of these post-listening activities:

- “I understood some parts, and I’m happy, but my level is not high so next time I want to catch more. I plan to use directed attention, elaboration and problem-identification [tomorrow]” (P1E).
- “Today I got 50% of the listening and the main idea, so I’m happy…tomorrow I will use double-check monitoring to make sure I didn’t have any misunderstanding points” (P2B).
- “Today I understood the main idea because I tried my best to catch all the details and I used many strategies: directed and selective attention, comprehension monitoring, and double-check monitoring, so I achieved my goal” (P1D).
- “I like [problem identification] because when I hear the first time I can hear some words but I can’t sure this is right—and the second time I have enough time to be sure this…is right” (P2E).

In the comparison class, on the other hand, listening practice was approached in three distinct ways, none of which exploited the pre- or post-listening phases in the same ways as the experimental methodology. First, the comparison participants engaged in
interactional listening practice which involved discussion with peers on a variety of
topics. This exercise, called *Conversation Starters*, did not incorporate any pre- or post-
listening practice activities. Second, when the learners practiced listening with the course
book (Mackey et al., 2007), they engaged in discussions and looked at pictures designed
to help them activate their background knowledge, occasionally, the instructor informed
them that they should think about what they already knew about a topic prior to listening,
and predictions were made on a voluntary basis. However, due to the fact that the
comparison group instructor was not attempting to raise the learners’ metacognitive
awareness or engage in explicit strategy training, these activities were not presented as
specific strategies and were not evaluated by the learners. Research has shown that by
explicitly informing learners of the values and potential benefits of strategic behaviors,
such as these two pre-listening activities, the knowledge of the teacher can be transferred
to the learners, empowering them with increased autonomy and awareness (Carrier, 2003;
Chamot, 2004; Chamot et al., 1999; Goh, 2008; Wenden, 1991).

The post-listening phase during course book listening practice consisted of
checking the multiple-choice comprehension questions provided by the book. Because
the pre-listening phase, first exposure, second exposure and post-listening phase were
completed in a 50 minute lesson, there was little time left over at the end. The
comparison group participants also did not engage in any sort of formal reflective or self-
evaluative activities, including the setting or checking of personal listening goals.

When the learners practiced listening with the *Fast Listening* technique, no pre-
listening activities were incorporated. This was intended to more closely mimic real
world listening experiences, and may have even given the comparison participants an advantage when taking the MPT listening test at the end of the term. However, repeated exposure to this type of listening practice activity may have had the effect of lowering the learners’ awareness of the value of the pre-listening phase. The post-listening phase was more involved during *Fast Listening* than during course book listening practice, due to the fact that the learners were tasked with collaborating, reconstructing the text, and comparing their oral summaries with those of other groups. However, formal opportunities to reflect on progress, to evaluate performance or evaluate the strategic behaviors which had helped them to parse and make sense of the input were not incorporated.

Therefore, it can be inferred that the heavy focus on preparing to listen and evaluating one’s comprehension just after listening had a powerful effect on the experimental learners’ level of metacognitive awareness regarding appropriate behaviors for before and after listening. Furthermore, the experimental participants’ scores on this category of the MALQ contrast significantly with those of the comparison participants’, for whom these aspects of the listening process were neither made explicit nor given a heavy amount of instructional focus. However, whether or not this has any effect on actual listening abilities is questionable, as the results from the listening comprehension portion of this investigation were inconclusive and did not show a significant level of improvement resulting from either methodology.
Directed Attention Category Scores

Directed Attention represents the other category on which the experimental participants’ post-test means were significantly lower (indicating significantly higher levels of metacognitive awareness) than those of the comparison participants. The comparison participants’ mean scores in this category also dropped, but their decrease was not examined in terms of its significance. The Directed Attention category involves “how listeners concentrate, stay on task, and focus their listening efforts” (Vandergrift & Tafaghodtari, 2010, p. 8), and the experimental treatment was shown to be significantly effective in this regard.

As was the case in the Planning and Evaluation category, this is likely due to the explicit nature of the instruction regarding those metacognitive strategies which work to focus and direct learners’ attention: Directed Attention and Selective Attention. Learners in both groups concentrated and focused on the input when engaging in listening practice, and it is possible that the challenging Fast Listening activity, which the comparison group participants deeply enjoyed, was at least partially responsible for their increased awareness of the benefit of paying close attention. However, only learners in the experimental group were informed that there are strategic behaviors which they could choose to employ to increase their focus and could work towards improving if they discovered themselves to be lacking in focus. Additionally, only learners in the experimental group practiced using these two strategic behaviors in the course of listening tasks and had the opportunity to evaluate their effectiveness. The following
learner comments, taken from reflective video-journals, illustrate some of the experimental learners’ opinions regarding the attention strategies:

- “Giving more focus when listening can allow for understanding more words” (P2D).
- “I can use [directed attention and selective attention] all the time, in class, on tests, in daily life, and they help me to catch what I want from the information” (P2K)
- “When you focus you will understand much better. Focusing makes it easier to hear important information and important details, and focusing on difficult parts helps you understand” (P1B).
- “Selective attention makes me focus and helps me get the correct answer” (P1E).

In conclusion, the explicit and cyclical metacognitive strategy training utilized in the experimental classes created statistically significant differences between the post-test mean scores of the experimental and comparison groups in terms of both Planning and Evaluation and Directed Attention. Furthermore, the means of the gain scores of the experimental group were significantly greater than those of the comparison group in the Planning and Evaluation category. The experimental pedagogical cycle was, therefore, very successful at raising the learners’ awareness regarding these two aspects of metacognition. However, the cycle was as effective at raising learners’ awareness regarding Mental Translation, at altering self-efficacy beliefs regarding listening, or at raising awareness of monitoring and problem-solving skills. The latter may potentially
be due to the use of listening selections specifically produced for use in L2 listening classes rather than real-time, authentic selections. Therefore, future researchers and educators interested in this cycle should utilize authentic listening selections, build in activities which raise awareness of the mental processes taking place during listening, include activities which directly address learner attitudes and self-efficacy beliefs, and incorporate instruction and practice with socio-affective strategies for dealing with negative emotions encountered during listening practice.

