Eye movements and reading comprehension while listening to preferred and non-preferred study music
Roger Johansson, Kenneth Holmqvist, Frans Mossberg and Magnus Lindgren
Psychology of Music published online 10 March 2011
DOI: 10.1177/0305735610387777

The online version of this article can be found at:
http://pom.sagepub.com/content/early/2011/03/06/0305735610387777

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
sempre:
Society for Education, Music and Psychology Research

Additional services and information for Psychology of Music can be found at:

Email Alerts: http://pom.sagepub.com/cgi/alerts

Subscriptions: http://pom.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav
Eye movements and reading comprehension while listening to preferred and non-preferred study music

Roger Johansson
Department of Cognitive Science, Lund University, Sweden

Kenneth Holmqvist
Centre for Languages and Literature, Lund University, Sweden

Frans Mossberg
Sound Environment Centre, Lund University, Sweden

Magnus Lindgren
Department of Psychology, Lund University, Sweden

Abstract
In the present study 24 university students read four different texts in four conditions: (1) while listening to music they preferred to listen to while studying; (2) while listening to music they did not prefer to listen to while studying; (3) while listening to a recording of noise from a café; and finally (4) in silence. After each text they took a reading-comprehension test. Eye movement data were recorded for all participants in all conditions.

A main effect for the reading-comprehension scores revealed that the participants scored significantly lower after they had been listening to the non-preferred music while reading, compared with reading in silence. No significant effects were found between the other conditions. No significant differences between conditions were found for the traditional eye movement measures in reading (fixation duration, saccadic amplitude, regressions, and first-pass and second-pass reading time). It is suggested that this result is a consequence of participants not being aware that their reading processes are disrupted by a non-preferred musical background. They do not make the necessary changes to the processes involved in reading required to compensate for increased cognitive load. The results are discussed in relation to study/reading habits, extraversion, arousal and working memory capacity.

Keywords
eye movements, music and cognition, music distraction, reading comprehension

Corresponding author:
Roger Johansson, Centre for Language and Literature, Lund University, Box 201, 221 00 Lund, Sweden.
[email: Roger.Johansson@ling.lu.se]
Introduction

The ability to understand information contained in a text is, in many ways, a prerequisite for success, for example in educational environments. Students read large amounts of literature to acquire knowledge and to be able to pass courses and examinations. The reading process involved is often a very cognitively-demanding task and there are a multitude of variables that may affect how well the written information is assimilated. Students often read and study in environments where certain distractions, voluntary or not, compete for their attention. Typical situations include studying while listening to music, watching television or listening to people talking in a café. As early as 1935, Cantril and Allport reported that 68% of the students they investigated listened to the radio while studying. Today, portable players have made it possible to listen to music anywhere, anytime. For example, Furnham, Trew and Sneade (1999) found that 90% of the 142 participants in their study listened to music while studying, and Ransdell and Gilroy (2001) reported that the majority of the college students in their study listened to music when using their computer for schoolwork. As part of the preparations for the present study, 150 students answered a questionnaire about their study habits; 81% of them said that they sometimes listened to music while studying. People often have strong beliefs about how they are affected in these situations. A very common claim is that they read and study better with certain music. In the preparatory survey, the majority (77%) of those who often listened to music while studying also believed that they studied better with music.

The interplay between reading and listening to music has been addressed in several studies and from different angles. In many respects, however, these studies have yielded different and conflicting findings, indicating that reading to music is a very complex and difficult phenomenon to study. The effects of music include that it arouses deep emotions (e.g., Juslin & Sloboda, 2001), influences mood (e.g., Eich & Forgas, 2003) and affects cognitive performance (e.g., Furnham & Allass, 1999). What is more, account must be taken of the task to be performed and the context in which the music is encountered as well as many characteristics intrinsic to the music: its loudness, tempo (bpm), form, complexity, variety, familiarity, genre, whether it is vocal or instrumental, etc. Previous studies have typically focused on one or more of these aspects and compared them with a silent condition. Kiger (1989) studied reading comprehension in the conditions of “low-information-load” music, “high-information-load” music and silence. Information load was calculated as a function of the amount of loudness, variety, complexity and tonal range in the music. Kiger found that listening to low-information-load music improved reading comprehension compared with both silence and listening to high-information-load music. Furnham and Allass (1999), however, did not find any difference in how participants performed in a reading-comprehension task when listening to music of different complexity with regard to tempo, repetition and instrumental layering. Fogelson (1973) reported that popular instrumental music impaired reading comprehension in 8th graders. Iwanaga and Ito (2002) found that both vocal and instrumental music impaired performance in a verbal-memory task.

