A Quasi-Experimental Investigation into the Impact of Typical Levels of Classroom Noise on Adolescents' Reading Comprehension

Daniel Connolly[1], Julie E. Dockrell[2,]
Charlie Mydlarz[3], Bridget Shield[4], Rob Conetta[4], Trevor Cox[3]

[1] Southampton Solent University
(Daniel.Connolly@solent.ac.uk)
[2]Institute of Education, London
[3] University of Salford
[4] London South Bank University

Abstract
This paper reports the findings of two quasi-experiments designed to investigate the effects of typical levels of classroom noise on the reading comprehension of pupils aged 11- to 16-years. A novel reading task using science texts was designed to assess reading comprehension and word learning. The task was completed by 1076 secondary school pupils (age range: 11- to 16-years) on laptop computers while edited recordings of classroom noise was played to pupils through headphones. Accuracy of responding, the time taken to read each article and time taken to respond to questions were recorded. In Experiment 1, pupils who completed the task with noise presented at 50 dB L_{Aeq} took significantly longer to respond and were more accurate than pupils who completed the task in 70 dB L_{Aeq}. Experiment 2 found no differences between groups of pupils who completed the task in 50 dB and 64 dB conditions. The findings are discussed in terms of the disruptive effects on reading comprehension and learning of classrooms with unfavourable acoustics.

Keywords: Classroom noise, reading comprehension, word learning, high school pupils.

PACS no. 01.40.ek, 43.50Jh, 43.50.Qp

1 Introduction

Research has found that adults' reading comprehension is negatively affected by background noise compared to quiet conditions, with the greatest disruption caused by unattended speech and speech-like sounds [1, 2, 3]. Studies with high school pupils have found negative effects of speech-like noise and road traffic noise on recall of text and for performance on free recall of visually presented sentences in background white noise compared to quiet [4]. However, little is known about the type of noise that is most detrimental to reading and the levels at which these disruptive effects become apparent. Similarly, the effects of classroom noise on adolescents' learning of complex new words from text have not been examined.

Reading differs from listening in that readers are able to re-read sections of text they did not understand on first reading or to vary their speed when reading passages that are difficult to understand [5]. The ability to reread and alter reading speed may reduce the impact of noise on the processing of written text. Alternatively, irrelevant noise may disrupt the reading process as has been demonstrated in [6, 7,8] or background speech may impact on the verbal process of reading as it does on office activities [8, 9, 10, 11].

Two experiments were designed to investigate the effects of classroom noise on high school pupils' reading comprehension. The reading task consisted of a series of graded passage and comprehension questions probed memory of factual information and ability to make inferences about information not explicitly stated in the texts. Word learning was also assessed by a question related to the meaning of a novel word that had been presented at the start of each passage.

In addition, the time taken to read the articles and the time taken to respond to each question was recorded. Word reading speed is an established predictor of reading comprehension [12, 13] and it was anticipated that noise levels would interfere with this process. Because individuals vary in terms of the time taken to process reading comprehension questions [14] and how susceptible they may be to distraction during testing, response latencies offer a sensitive way to measure the effects of noise level in cases where accuracy measures may not be sufficiently sensitive to identify changes. These measures would allow discrimination between the two possible ways in which background noise might interfere with reading. Unfavourable noise levels might produce a speed-accuracy trade off, such that reduced accuracy is reflected in shorter reading times and response latencies and poorer accuracy. By contrast, unfavourable noise levels might affect reading comprehension because screening out the noise places an additional load on cognitive processes, resulting in slower reading and longer response latencies and poorer accuracy.

