In: Applied Memory
Editor: Matthew R. Kelley

ISBN 978-1-60692-143-2

Chapter 2

STUDYING WITH MUSIC: IS THE IRRELEVANT SPEECH EFFECT RELEVANT?

Justin Kantner
Department of Psychology, University of Victoria
Victoria, BC, Canada

ABSTRACT

The detrimental impact of background speech on short-term serial recall, known as the irrelevant speech effect, has a long history as a valuable theoretical tool, but its applications have received less attention. Two experiments attempted to determine whether the effect would extend to the practice of studying for an exam in the presence of background music containing lyrics. In Experiment 1, participants listened to vocal or instrumental music while studying for a serial recall test under conditions designed to mimic a realistic study-test experience. Despite several methodological departures from the traditional irrelevant speech paradigm, the results were consistent with a classic irrelevant speech effect. In Experiment 2, highly naturalistic study materials and test questions were used to determine the generality of the irrelevant speech effect beyond short-term serial recall. Vocal music hurt test performance relative to instrumental music under these conditions, but only when the study materials consisted of a chart; when the materials were prose paragraphs, no irrelevant speech effect was found. Explanations for this pattern of results are considered in terms of processing strategies at study, and further aspects of a realistic examination of studying with background music are discussed.

INTRODUCTION

The irrelevant speech effect—the tendency for ambient human speech to impair short-term serial recall—began, and has remained, a quintessential laboratory phenomenon in the human memory literature (e.g., Colle & Welsh, 1976; Salamé & Baddeley, 1982). Generated almost exclusively within a single task domain using simple, highly controlled stimulus materials, irrelevant speech effect data are channeled largely towards theory development and assessment. Among the classic memory phenomena with long, fruitful careers under

laboratory scrutiny (e.g., the phonological similarity effect, the word length effect, articulatory suppression), the irrelevant speech effect is perhaps one of those most intuitively suited to real-world application. Despite its identity as a short-term serial recall phenomenon, the effect speaks to an utterly pervasive facet of everyday life—namely, environmental sound distraction. Thus, its practical implications for cognitive performance in applied settings are presumed to be considerable (Banbury, Tremblay, Macken, & Jones, 2001; Beaman, 2005; Jones, 1999).

The present investigation asks whether background music containing vocals can impair efforts to study for an exam by causing an irrelevant speech effect. The scenario of listening to music while studying for an exam has seldom been touched upon in past research, despite the fact that it bears a strong resemblance to a typical irrelevant speech experiment in many respects. Background music is analogous to the irrelevant speech played in traditional experiments, and both natural and experimental study tasks involve committing information to memory in the presence of a task-irrelevant auditory stream. Moreover, the population most often studied in irrelevant speech experiments is the one to which the current application is most directly relevant: college students. Exam preparation is a nearly perpetual matter for undergraduates, and listening to background music while studying is a popular habit. Given that the primary (if not exclusive) goal of studying is to encode information into memory ahead of an anticipated test, the potential of lyrics in vocal background music to constitute irrelevant speech, and thus to disrupt memorization, is a significant applied question.

THE IRRELEVANT SPEECH EFFECT IN THE LABORATORY

The vast majority of work on the irrelevant speech effect has concerned the effect of irrelevant speech on immediate (or short-term) serial recall. In a typical experiment, short lists of simple stimuli such as digits, letters, or words are presented one-at-a-time. Following the offset of the final list item (or a short delay thereafter), participants attempt to recall the list items in their presentation order. When human speech is played during list presentation, or during a subsequent pre-test rehearsal period (Miles, Jones, & Madden, 1991), serial recall performance suffers. This decrement—the irrelevant speech effect—has been reported with magnitudes ranging from 5% to 50% in the literature (Neath, 2000). In this short-term serial recall paradigm, the irrelevant speech effect is an extremely robust phenomenon, replicated numerous times with a wide variety of materials, designs, and procedures (e.g., Colle, 1980; Colle & Welsh, 1976; Ellermeier & Zimmer, 1997; Hanley & Broadbent, 1987; Jones, 1994; Jones, Madden, & Miles, 1992; Neath, Surprenant, & LeCompte, 1998; Salamé & Baddeley, 1982, 1986, 1989; Surprenant, LeCompte, & Neath, 2000).

One of the intriguing aspects of the irrelevant speech effect is that the speech can be deleterious to memory without being related in any way to the to-be-remembered material (hence the term "irrelevant"). Indeed, the first published report of the effect by Colle and Welsh (1976) demonstrated that it can be produced even when the speech is in a foreign language. In general, the content of the speech does not have much bearing on the magnitude of the effect (e.g., Buchner, Irmen, & Erdfelder, 1996), though some exceptions have been reported (Buchner & Erdfelder, 2005; Buchner, Rothermund, Wentura, & Mehl, 2004; Neely & LeCompte, 1999). The effect is largely indifferent to the phonological overlap of the

irrelevant speech and the to-be-remembered material (Jones & Macken, 1993, 1995; though see Salamé & Baddeley, 1982) and the volume at which the speech is played (Colle, 1980; Salamé & Baddeley, 1987). In sum, the evidence to date characterizes the irrelevant speech effect as a highly consistent phenomenon, robust with respect to numerous physical and psychological manipulations, and resistant, even, to habituation (Hellbrück, Kowano, & Namba, 1996; Tremblay & Jones, 1998).

The impenetrability of the effect has fostered great interest in its cognitive substrates. Examinations of the effect and its interactions with other classic short-term memory phenomena, such as the phonological similarity effect, the word length effect, and articulatory suppression have led to important insights as to the factors that affect short-term memory (e.g., Jones, Macken, & Nicholls, 2004; Neath, Farley, & Surprenant, 2003). The irrelevant speech effect has played an important role in model development; these models, in turn, have driven an ongoing debate as to how the effect is best explained (see Baddeley, 2000, Jones & Tremblay, 2000, and Neath, 2000, for one exchange, and Baddeley & Larsen, 2003, and Macken & Jones, 2003, for another).

An early and highly influential account of the effect was provided by Baddeley's Working Memory framework (Baddeley, 1986; Baddeley & Hitch, 1974; Salamé & Baddeley, 1982), which holds that irrelevant speech gains automatic access to a Phonological Store, where it interferes with to-be-remembered information also in the store. By contrast, the Object-Oriented Episodic Record (O-OER) model (Jones, 1993; Jones & Macken, 1993; Jones, Macken, & Murray, 1993) posits that interference from irrelevant speech occurs via automatically generated "pointers" that represent the serial order of any dynamic stimulus stream. The pointers from irrelevant speech interfere with those from the to-be-remembered information, which causes the loss of serial order information. Alternatively, the Feature Model, originally proposed by Nairne (1988, 1990) and augmented to account for the irrelevant speech effect by Neath (2000), characterizes the disruption in terms of feature adoption (i.e., the features of the irrelevant speech become part of the representation of to-be-remembered information) and divided attention (i.e., between the to-be-remembered information and the to-be-ignored speech). These varying conceptions of the phenomenon—and of short-term memory, more generally—give rise to competing predictions regarding its interaction with a host of stimulus manipulations.