**Correlation between Self-Reported Metacognitive Awareness and Listening Comprehension**

As this investigation was based upon the premise that increased levels of metacognitive awareness about listening can improve listening comprehension abilities (Field, 2000; Goh, 2008; Vandergrift, 2002; 2004; Vandergrift & Tafaghodtari, 2010), an analysis of the relationship between the two constructs (and the instruments used to measure them) was conducted. The Spearman’s bivariate correlation showed that there was no relationship between the listening portion of the Michigan Placement Test (MPT) and the Metacognitive Awareness Listening Questionnaire (MALQ) given at the beginning of the treatment. However, the same analysis, run on the post-test scores, showed a statistically significant, though weak, correlation between the two instruments. The MPT scores rose slightly, but the MALQ total scores dropped significantly, making the change from non-correlation to significant correlation likely due to changes in the MALQ scores.
A detailed analysis of Figure 6 reveals that 15 of the 26 participants’ pre-test scores followed the same correlation pattern (low MALQ and high MPT to high MALQ and low MPT) of the majority of the post-test scores (Figure 7). However, the remaining 11 participants had score combinations that varied so greatly from this pattern that overall the correlation was near zero. Then, after 10 weeks of instruction and practice, 24 of the total participant pool (n=28) had scores that fell into this pattern, which was enough to create a significant but weak correlation between the two instruments.

Separation of the participants’ scores into experimental and comparison groups reveals that the correlation held for both groups of learners. At the post-test, 94% of the experimental participants and 67% of the comparison participants had scores on the two instruments that fit the pattern of the correlation. Interestingly, the comparison participants’ scores generally fell into the lower right hand quadrant of the scatter plot (Figure 7), indicating both lower MPT scores and higher MALQ scores (lower self-reported metacognitive awareness), whereas the majority of the experimental participants’ MPT scores were higher and MALQ scores were lower (indicating higher levels of self-reported metacognitive awareness).

Additionally, as was mentioned in the Correlation between Self-Reported Metacognitive Awareness and Listening Comprehension section of the Results Chapter, of the four individuals whose post-scores did not correlate, three were comparison group participants and one was an experimental participant. Examination of these individuals’ scores is interesting, because each of these three comparison participants had high MPT scores and high MALQ scores (indicating lower levels of self-reported metacognitive
awareness). These individuals demonstrate that it is possible to have higher listening proficiency (as measured by the MPT) without having a strong knowledge of the metacognitive processes underlying the listening comprehension. On the other hand, the single experimental participant whose scores were uncorrelated (P1B) had the highest level of self-reported metacognitive awareness of anyone in the study on the post-test, but her MPT listening score remained relatively low. This illustrates the fact that strong understanding of those metacognitive processes does not always immediately translate into high listening scores, which is logical considering the number of individual variables which can influence the jump from agreement with strategy use to actual strategy use (overall L2 proficiency, prior experience with a strategy, self-efficacy beliefs, etc).

However, for the majority of participants in this study (86%) the post-test scores were correlated, demonstrating that knowledge or awareness of metacognitive strategy use can be related to listening proficiency as measured by the MPT. The fact that only 58% of the total participant pool had correlated pre-test scores may be attributed to the fact that many of those participants who increased their knowledge of metacognitive strategy use also either increased their MPT scores or retained MPT scores that were already relatively high. Likewise, many who did not increase their awareness of metacognitive strategies were the same learners who began and ended the treatment with lower MPT scores. It may therefore be the case that those learners with higher levels of listening comprehension during the treatment were better able to comprehend the strategy training, and this increased understanding of the concepts covered in the class translated into decreased scores on the MALQ post-test (indicating raised awareness). Likewise,
those learners who began the treatment with lower MPT scores, evidencing lower listening comprehension abilities, may have been less capable of understanding the strategy training, and therefore had less change in their metacognitive awareness as measured by the MALQ. This lends some evidence that the threshold hypothesis (Cross, 2009; Imhof, 2001; Macaro, 2006; Winne, 1995) is an important factor to consider when implementing strategy training of this nature, and the linguistic level at which learners are capable of comprehending instruction of this nature in the L2 may be higher than beginner or true beginner as stated in Vandergrift (1997). Additionally, there were two fewer participants at the time of the pre-test, because two individuals did not take the MPT test at the beginning of the term. Each of these individuals had post-test scores that fit the pattern of the correlation, and it is possible that their pre-test scores would have as well, which might have had some influence over the strength of the pre-test correlation.

Regardless, these results have implications for anyone desiring to use the MALQ for diagnostic purposes. It cannot be assumed that high MALQ scores (which represent low levels of metacognitive awareness) will correlate with low listening scores prior to metacognitive strategy training; the same is true of reverse scores. However, the significance of the post-treatment correlation lends some evidence that metacognitive awareness can have a positive relationship with L2 listening comprehension, especially after training. Considering the fact that the comparison participants’ scores were largely on the lower end (lower MPT listening scores and lower self-reported metacognitive awareness) at the post-test, these results do provide some support for the inclusion of cyclical metacognitive strategy training in the L2 listening classroom.
Participants’ Perceptions of Method of Instruction Received

The final component of this investigation to be discussed is that of the participants’ perceptions regarding the treatment, which were gauged through the use of an anonymous on-line survey created by the research for the purpose of this investigation (Appendix B). Participants were informed that their responses on the questionnaire would be innominate, in the hopes that their answers would be an accurate representation of their perceptions of L2 listening improvement as well as the instructional factors which led (or did not lead) to that improvement.

Research has shown that learners have positive reactions to strategic interventions of the type employed in the present investigation (Vandergrift, 2003b); however, the results of this aspect of the present investigation appear to only partially substantiate earlier findings. While 94% of the experimental participants rated themselves as having improved their L2 listening skills, only 50% of them perceived to have experienced “a good deal” or “a lot” of L2 listening improvement. The other 44% felt that the descriptor “somewhat” best described their level of improvement (no experimental participants chose “a little bit” in answer to this question). This would suggest that only about half of the learners in the experimental group perceived the metacognitive strategy training to have had a really positive impact on the development of their listening skills. Additionally, a slightly smaller (but similar) proportion of the comparison group participants (44%) perceived themselves to have made either “a good deal” or “a lot” of improvement. Therefore, the metacognitive strategy training did not have a considerable
impact on the participants’ perceptions of their level of improvement as compared with a practice-based methodology.