The study comprises two experiments: Experiment 1 compared effects of the maximum (70 dB L_{Aeq}) and minimum (50 dB L_{Aeq}) ambient classroom noise levels measured in a study of English classrooms [15]; Experiment 2 then compared the effects at the minimum recorded noise level (50 dB L_{Aeq}) with a more moderate upper noise level (64 dB L_{Aeq}). It was predicted that reading comprehension, as measured by response accuracy, would be worse in the louder conditions. Moreover, it was hypothesized that unfavourable noise levels would also hamper word learning from texts. Shorter reading times and response latencies in the louder noise conditions would indicate the operation of a speed/accuracy trade off in response to noisy conditions: pupils respond more quickly but less accurately. Conversely, longer reading times and response latencies in the louder conditions would indicate that high noise levels add to pupils' cognitive load, resulting in fewer correct answers and slower responding.

1.1 Experiment 1

Experiment 1 compares the effects of the maximum and minimum noise levels (70 dB L_{Aeq} and 50 dB L_{Aeq}) obtained from classroom recordings [15] at two time points.

1.2 Method

1.2.1.1 Participants

A total of 669 students with a mean age of 13.34 (SD = 1.5) participated in the study, from five English high schools: 83 11-year-olds (12.4%); 145 12-year-olds (21.7%); 133 13-year-olds (20%); 141 14-year-olds (21%); 104 15-year-olds (15.5%); and 63 16-year-olds (9.4%).

1.2.1.2 Design

Participants completed two test sessions separated by two weeks. Each test session took place on the same day of the week and at the same time of day on both occasions. The order of the quiet and loud conditions was counterbalanced across classes, with half of the participants in each year group receiving the quiet condition first and half receiving the loud condition first. Two sets of reading materials were counterbalanced across conditions, with half of the participants in each noise condition receiving set A first and half receiving set B first.

1.2.1.3 Materials

1.2.1.3.1 Classroom Noise Stimulus

The noise stimulus was constructed from recordings of classes of Year 8 pupils (12- to 13-year olds) as they engaged in 'individual work' (completing tasks set by the teacher but allowed to discuss the work with classmates) in a cellular classroom. The recordings consisted of a background of unidentifiable speech (babble) and sound events (chair scrapes, pencil drops, occupant movement). Eight unique but acoustically identical segments of background noise were combined to create a background noise stimulus with a total duration of 4 minutes 40 seconds. Levels of the background noise stimulus in the quiet and loud conditions were determined from the observed maximum and minimum L_{Aeq} levels of noise in classrooms as pupils engaged in individual work of 50 dB and 70 dB respectively [15]. Filters were then applied to the signal to correct for the frequency response of the headphones and to ensure that calibrated dB levels were reproduced faithfully.

1.2.1.3.2 Reading Task

Two sets of materials were developed, consisting of 4 articles adapted from a science news stories on children's science education websites. All articles were adapted to be 180 words in length and contain an average of 1.2 polysyllabic words per sentence. Five multiple-choice questions accompanied each article. Two of these questions probed memory for factual information contained explicitly in the text. Two more questions probed ability to infer information not explicitly stated in the text. A final question probed learning of a single polysyllabic word contained in the title page of each text.

1.2.2 Procedure

The task was programmed using E-Prime version 2.0.8.9 software. Classes of pupils completed the task simultaneously on Acer Aspire 1810TZ mini laptops. All tasks were triggered from a master laptop via a wireless network. The noise was replayed in mono through Beyerdynamic Dtime 100 headphones.

Test sessions took place in participating pupils' usual science room under the supervision of a teacher and two experimenters. An experimenter gave verbal introductions about the task and then instructed participants to enter their names and ages onto the laptops. Responses were anonymised once data from the two test sessions had been combined.

Each article was divided into three sections of text, each one 60 words in length. At the start of each text, participants read a title page featuring a polysyllabic word describing the subject matter of the text, along with an explanation of its meaning. For example 'Selenology: the Study of the Moon'.