A close examination of these theories and their predictions is beyond the scope of this work (see Neath & Surprenant, 2003, for a review). This purpose of the present chapter is to establish the generality of the irrelevant speech effect to a more realistic study-test scenario than is usually in place in irrelevant speech experiments—specifically, studying for an exam with music playing in the background. If background speech impairs serial recall performance, then background music containing speech (i.e., lyrics) could reasonably be expected to have a similar effect in this applied setting.

MUSIC AS IRRELEVANT SPEECH

While a number of studies have explored the effects of music on cognitive performance (e.g., Fogelson, 1973; Freeburne & Fleischer, 1952; Furnham & Bradley, 1997; Hallam, Price, & Katarsou, 2002; Ransdell & Gilroy, 2001), few have tested the potential for music to

produce an irrelevant speech effect. Demonstration of a musical irrelevant speech effect requires at least one vocal and one non-vocal (i.e., instrumental) background condition, so that memory performance in each may be compared. Past studies employing this design have provided strong evidence that vocals in music do produce a memory deficit closely resembling an irrelevant speech effect. Boyle and Coltheart (1996) presented participants with 50 10-item lists, half of which contained rhyming words and half of which contained non-rhyming words. Serial recall was initiated after each list. Study and recall of each list were accompanied by one of five auditory conditions: silence, instrumentals from Gilbert and Sullivan songs, singing plus instrumentals from the same songs, singing from those songs alone, and the lyrics from those songs read as normal speech. Boyle and Coltheart (1996) found that each condition containing speech (singing, singing plus instrumentals, and speech) significantly impaired serial recall performance, and to roughly the same degree. Instrumental music produced a directional but non-significant deficit in performance. Similarly, Salamé and Baddeley (1989) tested serial recall of random digit strings while participants listened to instrumental or vocal versions of opera pieces. Vocal music significantly lowered performance relative to both instrumental music and silence.

Salamé and Baddeley (1989) discussed the implications of their findings for settings such as school and the workplace, and suggested that further research should test the generalizability of their findings in such environments. The most direct extension of the Salamé and Baddeley (1989) paradigm beyond serial recall, however, was arguably that of Martin, Wogalter, and Forlano (1988), published one year earlier. In their Experiment 2, Martin et al. (1988) tested vocal versus instrumental music conditions while having participants read short passages on various topics (e.g., the propagation of sound, the creative process) from a GRE practice guide. Reading comprehension of each passage was tested after a 30-sec delay interval. The music consisted of the song "You Light Up My Life" by Joe Brooks; in the instrumental condition, the music from the song was played without lyrical accompaniment, while in the lyrics conditions, the words to the song were either sung or read by a vocalist from the university's music department, with and without backing instrumentation. Reading comprehension test scores were 6-7% lower for passages read in the lyrics conditions than for those read in the instrumental or silence conditions, a significant difference (Martin et al., 1988). Although these results do not indicate a large effect size, and are limited to a single experiment utilizing a single musical piece, they dovetail with those of Boyle and Coltheart (1996) and Salamé and Baddeley (1989) in suggesting that vocal music can cause an irrelevant speech effect. Further, they suggest that the effect might be felt under conditions more closely approximating a realistic study and test procedure than those generally used in irrelevant speech experiments.

EXPERIMENT 1

The goal of Experiment 1 was to begin to bridge the gap between the traditional irrelevant speech paradigm and a naturalistic examination of the effects of vocal music on studying for an exam. Therefore, Experiment 1 was built upon the list-learning, serial recall design of the Boyle and Coltheart (1996) and Salamé and Baddeley (1989) studies, but added several features meant to reflect a more realistic learning scenario. First, the nature of the

stimuli was changed. Lists in typical irrelevant speech experiments are composed of simple materials such as digits, letters, and words. Not only are such stimuli decidedly more elemental than those that would be learned in any university course, their inter-relationships differ from those of regular learning materials in ways likely to modulate the impact of irrelevant speech. That is, digit, letter, and word stimuli usually have no meaningful semantic relationship to one another. Lists of unrelated items are particularly vulnerable to irrelevant speech effects, because participants are likely to rely on phonological processing strategies such as subvocal rehearsal to facilitate learning (Jones & Tremblay, 2000). By contrast, related items may be processed in terms of meaning, leading to a diminished irrelevant speech effect. Hanley and Bakopoulou (2003) directly manipulated the learning strategy used by participants, instructing some to adopt a phonological rehearsal strategy and others to process items semantically. While the phonological group demonstrated a classic irrelevant speech effect (with an additive effect of phonological similarity), the semantic group exhibited a small, non-significant irrelevant speech effect, and no additional decrement from phonological similarity.

These results hold clear implications for the question of whether irrelevant speech in music will be detrimental to learning exam material. While the stimulus lists used in most irrelevant speech research consist of unrelated items, learning material in academic coursework is predominantly composed of highly inter-related information. The latter, then, might elicit semantic processing and be less prone to interference through irrelevant speech. In Experiment 1, study lists were composed of actual names, periods, or scientific terms organized according to a particular sequential order. For example, one list, labeled "20th Century U.S. Presidents (in the order they held office)" contained the items ROOSEVELT, TRUMAN, EISENHOWER, KENNEDY, JOHNSON, NIXON, and FORD. Thus, the items of each list adhered to a meaningful consecution, made apparent by the preceding list label.

Second, the music played was intended to approximate what current undergraduates might listen to independently. That is, modern themes and instrumentation were emphasized, and an array of songs was chosen, in order to simulate the experience of listening to a CD, iPod, or radio station while studying. The selection of songs was at the discretion of the researcher, and no claim is made that the songs used in the experiment are, themselves, commonly listened to by the participants. Indeed, the objective was to choose songs performed by current recording artists that were generally not "household names" amongst 18-22 year-olds, so as to minimize individual differences in familiarity with the music.

For the non-lyrical comparison condition, instrumental tracks were chosen from predominantly vocal albums of roughly the same genre as those used in the lyrics condition. This selection strategy was intended to help gather a set of instrumental songs with similar instrumentation, tempo, and length to those of the vocal songs. The highly complex nature of the range of songs used made precise matching of the vocal and instrumental tracks along these dimensions infeasible, but this was considered a reasonable exchange for the ecological validity of the materials.

Third, whereas most irrelevant speech experiments involve repeated study-test cycles wherein lists are recalled either immediately or shortly after presentation, academic exam preparation generally entails repeated review of all to-be-remembered material, followed by a single test. To simulate these factors in Experiment 1, participants studied each of eight lists on three separate occasions, and received only one test, on all of the lists, following the study phase.

Finally, music was played only at study; tests were taken in silence. This aspect of the format again mimicked that of academic testing, and ensured that participants would not derive a musical benefit of state-dependent memory at test (Balch, Bowman, & Mohler, 1992). In a third (control) condition, both study and test were undertaken in silence.

Method

Participants. Participants were 82 undergraduates at the University of Victoria who took part in the experiment in exchange for bonus credit in an introductory psychology course. Twenty-eight participated in the lyrics condition, 27 participated in the instrumental condition, and 28 participated in the silence condition. All were fluent in English.