However, strategy instruction was selected as an element contributing to this improvement by 82% of those learners (and by 100% of those learners who rated their improvement with the descriptors “a good deal” or “a lot”). Similarly, when asked to select one element of instruction that was the most beneficial for the improvement of listening skills, 66% of the experimental participants cited aspects of the pedagogical cycle, and no other category had a larger representation in the learners’ comments. Finally, it should be noted that those learners who perceived the strategy training to be beneficial had very strong positive opinions about it, as evidenced by this selection of comments taken from both the most beneficial and favorite activity portions of the survey:

- “All of the strategies are very helpful for me and really help me to provide my listening level.”

- “Because I think [strategies are] the thing that I was missed, And I haven't any idea about these strategies and how we can use them in the listenning before.”

- “[Listening strategies are the] most helpful… Because, this one give me many listening skill to improver my listen.”

- “These strategies are the key to get successful.”
• “I have never heard about strategies before and now i know and im really happy because it helps so much to prepare to listening when you even dont know about what the text about because of strategies. Thank you!”

• “My favorite in class is to study stadigy. i can from these to improve my listening [because]… i can [use] my knowledge to make guess, many times, it is useful, that can make me understand the listening clearly.”

• “These strategies [are helpful because] i can always use after, and very improve my listening ability.”

Positive comments regarding explicit strategy instruction, such as these, were made by 56% of the experimental participants. The other 44% of the learners who were exposed to this methodology felt that other instructional activities were more helpful (or more enjoyable) than the metacognitive strategy training. However, half of that 44% found other aspects of the cycle (such as summarizing or discussing the selection with a small group) to be the most helpful/enjoyable thing about the class. The remaining four learners (22%) did not make any positive comments on the survey about any aspects of the experimental pedagogical cycle. In addition, one of these learners claimed that strategy instruction was his/her least favorite thing about the class, though s/he did not elaborate about the reason behind this choice. It should be noted that the same proportion of learners in the comparison class (22%) made comments which suggested that their L2 listening skills had not improved as a result of their instruction. Of the 22% of experimental learners who did not support the training, none said that the instruction had not helped them to improve.
Nevertheless, there appears to be a contrast between these findings and those of Vandergrift (2003b), as it was claimed that the learners in that investigation “responded positively to the listening activities and the approach” (p. 434). This contrast will now be explored. Vandergrift assessed the perceptions of the university-aged participants (n=41) through a qualitative analysis of the learners’ reflective journals. These journals were not anonymous, and in the journals the learners reflected exclusively on those tasks and activities which were related to the metacognitive training. The subsequent analysis of their comments revealed that the learners were pleased with the effectiveness of the tasks and the content of the listening selections. In the present study, however, the participants were not simply asked to reflect on the training, but on their entire experience in the Listening & Speaking class, including, among other things, a variety of different types of homework assignments.

In fact, all of the experimental participants (22%) whose comments on the survey did not indicate that aspects of the experimental cycle were of the most benefit claimed that watching videos for homework and either answering questions about them or summarizing what had been understood had been more instrumental for the improvement of L2 listening skills. Therefore, it may be that the participants in the present appear to have less favorable attitudes towards their training than Vandergrift’s (2003b) participants because they had something else to compare the pedagogical cycle too.

One of the experimental participants made a comment on the anonymous survey which indicated an unfavorable response to the experimental methodology; 56% of them responded very positively in full knowledge that their responses were innominate.
Perhaps the results of this anonymous survey provide a more accurate portrayal of the participants’ attitudes towards listening instruction than Vandergrift’s (2003b) qualitative analysis of reflective journals, which were not anonymous and focused solely on perceptions of the training under examination. Learners are, after all, individuals with diverse learning styles, preferences, and opinions regarding instructional techniques and methodologies. What works well for one may not work as well for another; what is perceived as highly effective by one may not be by another. Therefore, the strong opinions of 56% of learners in the present study regarding metacognitive strategy training should be viewed in a positive light, and broad, general statements that make claims about the perceptions of entire groups of students are best approached with caution.

Moreover, considering individual variation and the impact that differences in learning styles, learner preferences, and learner backgrounds can have on learners’ perceptions, the strong support of explicit strategy training evidenced by 56% of the experimental learners, and the support of other elements of the cycle (including explicit strategy training) of 78% of the experimental learners can be regarded as positive. These facts, when combined with the significant increase in metacognitive awareness as measured by the MALQ, highlight the fact that training of this nature not only has the power to raise learners’ awareness of positive listening behaviors, but is also regarded positively by the majority of participants. Additionally, the insightful quality of many of the experimental participants’ comments on both the anonymous survey and throughout the term in the reflective video-journals reveals that emphasis on self-evaluation and reflection had a positive impact on these learners’ abilities to contemplate the listening
process and evaluate their personal progress as L2 listeners. While there is no empirical evidence to support that this ability can or will translate into improved listening scores, there is also no empirical evidence indicating that the opposite is true. Additional empirical investigations are still needed to ascertain whether or not training in the metacognitive processes underlying listening comprehension has the power to influence listening test scores both immediately after the treatment as well as on delayed post-tests.