Five multiple-choice response options (A, B, C, and D) were presented on the screen below each question. Questions 1 and 2 related to factual information contained explicitly in the preceding section of text. Questions 3 and 4 were inferential, and could only be answered correctly using information implied in the preceding section of text. Question 5 probed memory for the polysyllabic word introduced in the title page. The position of correct responses was randomized. In summary, each article was presented in the following sequence: Title page (containing science word tested subsequently); Text section 1 – Question 1 (factual); Text section 2 – Questions 2 (factual) and 3 (inferential); Text section 3 – Questions 4 (inferential) and 5 (word learning). Response latencies for each question and the reading time for each section of text were recorded automatically by the E-prime software. Timing for each item was initiated when the spacebar was pressed to advance to that item and terminated when the spacebar was pressed to advance to the next item.

When the three text sections and five questions from each article had been completed, participants progressed onto the next article. The task was time-limited to four minutes in total, timed from initiation of the first title page. If the time limit was reached when participants were halfway through a section of text, they were permitted to complete that section and its associated questions and these responses were included in the analysis. After reaching the time limit and completing the last question for that section, participants were presented with a screen containing the prompt 'That's it, all done'.

1.3 Results

1.3.1 Effect of Background Noise Level (50 dB versus 70 dB) at Time 1

Performance on the factual, inferential and word learning questions was assessed using the total number of correct responses for each question type. Mean article reading times were calculated by averaging participants' time to read the 60-word sections of text. To control for distraction and erroneous responses due to possible confusion or lack of motivation, only response latencies for only correct answers were analysed. Missing data from participants that attended only one testing session were replaced with the mean for each variable. Mean question response latencies for each type of question and article reading times were logarithmically transformed to correct for violations of assumptions of normality. To simplify presentation and interpretation, response latencies and article reading times were converted to z-scores. As the difference between the mean age of participants in Q1 ($n = 335$ $M = 13.58$, $SD = 1.49$) and in L1 ($n = 334$, $M = 13.10$, $SD = 1.49$), was statistically significant, $t\ (667) = 4.22$, $p < .001$, age was entered as covariate into the analyses. To investigate the effect of level of classroom noise at time 1, the number of questions attempted, the number of correct responses (response accuracy), and z-scores for reading times, response latencies and coefficients of variation in each condition were entered into a multivariate analyses of covariance (MANCOVA) with condition as between participant variables.

The numbers of questions attempted, correct responses, article reading times and response latencies to each question type at time 1 in Experiment 1 are displayed in Table 1. The MANCOVA revealed statistically significant main effects of condition, $F\ (12, 655) = 719.05$, $p < 0.001$, $P\eta^2 = 0.93$, and of the of the covariate age, $F\ (12, 655) = 6.913$, $p < 0.001$, $P\eta^2 = 0.11$. As shown in Table 1, there were more correct answers to all question types and longer reading times and response latencies in the 50 dB condition. Follow up univariate ANOVAs revealed a statistically significant effect of condition on correct responses to the factual questions, $F\ (1, 666) = 4.64$, $p = 0.03$, $P\eta^2 = 0.01$, with more correct responses to factual questions in the 50 dB condition, and on article reading times, $F\ (1, 655) = 4.60$, $p = 0.03$, $P\eta^2 = 0.01$, with longer reading times in the 50 dB condition. There was also a statistically significant effect of condition on response latencies to the inferential questions, $F\ (1, 666) = 5.02$, $p = 0.03$, $P\eta^2 = 0.01$, and word learning questions, $F\ (1, 666) = 9.60$, $p = 0.01$, $P\eta^2 = 0.04$, with longer response latencies in the 50 dB condition for both types of question. There were no statistically significant effects of condition on the total number of questions attempted, $F\ (1, 666) = 0.63$, ns, correct responses to inferential questions, $F\ (1, 666) = 1.70$, ns, correct responses to word learning questions, $F\ (1, 666) = 1.93$, ns, or on response latencies to correct responses to factual questions, $F\ (1, 666) = 1.49$, ns.