Materials. Eight lists of seven items each were created by the author. Lists were composed of real-life examples of names and terms occurring in a natural sequence. The eight lists were 20th Century U.S. Presidents (in the order they held office), Metallic Elements (in order of appearance in the Periodic Table), Layers of the Atmosphere (from highest to lowest altitude), the Caesars (in the order they ruled), Canadian Provinces (in the order of their incorporation into Canada), Geological Eras (from most recent to least recent), and Shakespearean Plays (in order of date of first performance). The full lists appear in Appendix A.

Vocal music selections were assembled by the author and consisted of various tracks from recent albums loosely definable as rock or pop. Songs were selected that contained contemporary instrumentation (e.g., drums and guitars) and themes but were not likely to be known to most first- or second-year University of Victoria undergraduates. These songs contained innocuous lyrics and instrumentals that were not likely to be perceived as overly heavy or grating. Ten vocal music tracks were used in the experiment, half containing male vocals and half containing female vocals. Ten non-vocal music selections were chosen on the basis of their similarity to the vocal tracks in instrumentation, tempo, and genre. In most cases, these tracks were drawn from instrumental pieces on albums predominantly composed of vocal rock or pop songs. A list of the songs used in the experiment appears in Appendix B. The music was played at a moderate volume level through stereo headphones.

Instructions and lists were presented via E-Prime software. Paper test forms were created and were completed with pen or pencil.

Procedure. Participants were tested one-at-a-time, with an experimenter present throughout the session. Instructions informed participants that they would study eight lists of seven items each and that they should try to learn each in its proper order. Participants were encouraged to approach the task as though they were studying for an exam on which the order of the lists would be tested. Each list was preceded by a label in a blue font indicating the topic of the list to follow and the organizational structure of the list (e.g., "20th Century U.S. Presidents (in the order they held office)"). Participants pressed the space bar after reading the list topic to initiate the presentation of the list. Each list item appeared in black lettering in the center of the computer screen for 2 sec. Following the presentation of the final item of a given list, the screen went blank for 3 sec, followed by the topic of the next list. While the list items appeared in their fixed order, the lists themselves appeared in a random order.

After the presentation of all eight lists, a screen appeared indicating that the lists would be presented again and that the task was to continue to memorize them in order. The

procedure of this second study block was identical to that of the first study block, except that the lists were presented in a new random order.

Following the second study block, a screen appeared indicating that each list would be presented for a final time, and that participants should make their best effort to memorize the order of each before the subsequent test. They were also encouraged to try and anticipate items in the third study block before they appeared on the screen as a check on their knowledge. The third study block was identical to the first two, except that the lists were presented a new random order, and the parenthetical component of the list topic was dropped, such that only the topic itself appeared on the screen (e.g., "20th Century U.S. Presidents").

All participants wore headphones throughout the study phase. Participants in the lyrics condition heard a randomly ordered assortment of as many of the ten vocal music songs as fit within the duration of the study session. Participants in the instrumental condition heard an assortment of the instrumental tracks, while participants in the silence condition heard nothing.

Following the third study block, participants removed the headphones and received test instructions. These instructions created a delay of roughly 30 sec between the end of the study phase and the administration of the test. Participants were then given a test sheet containing eight rows with seven blanks each. The first list to be recalled was indicated via the appearance of that list's topical label in the center of the screen. After participants had written down the items from the list to their satisfaction, they pressed the space bar to bring up the next list to be recalled. The eight lists were probed in a random order and subjects were given as much time as needed to recall each one.

Upon conclusion of the test, all participants were asked how many of the lists they felt they could have ordered with 100% accuracy before taking part in the experiment. None reported more than one such list out of the eight studied. Of the nine participants reporting extra-experimental knowledge of one list, four were in the lyrics condition, four were in the instrumental condition, and one was in the silence condition.

Results and Discussion

Prior to analysis, the data of three participants were removed due to exceedingly low test scores and/or a failure to understand task instructions. Thus, the analyses that follow are based on the data from 27, 26, and 27 participants in the lyrics, instrumental, and silence conditions, respectively.

Mean accuracy as a function of serial position is plotted for each of the three conditions in Figure 1. As is evident from inspection of the figure, performance in all three conditions was superior for the items at the beginning and end of the lists, yielding the classic bow-shaped serial position curve well known to characterize memory for order (Murdock, 1960). A repeated measures ANOVA with serial position (1-7) as a within-subjects factor and listening condition (lyrics, instrumental, or silence) as a between-subjects factor revealed a highly significant main effect of serial position, $F(6, 456) = 45.4$, $p < .001$. The serial position effect did not interact with listening condition ($F < 1$).