The comparison group participants’ opinions on the anonymous survey also yielded two interesting findings. The first was generated from the item on the survey which asked: *What were your least favorite in-class (or homework) listening practice activities?* More of the comparison participants’ responses to this question fell into the *Tests* category than any other (33%, or three learners). One learner specifically stated that this was “Because test makes me nervous.” This is in line with Arnold’s (2000) findings regarding the relationship between listening assessment and learner anxiety. Conversely, while 12% of the experimental participants (two learners) also stated that summaries, which were used as listening assessments in the experimental classes, were their least favorite elements of instruction, two other experimental learners selected summary writing as one of the most beneficial elements of the class. In contrast, no comparison group learners made any positive comments about assessment. These findings would suggest that while assessment is never a popular part of the instructional process, it remains a necessary and essential component. Utilizing summary writing to determine learners’ levels of comprehension not only allows instructors a more comprehensive and detailed picture of what was comprehended than is possible with
multiple-choice comprehension questions, but also has the ability to transform assessment into something perceived as beneficial by some learners. As one of the experimental learners explained on the survey: “[Summary writing] can help me which i really understand the listening detail or not.” However, neither form of assessment utilized in the experimental classes was referred to as a test. Additionally, all of the pre-, during and post-listening activities engaged in during assessments were the same as those engaged in during practice activities. It could, therefore, be the case that ensconcing assessments (regardless of assessment type) within everyday practice activities and removing the test label can decrease learner anxiety, while still providing necessary feedback to the learners regarding their progress.

The second interesting finding relates to the Fast Listening component of the comparison class. 78% of the comparison participants made positive comments about Fast Listening in response to both the most beneficial or favorite activity survey questions, and no negative comments about this activity were made. As was explained in the Participants’ Perceptions of Instructional Elements section of the Results Chapter, Fast Listening shared many important characteristics with the experimental pedagogical cycle, as illustrated by Table 15.
As Table 15 illustrates, the two approaches shared many important components. In fact, the primary pedagogical differences between *Fast Listening* and the listening practice activities of the experimental pedagogical cycle are explicit strategy instruction, bottom-up processing practice, and opportunities for learner reflection. Of these, only explicit strategy instruction was specifically mentioned by the experimental participants as being enjoyable or beneficial. It’s important to note that explicit strategy training is a critical and essential component of the cycle, and the experimental learners’ recognition of this is evident in both the quantity and quality of their comments about strategy training on the questionnaire. However, the fact that the comparison group participants responded so positively to this form of listening practice is also informative. Whether it

<table>
<thead>
<tr>
<th>Experimental Cycle Features</th>
<th>Shared Features</th>
<th>Fast Listening Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listening selections produced for intermediate level learners</td>
<td>Noting what could be understood during first &amp; second exposures</td>
<td>Authentic listening selections (NPR <em>Morning Edition</em> headlines)</td>
</tr>
<tr>
<td>Explicit strategy training</td>
<td>Collaboration with listening partners to reconstruct text</td>
<td></td>
</tr>
<tr>
<td>Exploitation of pre-listening phase</td>
<td>Purpose of listening: Understanding main idea, details, and making inferences</td>
<td></td>
</tr>
<tr>
<td>Opportunities for reflection</td>
<td>Non-evaluative nature</td>
<td></td>
</tr>
<tr>
<td>Bottom-up processing practice</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
was the non-evaluative nature, collaboration with peers, the verification of comprehension through noting what could be understood, the oral reconstruction of texts, the use of authentic selections, or some combination of these characteristics which made this aspect of listening practice so appealing is not evident in the learners’ comments:

- “Fast listen is my favorite. Because fast listen let me think quickly and I must catch main work, it help me listen when I will study my class.”
- “i like fast listening activities because help me better understand”
- “i like fast listening, i can try to catch some key words, it is very interesting.”
- “listening fast. because help me more.”
- “Fast listening is my favorite activity because this activity let me listen and try to gauss, it's improve my listening fast.”
- “fast listening! very helpful”

However, these comments provide some evidence that the learners liked trying to catch key words from the rapid, authentic stream of speech and then inferencing to fill in the gaps. As a result, it can be inferred that both the use of authentic materials and the method of noting what had been understood and then collaborating with reconstruct the text had much to do with the learners’ positive reactions and perceptions of the utility of this activity. This fact, in combination with the preferences of the experimental group, lends support to the activities of the First and Second Verification stages of the experimental cycle (Table 3) as well to the use of short, authentic listening selections in conjunction with the experimental pedagogical cycle. Therefore, instructors who are
interested in adopting this cycle should periodically incorporate authentic materials if they are using a course book which contains specially produced selections. Furthermore, textbook writers should reconsider the nature of the listening selections intended for practice with lower level learners. Learners, even at lower levels, appreciate the challenge inherent in decoding authentic speech samples and feel that practice with such materials has more relevance and applicability to real-life listening experiences, as is evidenced by the following comment taken from the anonymous questionnaire: “Fast listen… help me listen when I will study my class.”

In conclusion, the present study produced inconclusive results in terms of listening comprehension, statistically significant results regarding increases in self-reported metacognitive awareness, and demonstrated that about half of the learners exposed to this type of training perceived it to be influential for the improvement of their L2 listening skills. A number of variables may have been the cause of the inconclusive results on listening comprehension, including, but not limited to, the possibility that that the methodology was not superior to the one with which it was compared and the fact that the instrument used to measure the construct of listening was incapable of registering incremental changes in ability over the course of a 10-week term of study. The significant results in terms of metacognitive awareness reflect a combination of declarative and procedural knowledge regarding metacognitive strategy use, and the lack of stimulated-recall or think-aloud protocols limits the possibility of uncovering how many of these self-reported listening behaviors were actually being put to use. The findings regarding learners’ perceptions revealed that 78% of the experimental
participants approved of the cycle, and 56% had positive feelings about explicit strategy training, whereas 78% of the comparison participants had positive feelings about the Fast Listening activity, which mirrored the experimental methodology in several ways. These findings will be summarized in the upcoming conclusion, and pedagogical implications and potential avenues for future research will also be discussed.
CHAPTER 6: CONCLUSION, PEDAGOGICAL IMPLICATIONS AND AVENUES FOR FUTURE RESEARCH

The investigation detailed in this document investigated the effects of explicitly presented metacognitive, cognitive, and socio-affective listening strategies introduced via a pedagogical cycle designed to sensitize learners to the metacognitive processes underlying listening comprehension. The impacts of this instructional methodology were compared to an alternate approach which focused on large amounts of listening practice activities with a variety of types of L2 input. The two methods were examined in terms of their effects on listening comprehension, self-reported metacognitive awareness, and learners’ perceptions.