Several studies have also considered background noise instead of music (e.g., Hygge, 2003; Hygge, Evans, & Bullinger, 2002) and found various deleterious effects on cognitive-task performance. Clark and colleagues (2006) studied how exposure to noise from aircraft and road traffic affected reading comprehension in 9–10-year-olds at 89 schools in the Netherlands, Spain and England. They found that aircraft noise impaired reading comprehension but that road traffic noise did not. Furnham and Strbac (2002) found no significant difference between background music and office noise as regards their impact on reading
comprehension, proposing that this might be because the music and the noise used in their study were very similar in complexity.

The impact of music on people may also vary greatly between individuals, especially its potential to change arousal levels. Eysenck (1967) argues in his theory of personality that extraverts and introverts differ in how much external stimulation they require to reach an optimal level of arousal. Introverts experience greater arousal in response to lower-intensity stimulation than extraverts, and they experience an inhibition of excitation when their optimal level of arousal is exceeded (which thus happens at an earlier stage than for extraverts). Therefore, introverts typically exhibit an active aversion to situations of external stimulation (such as music) whereas extraverts actively seek out these situations. Background music would thus be expected to have a more negative effect on task performance in introverts than in extraverts. Several studies have indeed confirmed that introverts are more negatively affected by music and noise during cognitively demanding tasks (e.g., Cassidy & MacDonald, 2008; Daoussis & McKelvie, 1986; Furnham & Allass, 1999; Furnham & Bradley, 1997).

Besides personality type, another aspect reported as worthy of consideration is personal preferences. Etaugh and Ptasnik (1982) found that individuals who usually studied without background music scored higher on a reading-comprehension test while studying in silence than in the presence of music, whereas the case was the opposite for those who usually studied with background music. However, it is likely that those of their participants who usually did not listen to music while studying had a more introverted personality, and that those who usually did listen to music while studying had a more extraverted personality. For example, Daoussis and McKelvie (1986) reported that extraverts studied in the presence of music twice as often as introverts. Chamorro-Premuzic with colleagues have also in a number of cross-cultural studies (British, American, Spanish and Malaysian students) investigated how personality types correlate with different uses of music (Chamorro-Premuzic & Furnham, 2007; Chamorro-Premuzic, Gomà-i-Freixanet, Furnham, & Muro, 2009; Chamorro-Premuzic, Swami, Furnham, & Maakip, 2009). Results indicate that those who use music as a background to other activities (e.g., studying and working) are more likely to be extraverted.

Another factor that may influence the impact of listening to music while being engaged in another task is working-memory capacity. Working memory has limited resources (Baddeley & Hitch, 1974), meaning that different processes may compete for these resources. It has been proposed that reading comprehension depends on maintaining phonological representations of the written words in working memory while their meaning is derived (e.g., Baddeley, 1979, 2003; Levy, 1977). Background music or noise could occupy some of the available working-memory resources so that the person can no longer fully attend to the reading process. Individual differences in working-memory capacity might, therefore, be important for reading-comprehension performance in the presence of background noise.

**Aim of the study**

The aim of the present study is twofold. First, it investigates how reading comprehension is affected by different kinds of music and a certain type of background noise, and it examines whether the commonly-held belief that certain music improves reading comprehension is actually true. Second, while most studies have focused only on determining whether certain music or noise affects reading comprehension, this one also investigates how the reading process might be affected by music and noise.
To be able to investigate if the belief that certain music improves reading comprehension is correct, we have chosen not to focus on a specific kind of music or certain aspects of it, but on personal preferences. In this study, reading will therefore be performed while participants listen to music they prefer to listen to while studying/reading and while they listen to music they absolutely do not prefer to listen to while studying/reading. Given that students also very commonly study in cafés or similar environments, reading performance will also be studied in the presence of recorded noise from a café. Finally, reading will be studied in complete silence. Personal preferences as regards the auditory environment for reading have been studied before (Daoussis & McKelvie, 1986), but there the choice of preferred music was limited and related more to general music preferences. The music was not explicitly chosen as music the participants preferred to study/read to. In this study the participants were completely free to choose the music they preferred to study/read to and were also able to specify what kind of music they did not prefer to study/read to. To our knowledge, no previous studies have investigated personal preferences as regards music during reading to this extent.

To be able to study the reading process as it unfolds we used eye-tracking technology, which is well established in reading research and makes it possible to study reading in terms of both reading patterns and information processing (cf. Rayner, 1998). Examples of previous uses to which eye tracking has been put include the study of reading in terms of poor readers (e.g., Olson, Kliegel, & Davidson, 1983), reading strategies (e.g., Hyönnä, Lorch, & Kaakinen, 2002), speed-reading (e.g., Just & Carpenter, 1987), different reading materials (e.g., Rayner & Pollatsek, 1989), repeated readings of the same text (e.g., Hyönnä, 1995; Schnitzer & Kowler, 2006) and reading during writing (Johansson, Wengelin, Johansson, & Holmqvist, 2010). To our knowledge, this is the first study to use eye tracking to examine the ongoing process of reading against a background of music and noise.