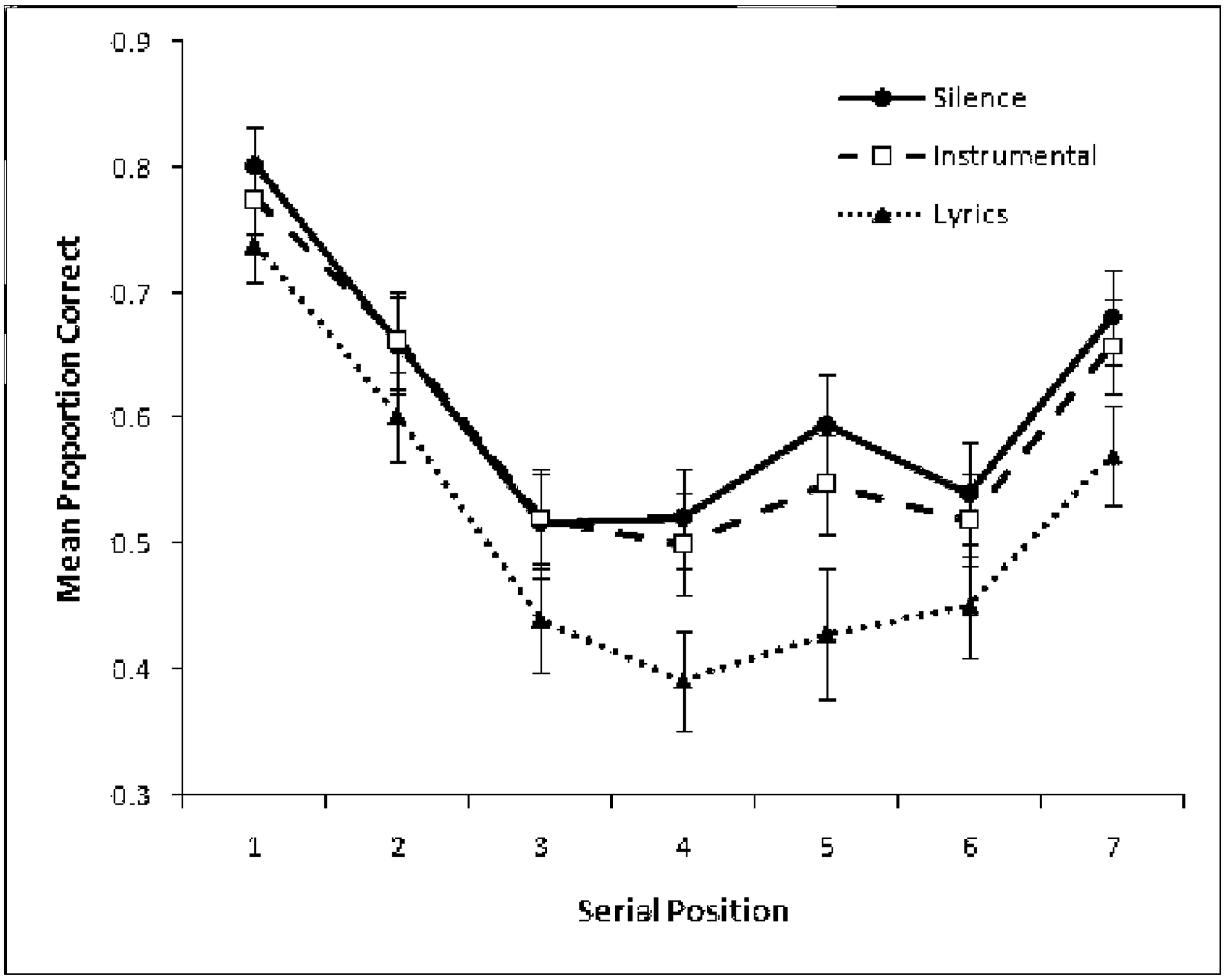

Figure 1. Serial recall performance as a function of item list position (1-7) and listening condition during study in Experiment 1. Error bars represent one standard error of the mean.

Participants in the lyrics condition performed less accurately than those in the other two conditions at every list position. Collapsed across serial positions, the mean percentage of correct responses was 61.6 for the silence group, 59.8 for the instrumental group, and 51.7 for the lyrics group. The main effect of listening condition was significant, $F(2, 76) = 3.372$, $p < .05$. Planned comparisons confirmed that the silence, $t(50) = 2.02$, $p < .05$, and instrumental, $t(51) = 1.86$, $p < .05$, groups were significantly more accurate than the lyrics group, but did not differ significantly from one another, $t(51) = 0.83$, $p = .53$.

In short, these results mirror those of more traditional irrelevant speech experiments in two important respects. First, all three conditions displayed bow-shaped serial position curves, a well-established pattern of accuracy in the irrelevant speech paradigm. Second, performance in the lyrics condition was impaired relative to the instrumental and silence conditions at a magnitude quite consistent with more typical irrelevant speech studies. That this correspondence between past results (e.g., Boyle & Coltheart, 1996; Salamé & Baddeley, 1989) and those of the current study was observed despite substantial design modifications in the present experiment suggests that the irrelevant speech effect may indeed generalize to the practice of studying with background music.

EXPERIMENT 2

The results of Experiment 1 suggest that lyrics in music can function as irrelevant speech and can impair serial recall performance. While the current experiment represents an extension of past work on the irrelevant speech effect, there are still many respects in which it does not reflect the circumstances of realistic exam preparation, nor a realistic academic test. Perhaps the most obvious is that students are not usually tested solely on rote serial recall. Instead, they are given questions that call for multiple types of mnemonic judgments, such as free recall (as in the fill-in-the-blank or short answer questions) and recognition (as in true/false or multiple choice questions).

Whether irrelevant speech would impact performance on these types of tasks, given naturalistic study materials, is an open question. The O-OER model (Jones, 1993) predicts that only tasks with a serial component (i.e., serial rehearsal of study items) will be subject to an irrelevant speech effect. Neither the Phonological Store hypothesis (Salamé & Baddeley, 1982) nor the Feature Model (Neath, 2000) makes this explicit prediction (Farley, Neath, Allbritton, & Surprenant, 2007). Generally, however, the evidence has favored the hypothesis that seriation is necessary for an irrelevant speech effect (Farley et al., 2007).

LeCompte (1994) demonstrated irrelevant speech effects in free recall, cued recall, and recognition using letter and word stimuli, which suggests that that the effect extends well beyond serial recall to tasks eliciting no serial processing. Beaman and Jones (1997), however, have argued that such tasks may elicit a serial rehearsal strategy, even though the goal is not serial recall per se (see also Jones and Tremblay, 2000). That is, although participants are not asked to memorize the order of to-be-remembered items, they may find rehearsal according to presentation order the best mnemonic strategy at their disposal.

By this view, the effects of irrelevant speech in music may be felt when studying for tests other than serial recall. To the extent that seriation is a prominent component of study strategies, one might expect irrelevant speech to have a negative impact on free recall, recognition, and other types of tests often represented on academic exams. Whether or not students engage in serial processing, however, may depend largely on the nature of the study materials. The serial rehearsal strategy encouraged by a list learning task, for example, may not be adopted in more natural learning situations, such as when one is studying from a textbook. Moreover, the processing strategy chosen is likely to follow from one's expectations as to the nature of the test questions. In a typical irrelevant speech experiment (and in the current Experiment 1), participants study with fore knowledge of the exact nature of the test, and, in effect, the questions to be asked. Such is not generally the case for undergraduates outside the laboratory.

In Experiment 2, participants studied three topics: Caesars, Canadian Prime Ministers, and geological eras. Information about the Caesars and Prime Ministers was presented via multiple pages of text interspersed with pictures, a format meant to approximate the appearance of pages from a textbook. Information about the geological eras was presented in chart format, with the names of the eras and their associated flora and fauna laid out in a list-like fashion. In each case, the material to be learned was distinctly serial in nature: the Caesars and Prime Ministers were discussed, and the geological eras were listed, in temporal order. In the former two cases, however, the information was embedded in prose. If prose presentation obviates or otherwise discourages the use of a serial processing strategy

necessary for the emergence of an irrelevant speech effect, music containing lyrics should pose no disadvantage to the study of the Caesars and Prime Ministers, but might still be expected to diminish recall of the geological eras. Memory for the different topics was assessed with serial reconstruction, fill-in-the-blank, multiple choice, and recognition questions. Participants were given no foreknowledge of the general format of the test, nor the individual questions; they were simply told to try their best to learn as much of the material as possible in the time allotted.

Another shortcoming with respect to the ecological validity of Experiment 1 was addressed in Experiment 2: the delay between study and test. Although this interval (approximately 30 sec) was considerably longer in Experiment 1 than is usually the case in irrelevant speech experiments, it was unrealistic in that students are not usually given exams so soon after study. While some review often occurs immediately before the test, students often prepare in the days leading up to the exam (particularly the day before). Thus, a stronger test of the practical implications of the irrelevant speech effect would hold study and test sessions on separate days. In Experiment 2, participants were tested twice: immediately after study, and on the following day. The first test served as an immediate measure of memory and as a means of preparation for the second test. The longer-term effects of irrelevant speech under conditions such as those of Experiment 2 remain to be established, though Knez and Hygge (2002) reported a detrimental effect of background conversation on recall from a text passage after a two-hr delay. Whether the effects of irrelevant speech would maintain after an approximately 24-hr delay is an open question.

Finally, the method of sound delivery was changed in Experiment 2. Participants in Experiment 1 listened to music through headphones, perhaps rendering the music particularly difficult to ignore. To the extent that students would normally play music through a more ambient, less invasive medium (e.g., external speakers) while studying, Experiment 1 may have exaggerated its impact. In Experiment 2, music was played through speakers and was audible throughout the testing room.

Method

Participants. Participants were 42 undergraduates at the University of Victoria who took part in the experiment in exchange for bonus credit in an introductory psychology course. All were fluent in English. Twenty-two participated in the lyrics condition and 20 participated in the instrumental condition. A silence condition was not included in Experiment 2.