The follow up ANOVAs yielded statistically significant effects of age on the number of questions attempted, F (1, 666) = 9.56, p = 0.002, Pη2 = 0.01: and on the number of correct answers to factual questions, F (1, 666) = 24.91, p < 0.001, Pη2 = 0.04; inferential questions, F (1, 666) = 41.87, p < 0.001, Pη2 = 0.06, and word learning questions, F (1, 666) = 18.81, p < 0.001, Pη2 = 0.03. There was also a statistically significant effect of age on response latencies to correct answers to the word learning questions, F (1, 666) = 24.56, p < 0.001, Pη2 = 0.04. There were no effects of age on article reading times, response latencies to correct factual questions, or inferential questions, *all ps* n.s.

Table 1: Age adjusted mean number of questions attempted and answered correctly, and z-scores of time based measures, Time 1, Experiment 1 [95% CI]

Dependent Variable		Condition			
		50 dB		70 dB	
	Question Type				
Total Number of Questions Attempted		11.71	[11.41, 12.02]	11.35	[11.05, 11.66]
Correct responses to questions	Factual	3.04*	[2.90, 3.18]	2.82	[2.68, 2.96]
	Inferential	2.08	[1.94, 2.21]	2.03	[1.90, 2.17]
	Word Learning	1.02	[0.94, 1.10]	0.94	[0.86, 1.02]
Reading Time		0.08*	[-0.02, 0.18]	-0.08	[-0.19, 0.02]
Response Latencies (Z-scores)	Factual	0.05	[-0.06, 0.15]	-0.05	[-0.15, 0.06]
	Inferential	0.08*	[-0.02, 0.18]	-0.08	[-0.18, 0.02]
	Word Learning	0.10**	[0.01, 0.18]	-0.10	[-0.18, -0.01]

Note: Age evaluated at 13.34
* P < 0.05; ** P < 0.005

In sum, there were subtle but statistically significant effects of background noise level on reading comprehension. The main effect of condition in the initial MANCOVA was statistically significant, with a large effect size. Performance on all types of question was less accurate in the 70 dB condition compared to the 50 dB condition, and this difference was statistically significant although small for the number of correct responses to factual questions. Time taken to read and process the information was also affected by background noise level: there was a consistent trend for response latencies to be longer in the 50 dB condition and this effect was statistically significant, although again small, for response latencies to the word learning question. Similarly, article reading times were significantly longer in the 50 dB condition.

1.4 Experiment 2

Experiment 2 compared performance on the reading task at 50 dB and a more moderate louder level of 64 dB.

1.4.1 Method

1.4.1.1 Participants

A total of 307 pupils with a mean age of 12.78 (SD = 1.24) participated in the study, from 3 English secondary schools: 37 11-year-olds (12%); 133 12-year-olds (43%); 33 13-year-olds (11%); 70 14-year-olds (23%); 34 15-year-olds (11%).

1.4.1.2 Design

As in experiment 1, participants completed two conditions, separated by two weeks. Half of the pupils in each year group were allocated to one of two session orders:50 dB followed by 64 dB) and 64 dB followed by 50 dB. Because of strong order effects, only the results for time 1 are analysed and reported here.

1.4.1.3 Materials and Procedure

The reading task was the same as that used in experiment 1.

1.4.2 Results

Data management and statistical analyses procedures were the same as in experiment 1. As in experiment 1, there was a statistically significant difference between the mean age of participants in the two conditions at time 1, (50 dB: n = 161, mean age = 12.44, SD = 1.20; and 64 dB: mean age = 13.14, SD = 13.14, SD = 1.19), $t\,(305) = -5.16$, $p < .001$. The number of questions attempted, number of correct responses, and z-scores of all time-based measures - article reading times, response latencies to questions answered correctly for each question type were entered as the dependent variables in a MANCOVA with age as the covariate.