The results regarding listening comprehension can be considered inconclusive due to the fact that neither the experimental nor comparison group made significant improvement in their listening comprehension abilities as measured by the MPT. Therefore, these results may have been influenced by measurement with an instrument which is not sufficiently sensitive to register incremental changes over the course of a 10-week term. Additionally, the small size of the sample, the cultural backgrounds of the participants, or the length of the treatment time may have influenced the results. It is also possible that the experimental methodology was not superior to the approach with which it was compared. In fact, a number of similarities between the experimental approach and the technique which received the largest number positive comments by comparison participants, *Fast Listening*, may indicate that the two approaches provided similar enough benefits to have eliminated the potential for a statistically significant difference
on the final measure of listening comprehension. However, paired-samples t-tests examining the growth of each group (experimental and comparison) independently uncovered a trend regarding a greater amount of growth on the part of the experimental group which may have reached statistical significance with a larger sample. Due to the large number of potential causes of these results, the inconclusive findings cannot be considered evidence that the experimental pedagogical cycle was ineffective. Additional empirical investigations regarding the effects of this type of training on learners’ listening abilities are still needed to confirm whether or not Vandergrift and Tafaghodtari’s (2010) results can be replicated in novel contexts and with different groups of learners studying different languages for different reasons.

However, investigations of this nature require the use of sensitive assessment tools capable of measuring the small changes characteristic of achievements made in a given term of study. Moreover, it is important to compare the experimental cycle to alternate forms of listening instruction which are also believed by the instructor to be “to the best of his or her ability” (S. Jarvis, personal communication, September 27, 2010). When the comparison group is artificially controlled, it can lead to statistically significant results which do not reflect the realities of classroom practice, which should be the goal of classroom-based, action research. Finally, incorporation of a delayed post-test, to measure the outcome of such training on listening comprehension abilities over time, would also provide the field with valuable findings that thus far have only been attempted by a single investigation (Graham & Macaro, 2008).
This investigation did provide support that strategy training conducted in a cyclical and repetitive manner, sensitizing learners to the metacognitive processes underlying listening comprehension, has a significant effect on learners’ self-reported levels of metacognitive awareness, specifically in the areas of Planning and Evaluation and Directed Attention. Since the effects on Person Knowledge, Mental Translation, and Problem-Solving were not significant, this pedagogical cycle could be improved by adding activities aimed at identifying and revising learners’ self-efficacy beliefs (Chamot et al., 1999; Wenden, 1991), including socio-affective strategies targeting negative feelings related to L2 listening (Vandergrift, 1997), using authentic selections, and incorporating think-aloud activities to increase learners’ awareness of monitoring strategies and instances of mental translation (Chamot et al., 1999).

Additionally, considering the complex and largely internal nature of metacognitive strategy use, future researchers interested in examining the effects of metacognitive strategy training should examine both metacognitive awareness as measured by the MALQ and actual metacognitive strategy use through the use of stimulated recall or think-aloud protocols. Triangulation of this nature can “help establish validity and reliability” (Chamot, 2004, p. 22) and will provide a more detailed analysis of the effects of the training on the actual listening behaviors of the learners.

Furthermore, though the participants in this study were not unanimous in their opinions regarding the value of this training, a majority of the experimental participants had strong opinions regarding its effectiveness. Similarly, the majority of the comparison group participants had strong positive opinions about an instructional technique (Fast
Listening) which mirrored the cycle in important ways. Fast Listening did not involve any pre-listening activities or explicit strategy training, but in terms of purpose for listening, method of determining comprehension, and lack of formal evaluation, it was similar to the experimental cycle. Also, the Fast Listening technique was practiced exclusively with authentic listening selections and has what could be considered a fun and motivating name. Comparison group support for this activity can be seen as support for the First and Second Verification Stages of the experimental cycle (Table 3) and can also be seen as support for practice with authentic selections. Educators and researchers interested in this cycle can benefit from this information in two ways. First, incorporation of authentic materials is essential for improving learner motivation, perception of benefit, and challenge. Second, the coining of compelling names for the stages in the experimental cycle (and the strategies to be included) might have an influence on learners’ perceptions of these elements.

Finally, additional research is needed regarding participants’ opinions and perceptions of cyclical, explicit metacognitive listening strategy training because few investigators have attempted to objectively measure participants’ opinions regarding instruction of this nature. However, future researchers interested in uncovering participants’ opinions should not only gather this data anonymously, but should also allow the learners to compare the experimental methodology to something else, whether it be additional instructional activities or those undertaken in previous learning environments. Both of these measures will provide the field with a more accurate and
honest appraisal of learners’ opinions than qualitative analysis of reflective listening diaries alone.

Considering the relative paucity of listening comprehension research in comparison with research into the reading skill (Dunkel, 1991; Morley, 1991; Vandergrift, 2007) it is perhaps not surprising that advances being made in listening research are taking so long to trickle down to individual classrooms. Mendelsohn (1998) criticized the gap between the findings of applied linguistics research and the inclusion of activities addressing those findings in L2 listening textbooks, and Vandergrift (2003b) leveled the same criticism five years later. The course book used in conjunction with the present study (Mackey et al., 2007) was produced five years after that, and still contained no information or practice activities related to listening strategies or the processes underlying listening comprehension.

Researchers have been advocating for explicit strategy instruction in the L2 listening classroom for many years (Carrier, 2003; Chamot, 2004; Chamot et al., 1999; Graham & Macaro, 2008; Lai, 2009; Macaro, 2006; O’Malley & Chamot, 1990; Rost, 2007; Rubin, 1994; Vandergrift, 2002; 2007; Wenden, 1991), but in the pilot observations that led up to the present investigation, explicit strategy training was not found to be a common component of listening instruction at the intermediate and high-intermediate levels in the IEP where this investigation took place. When strategies were practiced, they were generally practiced in isolation and often without the learners being informed of the name, value, or potential benefit of the strategy. In fact, more often than
not, the learners were not informed that they were using strategies at all, and in some 
cases, strategies were presented but not practiced.