Study Materials. Three topics for study were chosen by the author: the early Caesars, 20th Century Canadian Prime Ministers, and geological eras. Caesar materials were adapted from a Wikipedia article summarizing the writings of the historian Suetonius (http://en.wikipedia.org/wiki/Lives_of_the_Twelve_Caesars). Text and pictures from the article pertaining to the first six Caesars were fashioned into a document consisting of six sections, one for each of the Caesars discussed. Each section began with a picture of a bust of the Caesar along with his name, followed by four to six paragraphs (approximately 400-700 words) of history and anecdotes related to the Caesar. The Caesars were presented in the order of their rule from earliest to latest; this ordering was evident from the consecutive nature of events described in the text. Pictures of associated figures referred to in the text

(e.g., Mark Antony) were interspersed throughout the document. The Caesar material contained 10 pages in all, printed back-to-back and stapled together into a packet.

Prime Minister materials were drawn from the website "The Prime Ministers of Canada" (www.primeministers.ca) and were arranged similarly to the Caesar information. Six pages were constructed, each containing the name, years in office, and picture of a 20th Century Canadian Prime Minister, followed by a brief sketch (150-200 words) of the Prime Minister's term. As with the Caesar material, the Prime Ministers were presented in chronological order.

The geology materials consisted of a single-page chart from the website "Enchanted Learning" (www.enchantedlearning.com). A portion of the chart containing information about the 12 most recent periods of geological history was used in the experiment. The periods were listed in order from latest (at the top of the chart) to earliest (at the bottom). The chart indicated the flora and fauna associated with each period, the years of history encompassed by the period, and the geological Era and Eon to which each period belonged. All information was conveyed in single words and short phrases; no prose was included.

Test Materials: Day 1. A separate paper-and-pencil test for each of the three study topics was constructed. Each test began with a question requiring serial reconstruction of the names or periods studied. In the Caesars test, for example, the six Caesars were listed vertically in a random order with blanks next to each; participants were to enter a number 1-6 into each blank to indicate the order in which the Caesars held power. The question was analogous in the Prime Minister and geology tests; on the geology test, six of the 12 studied periods were included, so that each of the three tests called for the serial reconstruction of six items.

Following the reconstruction question on each test were seven fill-in-the-blank (FIB) and six multiple choice (MC) questions written by the author. Each FIB question could be answered with a single Caesar, Prime Minister, or period name. MC questions contained four alternatives, each the name of a Caesar, Prime Minister, or period.

The Caesar and Prime Minister tests contained an additional section testing recognition of the pictures presented in the study materials. In this section, the six Caesar busts (or Prime Minister portraits) from the study materials were randomly intermixed with six new items from each category. New Caesar busts were gathered from various locations on the internet; new Prime Minister portraits were taken from the website mentioned above. Participants were to indicate whether or not they had seen each item during study by writing "yes" or "no" beneath the picture. As there were no pictures in the geology materials, no recognition section was included on the geology test.

Both the questions within each section and the response options within each question (where applicable) appeared on the test in a random order. The test was identical across participants.

Test Materials: Day 2. The three Day 1 tests were combined into a single test to be given on Day 2. The questions and correct answers were identical to those included on the Day 1 test, but with new random orderings of the serial reconstruction items, multiple choice options, and faces on the recognition test. The reconstruction questions appeared on the front page of the test, followed by the FIB, MC, and recognition sections. Within the FIB and MC sections, questions from each of the three topics were randomly intermixed.

Music Materials: The music was identical to that used in Experiment 1 (see Appendix A). The songs were played in a random order using Windows Media Player, and were heard through standard external computer speakers.

Procedure: Day 1. Participation took place in groups of one to three, with an experimenter present at all times. Participants were seated in positions roughly equidistant from the computer speakers, and such that they were not easily able to view each other's test answers. Instructions characterized the task as one with which the participants were quite familiar: studying for an exam. Materials for each of the three topics were introduced, and participants were told that they would be given 15 min to study all three before receiving a test on their contents. It was emphasized that 15 min would not be sufficient to learn all of the material thoroughly, and that the interest of the experiment was simply in how much they could learn in the time allotted. No particular direction was given as to how the materials were to be studied, but participants were asked to devote some time to each of the three topics. The music (either vocal or instrumental) was playing when participants entered the room and continued throughout the study phase, but was never mentioned by the experimenter.

After 15 min had elapsed, the music was stopped and the study materials were collected. The three tests were then passed out to each participant, with instructions that only 10 min would be given to work on all three. Participants were asked to answer as many questions as possible while striving for accuracy, and were advised against dwelling on any given question.

At the end of the 10-min test period, the study materials were handed back, the music was re-initiated, and a review phase began. Participants were instructed to use the study materials to check their test answers, and to mark any corrections to the right or left of the corresponding answer blank (but not to change their original answers). Seven min were given for the review. The entire session lasted approximately 40 min.

Procedure: Day 2. Participants returned roughly 24 hr after the first session and were again tested in groups of one to three. The second session consisted solely of the administration of the Day 2 test. The test was distributed following a brief description of its general composition. Participants were given up to 15 min to complete the test, and were encouraged to take any extra time after finishing to check their answers as thoroughly as possible. Background music was not played at any point in the session.

Results and Discussion

The data of one participant who did not follow test instructions and another who achieved an extremely low Day 1 test score were removed prior to analysis, leaving 22 and 18 participants in the lyrics and instrumental conditions, respectively. Caesar, Prime Minister, and geology test scores were computed by averaging the proportion of correct responses on the serial reconstruction, FIB, and MC portions of the tests. Overall test scores were then calculated as the mean of the Caesar, Prime Minister, and geology test scores. Results from the recognition section did not figure into the final test scores due to the incidental nature of that test (i.e., participants were most likely not trying to commit the pictures to memory while studying), as well as the fact that the geology test did not contain a recognition section. However, the recognition results are reported below.