Individual differences are not a main focus of this study. Nevertheless, considering previous research on arousal (e.g., Juslin & Sloboda, 2001), introversion/extraversion (e.g., Cassidy & MacDonald, 2008; Daoussis & McKelvie, 1986; Furnham & Allass, 1999; Furnham & Bradley, 1997), working memory capacity (e.g., Baddeley, 1979, 2003; Levy, 1977) and music and reading habits (e.g., Etaugh & Ptasnik, 1982) these aspects were considered in post hoc analyses. Pupil size was used as a measure of arousal (e.g., Partala & Surakka, 2003), the Eysenck Personality Questionnaire – Revised (EPQ-R) personality test (Eysenck, Eysenck, & Barrett, 1985) was used to test participants for extraversion/introversion and the operation-span test (OSPAN) (Turner & Engle, 1989) was used to test them for working memory capacity. Finally, a questionnaire about music and study habits was developed to test participants for these factors.

**Eye tracking in reading studies**

Fixation duration, saccadic amplitude and regressions are the three major general eye-movement measures of relevance to reading (Rayner, 1998). **Fixation duration** is the time during which the eye is fixated on a word; a large body of research indicates that this measure can be used as an indicator of cognitive processing: that is, longer fixation duration indicates a more cognitively demanding task (e.g., Reichle, Pollatsek, Fisher, & Rayner, 1998). An average fixation in reading lasts about 200–250 ms (Rayner, 1998). **Saccadic amplitude** is the length of the rapid movements between fixations in which virtually no information at all is extracted. Average saccadic amplitudes in reading are typically 7–9 letter spaces. Readers with reading disabilities and poor readers typically have shorter saccadic amplitudes (e.g., Olson et al., 1983). About 10–15% of all saccades are **regressions**: they move backwards on the same line, move to previous lines or move within the word currently being fixated on. It is believed that short
within-word regressions occur when the reader is having difficulty in lexically activating a word, whereas longer regressions occur when the reader does not understand the text (Rayner, 1998). A reading task with high cognitive demands would typically result in longer fixation durations and more regressions.

It is also common to analyse eye movements in reading in terms of first-pass and second-pass reading time (cf. Rayner, 1998). First-pass reading time is the sum of the fixations during the initial reading of the text, and second-pass reading is the sum of the fixations during re-reading of the text, or parts of it. Reading texts with the purpose to remember their content would typically result in a relatively large amount of second-pass reading time.

Apart from these three major eye-movement measures in reading, the size of the pupil is also of interest. Pupil dilation has been used as an indication of arousal, emotion, stress and cognitive workload (e.g., Hess & Polt, 1960; Partala & Surakka, 2003; Pomplun & Sunkara, 2003). Larger pupil size indicates more arousal and/or higher cognitive demands. Partala and Surakka (2003) found that participants listening to both positive (baby laughing) and negative (baby crying) sounds had larger pupils than when listening to neutral sounds (office noise). However, no significant difference was found between positive and negative sounds. Ratings of arousal revealed that negative and positive sounds were experienced as equally arousing. Partala and Surakka (2003) therefore argue that the magnitude of the pupil response in the presence of sound is determined by the amount of emotional arousal.

**Hypotheses and expectations**

Since previous research has shown very diverse results when it comes to reading comprehension in the presence of music and noise it is hard to predict how preferred and non-preferred music as well as café noise will affect reading comprehension. Nevertheless, we propose the following hypotheses:

(1) Reading comprehension is improved by the preferred music;
(2) Reading comprehension is impaired by the non-preferred music and the café noise.

We expect that any noise conditions that turn out to impair reading comprehension will make the reading process more cognitively demanding. If any of the noise conditions turns out to impair the reading process, we therefore expect to find the following tendencies in the eye-movement data:

(a) Fixation durations will be longer;
(b) There will be more regressions;
(c) There will be more second-pass reading.

Since poor readers typically have shorter saccadic amplitudes in reading (e.g., Olson et al., 1983), a tendency in that direction would also be likely if the reading process is impaired.

**Method**

**Participants**

Twenty-four university students, 12 females and 12 males, participated in the experiment. All participants reported normal or corrected-to-normal vision (with contact lenses or glasses).
All participants were native speakers of Swedish. The mean age of the participants was 27.9 years ($SD = 7.7$).

**Data**

Eye-movement data were recorded for the participants while they were reading four different texts in four different conditions:

1. Reading while listening to music the participant preferred to listen to while studying/reading;
2. Reading while listening to music the participant did not prefer to listen to while studying/reading;
3. Reading while listening to recorded noise from a café;
4. Reading in silence.