While listening researchers are working to establish the best possible method for 
incorporating explicit strategy instruction into the L2 listening classroom, such as the 
cyclical approach examined in the present study, there is a good chance that practice-
based, product approaches with little emphasis on listening strategies remain the norm in 
ESL/EFL classrooms around the globe. Therefore, another important avenue for future 
research would be an investigation regarding the prevalence of listening strategy 
instruction in various ESL/EFL instructional contexts. This information would be 
especially relevant for university-affiliated IEPs, which are considered by many to 
promote cutting-edge pedagogical practices. An additional, related area in need of 
research would be an investigation into the inclusion of listening strategy and process-
based approaches in ESL/EFL teacher training programs. Finally, another 
comprehensive textbook review, similar to the once conducted by Mendelsohn (1998) is 
necessary in order to determine whether or not his findings regarding the inclusion of 
listening strategies in ESL/EFL textbooks have been remedied, and if so, to what degree. 
After all, if the majority of listening textbooks remain product-based and lack 
information about listening strategies, and pre-service teachers are not trained to 
approach listening as a process which learners can gain control of through strategy use, 
then there is no way for the information advocated by L2 researchers to reach the 
learners. These three lines of potential research, taken together, would provide the field 
with a very real look at what is happening in the classrooms the research is intended to
benefit, and would do much to strengthen the connection between applied linguistics research and the instructional practices it seeks to improve.
REFERENCES


study among Dutch students in grades 6, 8 and 10. *Language Learning, 48*, 71–106


Vandergrift, L. (2002). ‘It was nice to see that our predictions were right’: Developing metacognition in L2 listening comprehension. *The Canadian Modern Language Review, 58*(4), 555-575.


# APPENDIX A: METACOGNITIVE AWARENESS LISTENING QUESTIONNAIRE

<table>
<thead>
<tr>
<th>Strategy Belief/Perception</th>
<th>Strongly Agree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Before I start to listen, I have a plan in my head for how I am going to listen.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>2. I pay more attention to the text when I have trouble understanding.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>3. Listening in English is more difficult than reading, speaking, or writing in English.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>4. I translate in my head as I listen.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>5. I use the words I understand to help me with the words I don’t understand.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>6. While listening, if I realize I am thinking about other things, I pay attention again right away.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>7. As I listen, I compare what I understand with what I already know about the topic.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>8. It is a challenge for me to understand when I listen in English.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>9. I use my experiences and knowledge to help me understand.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>10. Before listening, I think of similar things that I have listened to before.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>11. I translate key words as I listen.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>12. When I realize I am not paying attention I quickly try to pay attention again.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>13. As I listen, my interpretation quickly changes once I realize that it was not correct.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>14. After listening, I think back to how I listened, and about what I might do differently next time.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>15. I am not nervous when I listen to English.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>16. When I have difficulty understanding what I hear, I give up and stop listening.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>17. I use the general idea of the text to help me guess the meaning of the words I don’t understand.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>18. I translate word-by-word as I listen.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>19. When I guess the meaning of a word, I think back to everything I have heard to check my guess.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>20. As I listen, I ask myself if I am happy with my level of understanding.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
<tr>
<td>21. I have a goal in my mind as I listen.</td>
<td>1 2 3</td>
<td>4 5 6</td>
</tr>
</tbody>
</table>
APPENDIX B: OPINION SURVEY

1. Click “Yes” below if you agree to have the information in this survey used for research.

a. Yes

2. Which Listening & Speaking class are you in?

a. AO1 (Mary’s 10:00-11:00 class)
b. AO2 (Mark’s 9:00-10:00 class)
c. AO3 (Mary’s 9:00-10:00 class)

3. How much do you think your listening ability has improved this quarter?

a. I don’t know.
b. It has not improved.
c. It has improved a little bit.
d. It has improved somewhat (so-so).
e. It has improved a good deal.
f. It has really improved a lot!

*Learners who answered (a) were directed to #7.
Learners who answered (b) were directed to #4.
Learners who answered (c), (d), (e), or (f) were directed to #5.

4. If you answered b:

Why do you think you weren’t able to improve?
________________________________________________________________________

How could your next listening class help you improve?
________________________________________________________________________
5. **If you answered d, e, or f:**

Why do you think your listening ability improved? (check all that apply)

- a. I practiced listening on my own outside of class.
- b. I practiced listening in homework activities outside of class.
- c. I practiced listening during in-class activities.
- d. I learned useful things about how to listen more effectively (better) in my listening class.
- e. Other __________________________________________________________________________

6. Why was this/were these helpful?

________________________________________________________________________

7. What did you do in your listening class or for homework for your listening class? (check all that apply)

- a. practiced listening to audio cds / audio recordings on the internet
- b. practiced listening to videos
- c. practiced listening to my classmates
- d. practiced listening to audio/video and answering questions about what I heard
- e. practiced listening to audio/video and wrote summaries about what I heard
- f. practiced listening to audio/video and talked about what I heard
- g. practiced listening skills (strategies) like guessing (inferencing), making predictions, using what I already know (activating background knowledge), etc.
- h. Other __________________________________________________________________________

8. Which ONE of these (a, b, c, d, e, f, g, or h) was the MOST helpful?

________________________________________________________________________

9. Why?

________________________________________________________________________

10. What were your favorite in-class (or homework) listening practice activities? Why?

________________________________________________________________________
11. What were your least favorite in-class (or homework) listening practice activities? Why?

________________________________________________________________________

12. AT THE BEGINNING OF THE WINTER QUARTER, how much time DID you spend practicing listening outside of class (*on your own*)?

*on your own = watching movies, watching TV, listening to the radio, talking to Americans, listening to Americans talk to each other, practicing listening on the Internet, etc.

a. 0-30 minutes a week __ 
b. 30 minutes—1 hour a week __
c. 1 hour – 2 hours a week __
d. 2 or more hours a week _

13. NOW, at the END of the WINTER QUARTER, how much time do you spend practicing listening outside of class (*on your own*)?