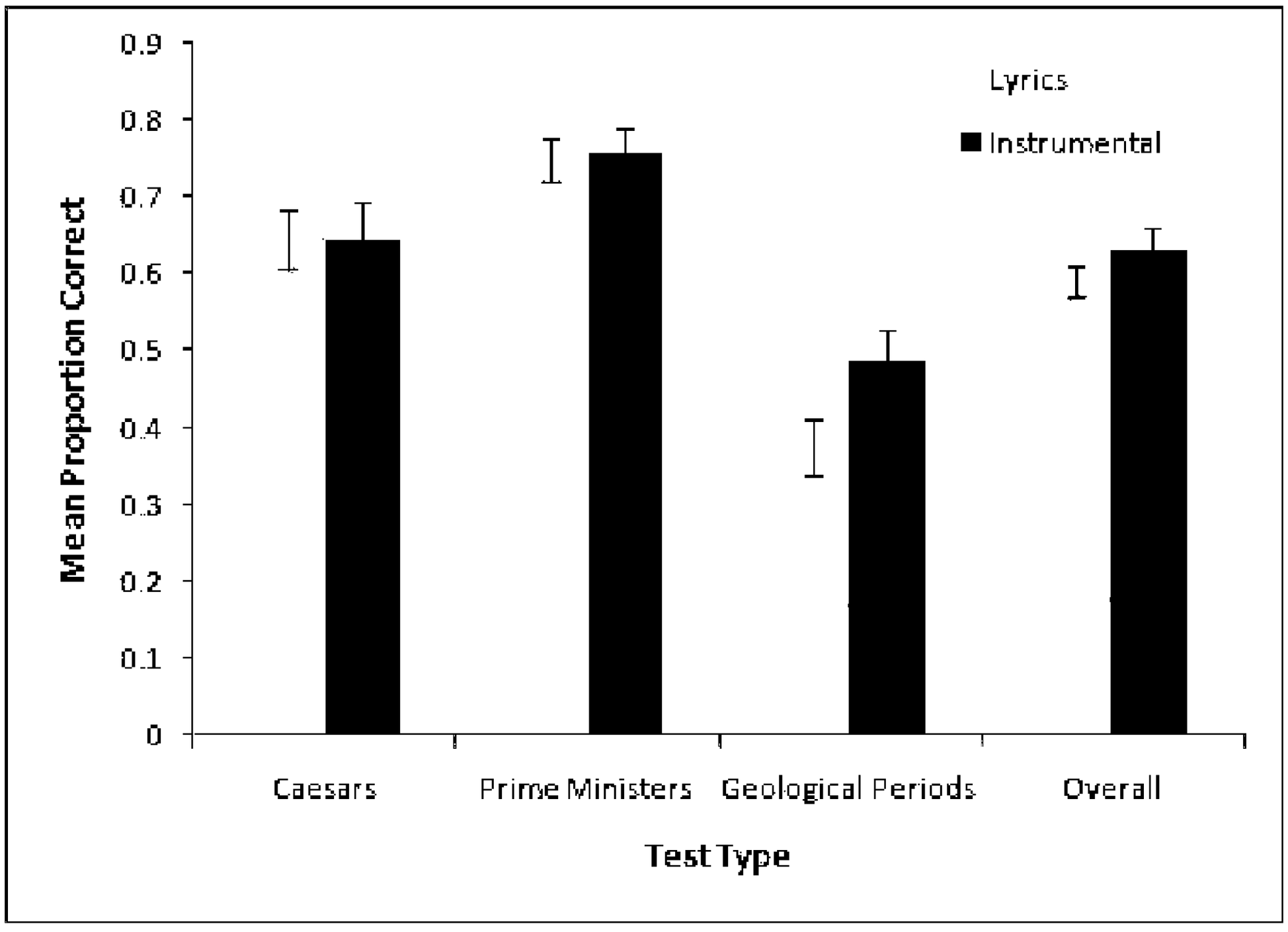

Figure 2. Test performance in Experiment 2 on Day 1 (top panel) and Day 2 (bottom panel). Error bars represent one standard error of the mean.

An additional repeated measures ANOVA was conducted to assess the performance of the instrumental and lyrics groups on the different types of test questions (serial reconstruction, FIB, and MC). This analysis revealed no interaction between listening condition and question type, $p = .75$. Planned comparisons yielded no significant differences between the listening conditions on any type of test question (all $ps > .25$), although there was a directional tendency towards better serial reconstruction performance in the instrumental condition. The only individual comparison yielding a significant difference between listening conditions concerned performance on MC geology questions, which was higher in the instrumental condition, $t(38) = 2.847$, $p < .01$. Although the instrumental group held moderate advantages in serial reconstruction of Prime Ministers (6.7%) and geological periods (10.5%), these differences did not reach significance, due in part to considerable variability in the data.

Finally, recognition test performance, measured as the hit rate minus the false alarm rate collapsed across the Caesar and Prime Minister tests, did not vary between the instrumental and lyrics groups, $p = .49$.

Day 2. The pattern of results on the second test closely followed that of the first test. Once again, the instrumental group held a slight overall performance advantage (68.3% to 64.3%), but this advantage failed to reach significance, $F(1,38) = 1.102$, $p = .30$. Listening condition did not interact with test type, $p = .14$, nor did it interact with question type, $p = .24$. Instrumental and lyrics participants performed equally on Caesar and Prime Minister tests (both $ps > .60$), but, as in the Day 1 test, the lyrics group performed more poorly on the geology test, $t(38) = 1.842$, $p < .08$. No other comparisons between the two groups approached significance, with the exception of superior performance for the instrumental group on FIB questions, $t(38) = 1.892$, $p < .07$.

These data were marked by a high degree of variability. The predominant source of this variability may have been strong individual differences in the ability to extract to-be-tested information from study materials in the brief time given. However, the naturalistic aims of the study required a lack of strict experimental control in several facets of the design, any of which could conceivably have contributed individual difference-based variability to the data. For example, participants likely varied in their study strategies (see General Discussion), pre-experimental familiarity with the topics studied, and personal reaction to the musical selections.

Despite this variability, the central result of Experiment 2 seems clear-cut: music with lyrics was detrimental to performance only on the geology test. Geology test scores were higher in the instrumental condition for all three question types on both administrations of the test, a consistency not seen in any other relevant aspect of the data. The magnitude of the vocal music detriment was virtually identical on Days 1 and 2 (11.3% and 11.0%, respectively), although the Day 2 result was only near-significant. Given that Caesar and Prime Minister test scores never differed by more than 2.5% across listening conditions, the effects observed on the geology test are compelling. Generally, they indicate that vocal music can cause an irrelevant speech effect while studying under relatively naturalistic circumstances. Moreover, they implicate the nature of the study materials as a critical determinant of the effects of the irrelevant speech: the geology materials were the only ones composed of a chart, and not paragraphs of prose. Explanations of these results are considered below.

GENERAL DISCUSSION

The object of the present experiments was to test a practical implication of the irrelevant speech effect; namely, that studying while listening to background music containing lyrics could impair memory for the studied material, thus lowering exam performance. To that end, both experiments were designed to more closely mimic an authentic study-test experience than traditional irrelevant speech experiments. In Experiment 1, lists were composed of names and terms following a meaningful, naturally occurring sequence. Furthermore, modern music was used and multiple reviews of the study lists were followed by a single test session, during which no music was played. Despite these departures from the established irrelevant speech paradigm, the overall pattern of results in Experiment 1 bore all the markings of an irrelevant speech effect: the presence of lyrics in music was associated with an approximately 8% decrease in accurate serial recall relative to instrumental music, and an approximately 10% decrease relative to silence. These trends parallel those of two other studies using music as irrelevant speech (Boyle & Coltheart, 1996; Salamé & Baddeley, 1989), even though both of those studies tested immediate serial recall after a single presentation of each list, and used lists of items with no systematic semantic relationship to one another.

One might have expected that repeated exposure to each list in Experiment 1 would have allowed participants in the lyrics condition to overcome the handicap posed by the vocal content of the music, yielding a weakened irrelevant speech effect. By contrast, the data suggest that the lyrics continued to interfere with serial encoding during each study presentation, resulting in a classic irrelevant speech effect. One might also have expected that the meaningful connections between the list items in Experiment 1 would have inspired a more semantic (and less phonological) processing orientation at study, rendering the lists less vulnerable to interference by irrelevant speech (e.g., Jones & Tremblay, 2000). The results suggest, however, that the strategy or combination of strategies chosen by participants did not diminish the detrimental impact of the lyrics. Considering that the naturalistic aspects of the Experiment 1 design failed to prevent the occurrence of an irrelevant speech effect, the present results may be taken as evidence that the effect can be produced by vocal background music, and can impact the efficacy of study efforts.