The four texts were randomized and balanced for order so that all of them were read the same number of times in the four conditions. After each text the participants took a reading-comprehension test consisting of four multiple-choice questions, each with five options. The reading-comprehension test was always performed in silence. Additionally, each participant completed a personality test (EPQ-R), was tested for working-memory capacity (Operation span [OSPA]), and completed a questionnaire about his or her music and study habits.

**Stimuli, apparatus and materials**

We first selected a number of texts that had been used in the Swedish Scholastic Assessment Test (SweSAT). This is a standardized test used to gain admission to higher education in Sweden, and includes a reading-comprehension module with texts designed to be as comparable as possible in regards of difficulty. The subject areas of the texts chosen were administration, economics and sociology. We computed several measures of readability (text length, LIX, vocabulary diversity (vocd) and lexical density) for 20 different candidate texts and identified the four most comparable ones; that is, no significant differences were found for any of the readability measures.

The texts were shown to the participants on a 19" Samsung GH19PS monitor with the resolution set to $1024 \times 768$ pixels and with a physical size of $380 \times 300$ mm. The texts were presented in a large Ariel font (21 points) and with 1.5 spacing. Each text consisted of 10–11 pages (screens) with about 1,000 words (min: 918, max: 1091) and was followed by four multiple-choice questions, each with five options (only one option was correct). The participants were seated 0.67 metres from the monitor.

The music or noise was played from speakers placed in front of the participant next to the monitor. All noise and music played during the experiment had been normalized to the same sound level in the software Adobe Audition 3.0. On average the sound level for the noise and music varied between 60–70 dB and never exceeded 80 dB (measured in the position where the participants were seated). The café noise was recorded in a real café and consisted of a relatively homogeneous buzz, but with frequent and audible conversations between people.

Eye tracking was performed during the reading of all texts with a SMI iView X Hi-Speed 1250 Hz system. Presentation of stimuli and communication with the eye tracker were performed with Matlab and the Psychophysics Toolbox Version 2 (PTB-2; Brainard, 1997).
The OSPAN test (Turner & Engle, 1989) was performed without eye tracking at the computer. In this test participants were required to judge as quickly as possible if mathematics operations were false or correct while trying to remember a set of unrelated words. First a mathematics operation was presented on the screen and, when the participant had judged it as true or false, a single syllable word was shown at the centre of the screen for exactly one second. The set of mathematics operations and words were presented in random order for two-item sets up to five-item sets in three trials. For example, a three-item set might have been:

\[
\begin{align*}
(1/1) - 1 &= 0? & \text{JOB} \\
(5 \times 4) + 3 &= 22? & \text{BOAT} \\
(3 \times 1) - 3 &= 1? & \text{CAR}
\end{align*}
\]

After each item-set the participants were required to type the words in the presented order. The total score from each correct word (maximum: 42) was used as the working memory capacity measure.

The EPQ-R personality test (Eysenck et al., 1985) was performed by pencil and paper. This test measures personality in the psychometric scales of psychoticism, extraversion, neuroticism and lie. The test consists of 100 yes or no questions. Twenty-three of these questions belong to the extraversion scale. For example, “Do you enjoy meeting new people?”. The score (maximum: 23) was used as the extraversion measurement.

The questionnaire about music and reading habits was performed by pencil and paper. For example, they were to judge if they preferred to study in silence (yes/no/doesn’t matter), how often they listened to music while studying (often/sometimes/rarely/never), and if they felt they studied better with music (yes/no).

**Procedure**

The participants were given a cinema ticket in exchange for participating in the experiment. All participants had been told to bring 20–30 minutes of music that they preferred to listen to while studying/reading. Their choices ranged from instrumental classical music to different kinds of modern pop, rock or rap music, with or without vocals (all choices are listed in Appendix A). The participants had also told the experiment leader what kind of music they absolutely did not prefer to listen to while studying/reading. The experiment leader selected typical music in those categories and brought it to the experiment. Again the choices varied a great deal (all choices are listed in Appendix A), but common non-favourites were heavy metal, gangster rap, techno and schlager.²

All participants began the experiment by performing the working-memory test (OSPA). While the participants were doing that, the experiment leader had time to normalize the volume of the music brought by the participants to ensure that it would be played at the same volume as the non-preferred music and the café noise. The participants were then introduced to the eye-tracking equipment and were seated as comfortably as possible. The experiment leader informed them that they were about to read four different texts and that, immediately after reading each text, they would answer four multiple-choice questions about that text. They were also informed that they would read one of the texts to their own preferred music, one to the non-preferred music, one to recorded noise from a café and one in silence, and that they
Psychology of Music

would always answer the questions in silence. They did not, however, know the order of the four conditions. They were also informed that they could only move forward in the texts (by pressing the mouse), that they had a maximum of 20 minutes to finish each text with the corresponding questions, and that if the time was about to run out (which it never actually did in the experiments) the experiment leader would inform them when five minutes remained.