*on your own = watching movies, watching TV, listening to the radio, talking to Americans, listening to Americans talk to each other, practicing listening on the Internet, etc.

a. 0-30 minutes a week __ 
b. 30 minutes—1 hour a week __
c. 1 hour – 2 hours a week __
d. 2 or more hours a week _

14. Did this Listening & Speaking class influence/affect how much time you spent practicing listening outside of class?

a. Yes __________ (Go to Question #15)
b. No __________ (survey complete)

If you answered YES to #14:

15. What about your Listening & Speaking class caused the change?
APPENDIX C: CURRICULAR OBJECTIVES FOR INTERMEDIATE LEVEL

LISTENING INSTRUCTION

| RECOGNITION | 1. Can identify/discriminate phonemes, inflectional markers, and question forms in isolated words and carefully spoken sentences |
|            | 2. Can identify/discriminate basic stress and intonation patterns |

| COMPREHENSION | 1. Associates appropriate meaning with grammatical forms |
|               | 2. Associates appropriate meaning with basic stress and intonation patterns |
|               | 3. Comprehends the gist of a short narration, description, or dialogue which contains a modest amount of unfamiliar language |

| COMPREHENSION: INTERACTIONAL | 1. Comprehends contractions and reductions |
|                             | 2. Comprehends main ideas and most details in conversational discourse on a variety of topics beyond the immediate situation |

| COMPREHENSION: TRANSACTIONAL | 1. Understands key words in a short piece of extended discourse |
|                            | 2. Identifies and understands some transition signals in a short piece of extended discourse |
|                            | 3. Identifies and understands some cohesive devices in a short piece of extended discourse |
|                            | 4. Identifies and comprehends thought groups in a short piece of extended discourse |
|                            | 5. Understands the basic organizational pattern of a short mini-lecture |
|                            | 6. Responds to questions based upon discourse which contains little or no unfamiliar language |
|                            | 7. Given a pre-listening activity, discerns specific information needed to respond to follow-up questions or tasks |

Adapted from: AE45 Listening & Speaking Curriculum Guide
APPENDIX D: SAMPLE STRATEGY HANDOUT

Making Predictions & Activating Background Knowledge

MAKING PREDICTIONS

→ Thinking of the words, phrases, and information that you might hear.

EXAMPLE: You want to buy a coffee. Think about what you need to say, and predict what the cashier will say to you.

WHY: Predicting what you might hear makes it easier to understand when you are listening.

WHEN: Use prediction when you have knowledge of the topic. When you get new information (during listening) you can change your predictions or make new ones.

ACTIVATE BACKGROUND KNOWLEDGE

→ Bringing information that you already know about a topic into your mind before you listen.

EXAMPLE: Your teacher tells you that you will listen to a radio program about fast food. Think about everything you know about fast food: where you buy it, how much it costs, how it tastes, how healthy it is, how popular it is, etc.

WHY: Thinking about what you already know helps you get ready to listen. Having this information in your mind makes it easier to understand new information coming in while you listen.

WHEN: Activate background knowledge when you know what the topic will be, and you already know something about the topic.

Adapted from information presented in Chamot et al. (1999) and Vandergrift (1997)
<table>
<thead>
<tr>
<th></th>
<th>My Predictions</th>
<th>First Listening</th>
<th>Second Listening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topics I might hear:</td>
<td></td>
<td>I also understood:</td>
<td>Now I need to focus on:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Words I might hear:</td>
<td></td>
<td>My partner understood:</td>
<td>This time I understood:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### APPENDIX F: EVENT SAMPLING OBSERVATION CHECKLIST

<table>
<thead>
<tr>
<th>STRATEGY</th>
<th>Strategy Mentioned/Presented/Referenced but Not Practiced</th>
<th>Strategy Used/Practiced but Unnamed</th>
<th>Strategy Used/Practiced and Named</th>
<th>Learners Informed of Strategy's Value</th>
<th>Learners Informed of Other Potential Uses</th>
<th>Evidence that Tchr Support Gradually Removed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>METACOGNITIVE: PLANNING</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal Setting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activate Background Knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Directed Attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selective Attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>METACOGNITIVE: MONITORING</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension Monitoring</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-Check Monitoring</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>METACOGNITIVE: EVALUATING</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verify Guesses/Predictions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problem Identification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Check Goals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>COGNITIVE: ELABORATION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Elaboration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>World Elaboration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questioning Elaboration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imagery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>COGNITIVE: INFERENCING</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linguistic Inferencing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voice/Paralinguistic Inferencing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Parts Inferencing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Event Sampling Observation Checklist: Continued

<table>
<thead>
<tr>
<th>STRATEGY</th>
<th>Strategy Mentioned/Presented/Referenced but Not Practiced</th>
<th>Strategy Used/Practiced but Unnamed</th>
<th>Strategy Used/Practiced and Named</th>
<th>Learners Informed of Strategy's Value</th>
<th>Learners Informed of Other Potential Uses</th>
<th>Evidence that Tchr Support Gradually Removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>COGNITIVE: INDUCTION/DEDUCTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COGNITIVE: SUMMARIZATION</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COGNITIVE: RESOURCING</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COGNITIVE: GROUPING</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOCIOAFFECTIVE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooperation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questioning for Clarification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowering Anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-Encouragement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EVENT</th>
<th>SPONTANEOUS STUDENT-LED OCCURANCE</th>
<th>SPONTANEOUS TEACHER-LED OCCURANCE</th>
<th>PLANNED INTO THE LESSON</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREDICTION/PRE-LISTENING</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher Informs about Content</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prediction of Content</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prediction of Specific Lexical Items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predictions Written Down by Teacher</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Event Sampling Observation Checklist: Continued

| Predictions Written Down by Ss |  |  |
| Predictions Verified/Revised/Rejected |  |  |

**EXPOSURE TO TEXT**
- One Listening Per Text
- Two Listenings Per Text
- 3+ Listenings Per Text
- Learners Collaborate Between Listenings
- Subsequent Listenings Used for Verification of Previous

**BOTTOM-UP PROCESSING**
- Listen and Read Along
  - Segmenting words from speech Stream/Listening for Key Words
  - Listening for Inflectional Markers/Phonemes
  - Listening for Stress/Intonation
  - Prior to Other Listening Activities
  - After Other Listening Activities