Experiment 2 was designed to test these implications under highly realistic conditions. While all study materials remained serially structured, they appeared in a prose format accompanied with illustrations, mimicking the style of a textbook, or as a chart. Tests were given immediately and approximately 24 hr after study, and were designed to emulate the format of a university exam. The results indicated that vocal music only reduced performance on the portion of the test for which the study material was a chart (geological periods); test scores on the topics studied with prose materials (Caesars and Prime Ministers) were virtually identical in the lyrics and instrumental conditions. Interestingly, this pattern of results held for multiple choice and fill-in-the-blank questions as well as questions calling for the reconstruction of order. The general pattern of results also held both immediately following study and after a 24-hr delay. In terms of the interfering effects of lyrics, then, the critical factor appeared to be the format of the study materials: information studied in prose format was resistant to the effects of irrelevant speech, while information studied in a chart format was not.

Why would irrelevant speech interfere with memory for information gleaned from a chart, but not from prose? An intuitive explanation is that the rich, elaborative nature of prose encourages semantic processing, while the list-like features of a chart elicit the phonological processing strategy known to be more vulnerable to effects of irrelevant speech (e.g., Hanley & Bakopoulou, 2003; Jones & Tremblay, 2000; Salamé & Baddeley, 1986). A second possibility is that the discrete, hierarchical structure of information presented in the chart encouraged serial processing, while the continuous nature of the prose rendered such processing less likely. This explanation may be particularly well-suited to account for the present data, as serial processing is a benchmark concomitant (if not prerequisite) of the irrelevant speech effect (e.g., Jones, 1993). Performance on the serial reconstruction portions of the Caesar and Prime Minister tests clearly indicates that order information was learned from prose, but such learning may not have arisen from explicitly serial study processing (e.g., serial rehearsal).

Whether either or both of the above explanations bear on the current results remains speculative until more is understood about the strategies engaged by participants at study. Indeed, the importance of strategy as a mediator of the irrelevant speech effect has received attention in the short-term serial recall irrelevant speech literature (e.g., Baddeley, 2000; Baddeley & Larsen, 2003; Hanley & Bakopoulou, 2003; Salamé & Baddeley, 1986), and might be of greater consequence under the more complex demands of exam studying. Participants appear to utilize a range of strategies even when faced with a simple serial recall task. Hanley and Bakopoulou (2003) explicitly manipulated study strategies by having participants engage in phonological or semantic processing of letter stimuli (discussed above), but included a third condition in which participants were free to choose their own strategies. These participants reported using five different strategies, classified as phonological (27% of participants), semantic (18%), a mixture of phonological and semantic (14%), representing each letter with a word and rehearsing the words subvocally (14%), and merging letters into short words or nonsense syllables and repeating these subvocally (27%). Each of these strategies likely entails a unique balance of phonological and semantic processing, and thus, potentially, a unique level of susceptibility to irrelevant speech effects. The results of the current experiments undoubtedly reflect a distribution of approaches as well. Assessing the strategic approach of participants under circumstances such as those of Experiment 2 is therefore an important direction for future research. To the extent that the nature of the study materials modulates processing strategy, it modulates susceptibility to the irrelevant speech effect.

Several other factors that might influence the extent of musical irrelevant speech effects on exam studying are not addressed by the present research. For example, a number of participants in Experiment 2 asked whether they would be allowed to take notes as they studied. While this option was not extended in the current experiments, it is a pervasive learning strategy, and should be incorporated into future experiments. Studying from notes, which often bear more resemblance to outlines and charts than to prose materials, might be more susceptible to irrelevant speech effects than other methods. The role of past experience studying with music is another factor left unexamined by the current experiments. Many participants in Experiment 1 claimed that they commonly listened to music while studying, and that they were proficient at "tuning it out" such that it did not affect their study efforts (though several others admitted to studying with music in spite of their sense that it was a hindrance to learning). Evidence that the irrelevant speech effect is resistant to habituation

(Hellbrück et al., 1996; Tremblay & Jones, 1998) would seem to argue against the claim that one can develop immunity to the effect, but this possibility has not been tested in the current context. A related question concerns the role of past experience with the music itself. While participants in the current experiments listened to songs that were largely unfamiliar to them, students often choose highly familiar (and highly enjoyed) songs as background music while studying. If such music is more easily ignored than novel music, the lower demand on attentional resources might reduce susceptibility to irrelevant speech effects (e.g., Neath, 2000).

The practice of studying with background music takes place in a complex real-world context. The role of the experiments reported in this chapter has been to suggest that the often laboratory-bound irrelevant speech effect can extend beyond its normal empirical range and hold influence in such a context. Future research will help determine when this influence is benign, and when it may cost a letter grade.

APPENDIX A

20th Century U.S. Presidents (in the order they held office)
Roosevelt
Truman
Eisenhower
Kennedy
Johnson
Nixon
Ford

Canadian Provinces (in the order they became part of Canada)
Nova Scotia
Manitoba
British Columbia
Prince Edward Island
Yukon
Alberta
Newfoundland

Metallic Elements (in order of appearance in the Periodic Table)
Chromium
Manganese
Iron
Cobalt
Nickel
Copper
Zinc

Shakespearean Plays (in order of date of first performance)
Henry VI
Richard III
Two Gentlemen of Verona
King John
As You Like It
The Merry Wives of Windsor
The Winter's Tale

The Caesars (the first seven)
Augustus
Tiberius
Caligula
Claudius
Nero
Vespasian
Titus

Layers of the Atmosphere (lowest to highest altitude)
Troposphere
Stratosphere
Mesosphere
Thermosphere
Ionosphere
Magnetosphere
Exosphere

Geological Eras (most recent to least recent)
Quarternary
Neogene
Paleogene
Cretaceous
Jurassic
Triassic
Permian

APPENDIX B

Songs used in the Lyrics condition
Alison Krauss – Baby, Now That I've Found You
Cowboy Junkies – Crescent Moon
Eels – I Like Birds
Eels – Packing Blankets
Last Forever – Hide and Seek

Sean Lennon – Bathtub
Sean Lennon – Breeze
Sean Lennon – Spaceship
Shawn Colvin – The Facts about Jimmy
The Sundays – Here's Where the Story Ends

Songs used in the Instrumental condition
311 – Colors
Alice In Chains – Whale & Wasp
Eels – Dusk: A Peach in the Orchard
Eels – Going to Your Funeral Part II
Orbital – New Style
Sting – Saint Agnes and the Burning Train
Stone Temple Pilots – No Memory
The Appleseed Cast – Doors Lead to Questions
The Appleseed Cast – View of a Burning City
U2 – Bass Trap

REFERENCES

Baddeley, A. (1986). *Working memory*. New York, NY: Clarendon Press/Oxford University Press.

Baddeley, A. D. (2000). The phonological loop and irrelevant speech effect: Some comments on Neath (2000). *Psychonomic Bulletin & Review, 7*(3), 544-549.

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A., & Larsen, J. D. (2003). The disruption of STM: A response to our commentators. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 56*(8), 1301-1306.