The eye tracker was calibrated before the first text and re-calibrated if necessary before each following text. The experiment leader was never present in the room while the participant performed the reading or answered the questions, but always entered the room between the texts to check that everything was in order and to see if a re-calibration was needed. The eye-tracking system requires participants to hold their head still during reading, but the participants were free to move their head and stretch their neck while answering the questions about the texts (when no eye-movement data were being recorded) and in between texts. This meant that the system almost always needed re-calibration before the next text and condition were introduced. When all four texts (with the corresponding questions) had been completed, the participants rated how disruptive they had found the non-preferred music: 1 = Not disruptive at all; 2 = Slightly disruptive; 3 = Disruptive; 4 = Very disruptive; 5 = Neutral/Don’t know. Finally, all participants answered the questionnaire about their music and study habits (Appendix A) and completed the personality questionnaire (EPQ-R).

Analysis

Reading-comprehension scores (average proportion of correct responses) and eye-movement data were analysed using a within-subjects analysis of variance (ANOVA) with condition (preferred music/non-preferred music/café noise/silence) as the within-subjects factor. Dependent variables in the eye-movement data were the three major measures of eye movements in reading – fixation duration, saccadic amplitude and regressions, as well as the proportion of second-pass reading. Fixation durations and saccadic amplitudes were also broken down depending on whether they belonged to the first-pass or second-pass reading time. Regressions were separated into regressions on the same line, regressions to previous lines and regressions within words. Additionally, on the basis of the results by Partala and Surakka (2003), pupil size was used as a measure of arousal and was analyzed as a dependent variable within the conditions. If the assumption of homogeneity of variance was violated, Friedman and Wilcoxon analyses were performed.

To test the results for underlying individual differences, each analyzed factor in each condition was correlated (Bivariate Spearman’s Rho) with the extraversion dimension in the EPQ-R scores, the OSPAN scores and with the answers from the questionnaire about music and study habits. Furthermore, the rating as to how disruptive the participants had found their non-preferred music was correlated with each analyzed factor in that condition.

Results

Reading comprehension

For the reading-comprehension score, a main effect of condition was found ($F(3,66) = 3.191$, $p = 0.03$). Mean scores and standard deviations in all four conditions are shown in Table 1. The average score for Swedish students on this part of the SweSAT is about 55–60 percent (Wallin & Eriksson, 2002). Bonferroni post-hoc tests revealed that the participants scored significantly
lower \( (p < 0.01) \) in the non-preferred music than in silence. No significant difference was found between preferred music and silence or between café noise and silence, and no significant differences were found among the three noise conditions.

**Reading process: eye-movement data**

Table 2 (a–d) shows means and standard deviations for eye-movement data in all four conditions. No significant main effects for fixation durations or saccadic amplitudes were found for either first-pass or second-pass reading. Neither did proportion of second-pass reading time, regressions on the same line, regressions to a different line or regressions within words result in any main effects.

**Arousal: pupil size**

For the arousal measure of pupil size (the horizontal radius of the pupil), there was a main effect of condition \( (F(3,66) = 3.122, p = 0.03) \). Mean scores and standard deviations in all four

### Table 1. Average scores on the reading-comprehension test; standard deviations in brackets

<table>
<thead>
<tr>
<th>Condition</th>
<th>Preferred music</th>
<th>Non-preferred music</th>
<th>Café noise</th>
<th>Silence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scores (%)</td>
<td>45.8 (35.9)</td>
<td>40.2 (21.0)</td>
<td>50.0 (29.5)</td>
<td>62.5 (24.5)</td>
</tr>
</tbody>
</table>

### Table 2. Average values for eye-tracking data; standard deviations in brackets

#### a. First-pass reading: fixation duration and saccadic amplitude

<table>
<thead>
<tr>
<th>Condition</th>
<th>Preferred music</th>
<th>Non-preferred music</th>
<th>Café noise</th>
<th>Silence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixation duration (ms)</td>
<td>204 (19)</td>
<td>202 (18)</td>
<td>202 (21)</td>
<td>199 (18)</td>
</tr>
<tr>
<td>Saccadic amplitude (letter spaces)</td>
<td>6.9 (3.8)</td>
<td>6.9 (2.1)</td>
<td>7.2 (4.6)</td>
<td>6.4 (1.4)</td>
</tr>
</tbody>
</table>