**COLLABORATION**
- Answers Checked with Partner
- Partners Collaborate to Solve Problems
- Whole Group Collaborates to Solve Problems
### DEMONSTRATION OF COMPREHENSION

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple-Choice Questions</td>
<td></td>
</tr>
<tr>
<td>Fill-in-the-blank</td>
<td></td>
</tr>
<tr>
<td>Open-ended questions</td>
<td></td>
</tr>
<tr>
<td>Chart Completion</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>Individual Reconstruction</td>
<td></td>
</tr>
<tr>
<td>Reconstruction with a Partner</td>
<td></td>
</tr>
<tr>
<td>Whole Group Reconstruction</td>
<td></td>
</tr>
<tr>
<td>Reconstruction with Teacher Intervention</td>
<td></td>
</tr>
</tbody>
</table>

### COMPREHENSION BREAKDOWNS

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ignored</td>
<td></td>
</tr>
<tr>
<td>Correct Answer Provided</td>
<td></td>
</tr>
<tr>
<td>Teacher Gives Explanation</td>
<td></td>
</tr>
<tr>
<td>Other Students Explain</td>
<td></td>
</tr>
<tr>
<td>Turn to Tapescript</td>
<td></td>
</tr>
<tr>
<td>Listen Again (Strategically)</td>
<td></td>
</tr>
</tbody>
</table>

### CLASS DISCUSSIONS

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE: Content of the Text</td>
<td></td>
</tr>
<tr>
<td>RE: Strategy Use/Effectiveness</td>
<td></td>
</tr>
</tbody>
</table>

### LEARNER REFLECTION

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learners Discuss Listening Skill</td>
<td></td>
</tr>
<tr>
<td>Reflection Related Homework Given/Received</td>
<td></td>
</tr>
</tbody>
</table>
Title of Research: Listening Comprehension Research

Researcher: Mary Freday

You are being asked to participate in research. This form describes the purpose, procedures, possible benefits, and risks. It also explains how your personal information will be used and protected. Once you have read this form and your questions about the study are answered, you will be asked to sign it. This will allow you to participate in this study.

EXPLANATION OF STUDY

This study is about ways improving second language listening. If you agree to participate, you will complete a survey 2 times during the quarter about what you do while you listen in English, and you will be asked to participate in a focus group interview at the end of the quarter. Your Michigan Placement Test listening scores from the beginning and end of the quarter will also be used. Your participation in this study will be during the Winter Quarter.

Risks and Discomforts

No risks or discomforts are anticipated.

Benefits

This study is important to society because listening is an important skill and teachers need to know how to help students improve their listening abilities. Individually, you may benefit. Your listening comprehension might improve.

Confidentiality and Records

Your study information will be kept confidential. Your name will not be recorded and your listening test scores and survey results will be anonymous.

Additionally, while every effort will be made to keep your study-related information confidential, there may be times where this information must be shared with:

* Federal agencies, for example the Office of Human Research Protections, whose responsibility is to protect human subjects in research;
* Representatives of Ohio University (OU), including the Institutional Review Board, a committee that oversees the research at OU;
* Dr. Greg Kessler, Linguistics Department (chair of the thesis committee)

**Contact Information**

If you have any questions regarding this study, please contact Mary Freday. Office: Gordy 346; Phone: 740-707-6936; email: ml168709@ohio.edu

If you have any questions regarding your rights as a research participant, please contact Jo Ellen Sherow, Director of Research Compliance, Ohio University, (740)593-0664.

By signing below, you are agreeing that:

- you have read this consent form and have been given the opportunity to ask questions and have them answered
- you have been informed of potential risks and they have been explained to your satisfaction.
- you understand Ohio University has no funds set aside for any injuries you might receive as a result of participating in this study
- you are 18 years of age or older
- your participation in this research is completely voluntary
- you may leave the study at any time. If you decide to stop participating in the study, there will be no penalty to you and you will not lose any benefits to which you are otherwise entitled.

Signature________________________ Date_______

Printed Name________________________

Version Date: 12/13/2010
IRB Approval

A determination has been made that the following research study is exempt from IRB review because it involves:

Category 1: research conducted in established or commonly accepted educational settings, involving normal educational practices

Project Title: In Support of Strategic Listening: A Pedagogical Cycle for Teaching the Listening Process, Increasing Metacognitive Awareness, and Improving Comprehension

Primary Investigator: Mary Elizabeth Freday

Co-Investigator(s): 

Advisor: Greg Kessler

Department: Linguistics

Rebecca Cale, AAB, CIP
Office of Research Compliance

Date: 11/09/10

The approval remains in effect provided the study is conducted exactly as described in your application for review. Any additions or modifications to the project must be approved (as an amendment) prior to implementation.
APPENDIX H: COPYRIGHT PERMISSION

From: Mary Lynch [mailto:wanderingwonderwoman@yahoo.com]
Sent: April-23-11 2:21 PM
To: Larry Vandergrift
Subject: Copyright Permission

Dear Dr. Vandergrift,

My name is Mary, and in case you don't remember me, I contacted you in February regarding scoring and interpretation of the MALQ. I am now in the process of writing my thesis, and I realized that I may need copyright permission from you for some of the items in the document. First, I used the MALQ, but I adapted it slightly to increase its comprehensibility for the ESL learners in this context. I have attached the adapted questionnaire to this message. Second, I am using a (very) adapted version of Figure 1: Listening Comprehension Strategies and their Definitions with Representative Examples (Vandergrift, 1997) in my literature review, as well as an adapted version of Figure 1: Listening Instruction Stages and Related Metacognitive Strategies (Vandergrift, 2004) in my methodology chapter. My adapted versions of both of these figures are also attached to this email.

I would sincerely appreciate you granting me permission to utilize each of these items in my thesis, as they were all instrumental in my research project and are valuable and important parts of the document.

Thank you very much for your time (and for your research!!)

Have a great weekend,

Mary

From: Larry Vandergrift <lvdgrift@uottawa.ca>
View Contact
To: Mary Lynch <wanderingwonderwoman@yahoo.com>

No problem from my perspective. Best wishes on your research, lvg