Balch, W. R., Bowman, K., & Mohler, L. A. (1992). Music-dependent memory in immediate and delayed word recall. *Memory & Cognition, 20*(1), 21-28.

Banbury, S. P., Tremblay, S., Macken, W. J., & Jones, D. M. (2001). Auditory distraction and short-term memory: Phenomena and practical implications. *Human Factors, 43*(1), 12-29.

Beaman, C. P. (2005). Auditory distraction from low-intensity noise: A review of the consequences for learning and workplace environments. *Applied Cognitive Psychology, 19*(8), 1041-1064.

Beaman, C. P., & Jones, D. M. (1997). Role of serial order in the irrelevant speech effect: Tests of the changing-state hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(2), 459-471.

Boyle, R., & Coltheart, V. (1996). Effects of irrelevant sounds on phonological coding in reading comprehension and short-term memory. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 49*(2), 398-416.

Buchner, A., & Erdfelder, E. (2005). Word frequency of irrelevant speech distractors affects serial recall. *Memory & Cognition, 33*(1), 86-97.

Buchner, A., Irmen, L., & Erdfelder, E. (1996). On the irrelevance of semantic information for the 'irrelevant speech' effect. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 49*(3), 765-779.

Buchner, A., Rothermund, K., Wentura, D., & Mehl, B. (2004). Valence of distractor words increases the effects of irrelevant speech on serial recall. *Memory & Cognition, 32*(5), 722-731.

Colle, H. A. (1980). Auditory encoding in visual short-term recall: Effects of noise intensity and spatial location. *Journal of Verbal Learning & Verbal Behavior, 19*(6), 722-735.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning & Verbal Behavior, 15*(1), 17-31.

Ellermeier, W., & Zimmer, K. (1997). Individual differences in susceptibility to the 'irrelevant speech effect.'. *Journal of the Acoustical Society of America, 102*(4), 2191-2199.

Farley, L. A., Neath, I., Allbritton, D. W., & Surprenant, A. M. (2007). Irrelevant speech effects and sequence learning. *Memory & Cognition, 35*(1), 156-165.

Fogelson, S. (1973). Music as a distractor on reading-test performance of eighth grade students. *Perceptual and Motor Skills, 36*(3), 1265-1266.

Freeburne, C. M., & Fleischer, M. S. (1952). The effect of music distraction upon reading rate and comprehension. *Journal of Educational Psychology, 43*(2), 101-109.

Furnham, A., & Bradley, A. (1997). Music while you work: The differential distraction of background music on the cognitive test performance of introverts and extroverts. *Applied Cognitive Psychology, 11*(5), 445-455.

Hallam, S., Price, J., & Katsarou, G. (2002). The effects of background music on primary school pupils' task performance. *Educational Studies, 28*(2), 111-122.

Hanley, J. R., & Bakopoulou, E. (2003). Irrelevant speech, articulatory suppression, and phonological similarity: A test of the phonological loop model and the feature model. *Psychonomic Bulletin & Review, 10*(2), 435-444.

Hanley, J. R., & Broadbent, C. (1987). The effect of unattended speech on serial recall following auditory presentation. *British Journal of Psychology, 78*(3), 287-297.

Hellbrück, J., Kuwano, S., & Namba, S. (1996). Irrelevant speech and human performance: Is there long-term habituation? *Journal of the Acoustical Society of Japan, 17,* 239-247.

Jones, D. M. (1993). Objects, streams, and threads of auditory attention. In A. D. Baddeley, & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control: A tribute to Donald Broadbent.* (pp. 87-104). New York, NY, US: Clarendon Press/Oxford University Press.

Jones, D. M. (1999). The cognitive psychology of auditory distraction: The 1997 BPS Broadbent lecture. *British Journal of Psychology, 90*(2), 167-187.

Jones, D. M. (1994). Disruption of memory for lip-read lists by irrelevant speech: Further support for the changing state hypothesis. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 47*(1), 143-160.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(2), 369-381.

Jones, D. M., & Macken, W. J. (1995). Phonological similarity in the irrelevant speech effect: Within- or between-stream similarity? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 103-133.

Jones, D. M., Macken, W. J., & Murray, A. C. (1993). Disruption of visual short-term memory by changing-state auditory stimuli: The role of segmentation. *Memory & Cognition, 21*(3), 318-328.

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(3), 656-674.

Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: The role of changing state. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 44*(4), 645-669.

Jones, D. M., & Tremblay, S. (2000). Interference in memory by process or content? A reply to Neath (2000). *Psychonomic Bulletin & Review, 7*(3), 550-558.

Knez, I., & Hygge, S. (2002). Irrelevant speech and indoor lighting: Effects of cognitive performance and self-reported affect. *Applied Cognitive Psychology, 16*(6), 709-718.

LeCompte, D. C. (1994). Extending the irrelevant speech effect beyond serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1396-1408.

Macken, W. J., & Jones, D. M. (2003). Reification of phonological storage. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 56*(8), 1279-1288.

Martin, R. C., Wogalter, M. S., & Forlano, J. G. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory and Language, 27*(4), 382-398.

Miles, C., Jones, D. M., & Madden, C. A. (1991). Locus of the irrelevant speech effect in short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(3), 578-584.

Murdock, B. B. (1960). The distinctiveness of stimuli. *Psychological Review, 67*(1), 16-31.

Naime, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory & Cognition, 16*(4), 343-352.

Naime, J. S. (1990). A feature model of immediate memory. *Memory & Cognition, 18*(3), 251-269.

Neath, I. (2000). Modeling the effects of irrelevant speech on memory. *Psychonomic Bulletin & Review, 7*(3), 403-423.

Neath, I., Farley, L. A., & Surprenant, A. M. (2003). Directly assessing the relationship between irrelevant speech and articulatory suppression. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 56A*(8), 1269-1278.

Neath, I., & Suprenant, A. M. (2003). *Human memory: Second Edition*. Belmont, CA: Wadsworth/Thomson Learning.

Neath, I., Surprenant, A. M., & LeCompte, D. C. (1998). Irrelevant speech eliminates the word length effect. *Memory & Cognition, 26*(2), 343-354.

Neely, C. B., & LeCompte, D. C. (1999). The importance of semantic similarity to the irrelevant speech effect. *Memory & Cognition, 27*(1), 37-44.

Ransdell, S. E., & Gilroy, L. (2001). The effects of background music on word processed writing. *Computers in Human Behavior, 17*(2), 141-148.

Salamé, P., & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning & Verbal Behavior, 21*(2), 150-164.

Salamé, P., & Baddeley, A. D. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society, 24*(4), 263-265.

Salamé, P., & Baddeley, A. D. (1987). Noise, unattended speech and short-term memory. *Ergonomics, 30*(8), 1185-1194.

Salamé, P., & Baddeley, A. D. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 41*(1), 107-122.

Surprenant, A. M., LeCompte, D. C., & Neath, I. (2000). Manipulations of irrelevant information: Suffix effects with articulatory suppression and irrelevant speech. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 53*(2), 325-348.

Tremblay, S., & Jones, D. M. (1998). Role of habituation in the irrelevant sound effect: Evidence from the effects of token set size and rate of transition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(3), 659-671.