#### b. Second-pass reading: proportion thereof, fixation duration and saccadic amplitude

<table>
<thead>
<tr>
<th>Condition</th>
<th>Preferred music</th>
<th>Non-preferred music</th>
<th>Café noise</th>
<th>Silence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of second-pass reading (%)</td>
<td>42.9 (10.6)</td>
<td>42.4 (10.1)</td>
<td>41.9 (10.0)</td>
<td>39.9 (9.1)</td>
</tr>
<tr>
<td>Fixation duration (ms)</td>
<td>200 (17)</td>
<td>200 (17)</td>
<td>199 (22)</td>
<td>198 (17)</td>
</tr>
<tr>
<td>Saccadic amplitude (letter spaces)</td>
<td>5.1 (1.6)</td>
<td>5.5 (2.2)</td>
<td>5.1 (1.1)</td>
<td>5.0 (1.2)</td>
</tr>
</tbody>
</table>

#### c. Regressions as mean percentages of all saccades

<table>
<thead>
<tr>
<th>Condition</th>
<th>Preferred music</th>
<th>Non-preferred music</th>
<th>Café noise</th>
<th>Silence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion to same line (%)</td>
<td>15.1 (4.2)</td>
<td>14.8 (4.5)</td>
<td>14.6 (4.7)</td>
<td>15.6 (4.3)</td>
</tr>
<tr>
<td>Proportion to different line (%)</td>
<td>2.5 (1.3)</td>
<td>2.4 (1.0)</td>
<td>2.5 (1.1)</td>
<td>2.6 (1.4)</td>
</tr>
<tr>
<td>Proportion within word (%)</td>
<td>9.8 (3.5)</td>
<td>10.4 (3.2)</td>
<td>10.1 (2.7)</td>
<td>10.1 (4.1)</td>
</tr>
</tbody>
</table>
conditions are shown in Table 3. Bonferroni post-hoc tests revealed that, compared with the silent condition, the pupil was significantly larger for preferred music ($p = 0.004$) and non-preferred music ($p = 0.03$) but not for the café noise. No significant differences were found between the three noise conditions.

**Individual differences**

From the scores in the EPQ-R test it was revealed that there was a negative correlation for the extraversion dimension and the reading comprehension scores for preferred music ($\rho = -0.463$) and silence ($\rho = -0.406$). There was also a positive correlation for the extraversion dimension and within word regressions ($\rho = 0.481$) in the preferred music condition, and a negative correlation for the extraversion dimension and the saccades in all conditions (preferred: $\rho = -0.433$; non-preferred: $\rho = -0.441$; café: $\rho = -0.484$; silence: $\rho = -0.446$).

No significant correlations were found between either the scores on the reading-comprehension test or any of the variables in the eye-movement data and the OSPAN scores and the answers to the questionnaire. Nor were there any correlations between the degree of disruptiveness ascribed by the participants to the non-preferred music and either their scores on the reading-comprehension test or their eye-movement data for that condition.

Additionally, since 12 of the participants had chosen instrumental music and 12 had chosen music with vocals as their preferred music, the relationship between this factor and the reading-comprehension scores and the eye-movement data was also investigated. However, no significant correlations were found for this factor either.

**Discussion**

The objective of this study was twofold. First, it investigated how reading comprehension was affected by different kinds of music and background noise. The investigated conditions were music the participants preferred to study/read to, music they did not prefer to study/read to, café noise and silence. Second, eye-tracking was used as a method to investigate how the reading process was affected by these conditions. Individual differences were not a main focus of this study. But based on previous findings, the reading-comprehension scores and the eye-tracking measures were correlated post hoc with the extraversion dimension from the EPQ-R personality test (Eysenck et al., 1985), with the working memory capacity score from the operation-span test (OSPAN) (Turner & Engle, 1989) and with the participants’ music and study habits (as reported in the questionnaire). Furthermore, pupil size was used as a measure of arousal (Partala & Surakka, 2003).

The most notable finding from the reading comprehension scores revealed that, compared to reading in silence, participants performed significantly worse when reading to the non-preferred music. For the preferred music and the café noise the scores were not significantly
different compared to silence. We therefore confirm the hypothesis that non-preferred music impairs reading comprehension and reject the hypothesis that a person’s reading comprehension is improved by listening to music that he or she prefers to study/read to. It should, however, be noted that the standard deviations for the reading-comprehension scores are large, especially for the preferred music condition. As stated before, there are a great many characteristics intrinsic to music that may have different implications for how people are affected by it, such as tempo (bpm), form, complexity, variety and genre, and whether it is instrumental or vocal. It is therefore possible that certain music might be suitable during reading (in the sense of improving reading comprehension) for certain people. Additionally, since this study was carried out in shorter sessions, we do not know how the music would affect the readers in longer sessions.

Unexpectedly, none of the three major eye-movement measures in reading – fixation duration, saccadic amplitude and regressions – showed any significant main effects of condition in either first-pass or second-pass reading. The proportion of second-pass reading also did not show any significant effects of condition. Our expectations of longer fixation durations, more regressions, more second-pass reading and possibly shorter saccadic amplitudes, factors associated with impaired reading processes, were therefore not supported.

For the arousal measure of pupil size it was shown that both preferred and non-preferred music evoked higher levels of arousal compared to silence. No significant difference was found between café noise and silence, or between the noise conditions. Assuming that the non-preferred music engaged the participants in a negative way and the preferred music in a positive way, this is consistent with previous findings that negative and positive sounds are experienced as equally arousing (Partala and Surakka, 2003). This measure cannot, however, disambiguate negative experiences from positive experiences or say anything about emotional states. A more thorough investigation of how the noise conditions affected emotion and mood could have revealed other factors that may have influenced the results. For example, the non-preferred music is likely to have induced moods and emotions that affected the reading process negatively. Nonetheless, high “disruptiveness ratings” for the non-preferred music did not correlate with low reading-comprehension scores for that condition.

The results for extraversion revealed that there was a negative correlation between extraversion and the reading comprehension scores in the preferred music and the silent condition; that is, the group with a higher score in the extraversion dimension performed weaker in the reading comprehension. We suggest that this correlation does not say anything about how background music and extraversion interacts, but is rather an indication that those with a higher score on the extraversion dimension consisted of less skilled readers. It was also revealed that there was a positive correlation between extraversion and within-word regressions in the preferred music condition and a negative correlation between extraversion and the saccadic amplitude in all conditions; that is, the group with a higher score in the extraversion dimension performed more regressions within words during the preferred music and performed shorter saccades in general. Since frequent within-word regressions and shorter saccades tend to be typical of poor readers (e.g., Olson et al., 1983; Rayner, 1989), these findings strengthen the argument that those with a higher score in the extraversion dimension were in fact less skilled readers. A study where participants are selected on the basis of extremes in the EPQ-R distribution of the extraversion dimension will have to be conducted before any claims can be made in this respect. In this study only one participant had an EPQ-R score below eight, and only four scored above 19 (the lowest possible score is 0 and the highest is 23), which suggests that the included participants constituted a rather homogeneous group who were not at the extremes.
Psychology of Music

of extraversion or introversion (cf., Cassidy & MacDonald, 2008; Furnham & Allass, 1999; Furnham & Bradley, 1997).

Finally, no significant correlations were found on the basis of the OSPAN test or the questionnaire about study and music habits. All the participants were university students, and it could therefore be argued that the reason why there were no correlations with regard to OSPAN results was that the participants constituted a rather homogeneous group with relatively high working-memory capacity (OSPA N scores: mean = 27.6, SD = 5.9, max = 42). However, previous research seems to show that working-memory capacity is not extremely important for reading comprehension in the presence of aural distraction. For example, Martin, Wogalter and Forlano (1988) reported that the decisive factor for reading comprehension is the ability to understand meaning, not the ability to maintain phonological representations of written words in working memory.

As regards the questionnaire, it is notable that there were no correlations between reading-comprehension scores in the preferred music condition and the participants’ perception as to whether they studied better with music. This suggests that peoples’ personal beliefs about conditions which facilitate reading do not correspond to the actual effects.

Nonetheless, there are other individual factors that might interact with noise conditions in an experiment of this kind. Other examples include motivation and reading ability. Motivation is hard to measure, but it is a reasonable assumption that motivation will have decreased towards the end of the experiment. However, no correlations were found between scores on the reading-comprehension test and the order in which the texts were read, which could be interpreted to indicate that the participants remained equally motivated throughout the experiment even though nothing can be said about their relative levels of motivation. We also tried to revise the eye movement analyses by classifying high and low achievers on the basis of the reading comprehension scores in all conditions (median splits). But these analyses offered no further insights; that is, no significant differences in regards of fixation durations, saccadic amplitudes or regression were found in these analyses either.

Our results could be criticized for being an effect of state-dependent memory (e.g., Blaney, 1986); that is, the results could be perceived as an artefact deriving from the fact that, in three-fourths of cases, the reading process was performed in a different condition (with background music/noise) from the reading-comprehension test (in silence). Nevertheless, if you are studying for a real examination you are not allowed to listen to any kind of music during the actual test. Based on this fact we made the choice to have participants’ answer the reading comprehension in silence.

But how should we interpret the eye-movement data, where no significant differences were found between the silent condition and the non-preferred music condition for any of the three major eye-movement measures during reading? The results from the reading-comprehension scores clearly show that the reading process is disrupted by the non-preferred music. One interpretation would be that the readers are indeed not aware that their reading is affected and impaired by the non-preferred music. They do not make any changes to their reading process because they are simply not conscious that their reading comprehension deteriorates in the presence of the non-preferred background music. To obtain the same level of reading comprehension as in silence, they would probably have to make some changes to compensate for the increased cognitive load. For example, it would seem appropriate to slow down (longer fixation durations) and to do more re-reading (more regressions and more second-pass reading). It should be noted that the current study included an experimental restraint in that the participants were not able to go back to previous pages. This greatly reduced their opportunities to
re-read. It is possible that re-reading of certain parts of the texts could have compensated for the loss of reading comprehension.

One problem with the interpretation that conscious changes must be made to the reading process in order for reading comprehension to improve is that this assumes that we have a great deal of control over our eye movements while we are reading. The extent of eye-movement control in reading is an issue of debate (cf. Starr & Rayner, 2001), but there is much empirical evidence in favour of the assumption that fixation duration is automatically related to the amount of time required to process a word (e.g., Rayner, 1998; Reichle, 1998). Therefore it seems unlikely that the reading-comprehension impairment should be related to the lexical activation of words or sentences; the derivation and storage of the meaning and content of a text seems a more probable candidate.

It is often argued that reading comprehension requires you to maintain phonological representations of written words in working memory while the meaning is derived and stored in long-term memory (e.g., Baddeley, 1979, 2003; Levy, 1977). A possible explanation for why reading comprehension is impaired could therefore be that the background noise reduces the ability to maintain phonological representations in working memory. Since no significant correlations were found between the OSPAN scores and the reading-comprehension scores, this explanation was not supported. However, as mentioned previously, the participants constituted a rather homogeneous group with relatively high working-memory capacity. Therefore, this study does not attempt to make any strong claims in this respect. Nevertheless, in the study by Martin and colleagues (1988), where it was reported that reading comprehension decreased significantly with different kinds of speech in the background, it was concluded that the degree of disruptiveness was determined not by the phonological properties of the background speech but by the semantic ones. The authors concluded that the key to reading comprehension is not the ability to remember exact words and sentences, but rather the ability to understand the meaning of words and sentences. These findings and interpretations are consistent with a study by Boyle and Coltheart (1996), which investigated how irrelevant noise, such as singing, speech and instrumental music, affected comprehension of sentences at different levels of syntactic complexity and working memory in terms of the ordered recall of five-word lists. They found that the ability to recall the word lists was impaired by the irrelevant vocal noise but found no effect for instrumental music. No impact on sentence comprehension was found for any of the noise conditions.

Further studies will have to be conducted to identify the processes that are impaired by background noise during reading. Even so, however, our finding that the eye-movement measures were unaffected by the disruptive condition of non-preferred music together with the findings of Martin and colleagues (1988) and Boyle and Coltheart (1996) do indicate that the impairment of reading comprehension is not related to the maintenance of information in working memory, but rather to the processes involved in understanding meaning and content as well as to the ways in which this information is stored in long-term memory.

Outlook

This is the first study to investigate eye movements when reading under exposure to different kinds of music and noise. This study was designed to get a general idea of how eye movements are affected by background music/noise during reading. Future research will include more controlled studies focusing on specific details of both the texts and the background noise. For example, we will investigate how the reading process is affected by noise of increasing volume...
and the differential impact of exposure to speech versus instrumental music. Detailed analyses of eye movements in regards of syntactic and semantic aspects of the texts will also be considered in the future.

Acknowledgements
Thanks to Stipendiejuryn för Sparbanken Finn Framtidsstiftelses Forskarstipendium i musik for financing this project, to Högskoleverket (Swedish National Agency for higher Education) for providing the texts, to Marcus Nyström for programming help and of course to all the participants who made this study possible.

Notes
1. A Swedish readability index (Björnsson, 1968) based on the proportion of long words and the number of words per sentence.
2. Sweet, highly sentimental ballads with a simple, catchy melody or light pop tunes (Wikipedia, 2009).

References


### Appendix A. Preferred and non-preferred music

<table>
<thead>
<tr>
<th>Participant</th>
<th>Preferred music</th>
<th>Non-preferred music</th>
</tr>
</thead>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Participant</th>
<th>Preferred music</th>
<th>Non-preferred music</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.</td>
<td>Belle and Sebastian – The boy with the arab strap (tracks: “It could have been a brilliant career”, “Sleep the clock around”. Is it wicked not to care?)</td>
<td>Slayer – Christ Illusion (tracks: “Catalyst”, “Skeleton Christ”, “Eyes of the insane”, “Consfearacy”)</td>
</tr>
</tbody>
</table>
Roger Johansson is a PhD student in cognitive science, Lund University.

Kenneth Holmqvist is an associate professor and technical manager for the Humanities Laboratory, Lund University.

Frans Mossberg is a researcher at the division of musicology, Lund University.

Magnus Lindgren is an associate professor in neuropsychology, Lund University.