1.4.2.1 Effect of Noise Condition (50 dB versus 64 dB) at Time 1

The means for the number of questions attempted, number of questions answered correctly, z-scores of article reading times and response latencies to all question types are displayed in Table 2. The MANCOVA revealed statistically significant main effects of condition, $F\,(12, 293) = 447.91$, $p < 0.001$, $P\eta^2 = 0.95$, and of the covariate age, $F\,(12, 293) = 2.92$, $p = 0.001$, $P\eta^2 = 0.11$. However, there were no statistically significant effects of condition on any of the individual variables, all *p*s n.s. The follow up ANOVAs revealed statistically significant effects of the covariate age on all variables: all *p*s < 0.05: older children achieved more correct answers and shorter response latencies.

In sum, there were few differences in performance between pupils in the 50 dB and 64 dB conditions. All measures were significantly affected by the covariate age.

Table 2: Age adjusted mean numbers of questions attempted and answered correctly, and z-scores of time-based measures (presented as z-scores), Experiment 2

Dependent Variable		Condition 50 dB		64 dB	
Total Number of Questions Attempted		11.78	[11.36,12.19]	11.43	[10.97, 11.89]
Number of Questions Answered Correctly	Factual	3.39	[3.18, 3.60]	3.59	[3.36, 3.81]
	Inferential	2.36	[2.17, 2.55]	2.36	[2.16, 2.56]
	Word Learning	1.16	[1.03, 1.29]	1.21	[1.08, 1.35]
Article Reading Time		0.03	[-0.12, 0.18]	-0.03	[-0.19, 0.13]
Response Latencies	Factual	-0.04	[-0.19, 0.12]	0.04	[-0.12, 0.20]
	Inferential	0.05	[-0.10, 0.20]	-0.05	[-0.21, 0.10]
	Word Learning	0.03	[-0.11, 0.17]	-0.03	[-0.18, 0.11]

Note: Age evaluated at 12.78

2 Conclusions

The findings of this study have supported the hypotheses that high levels of background noise in high school classrooms have disruptive effects on adolescents' reading comprehension and word learning from text. Experiment 1 compared performance of high school pupils on a reading task in relatively quiet (50 dB LAeq) and loud (70 dB LAeq) conditions. Better performance in the quiet condition was revealed in significantly more correct answers to factual questions; there was also a non-significant trend towards more accurate responding to inferential and word learning questions in the quiet condition. Experiment 1 also produced evidence that better accuracy in the quieter noise condition was underpinned by a speed/accuracy trade-off: article reading times were significantly longer in the 50 dB condition, whereas pupils in the 70 dB condition read more quickly and accuracy was lower. The time taken for pupils to respond to questions also reflected this trade-off: response latencies to inferential and word learning questions were longer and there was a non-significant trend towards more questions answered correctly in the 50 dB condition. Similarly, significantly more factual questions were answered and there was a trend towards longer response latencies correctly in the 50 dB condition.

The disruptive effects of the loud condition were mostly evident when classroom noise was presented at 70 dB LAeq, representing the upper level of noise recorded in English high school classrooms during individual work [15]. In experiment 2, when noise was presented at a more moderate 64 dB LAeq, there were no significant differences in either the numbers of questions answered correctly or in response latencies compared to the 50 dB condition.

The findings of the experiments reported here underline the importance of good acoustical conditions in classrooms for adolescent learners. Much of teaching and learning in secondary schools occurs through written text in the form of books, worksheets or instructions. Competency in reading enables the learner to develop in-depth understanding of the material being studied and facilitates the learning of new, topic-relevant vocabulary. Future research needs to examine the differential effects of poor acoustics on the reading comprehension of learners with additional needs and specific learning disorders such as dyslexia.

Acknowledgements

This research was funded by the U.K. Engineering and Physical Sciences Research Council.

References

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

[2] Oswald, C. J. P., Tremblay, S., & Jones, D. M. Disruption of comprehension by the meaning of irrelevant sound. *Memory,* 8(5), 2000. 345-350.

[3] Sorqvist, P., Halin, N., & Hygge, S. Individual Differences in Susceptibility to the Effects of Speech on Reading Comprehension. *Applied Cognitive Psychology*, 24(1). 2010. 67-76.

[4] Hygge, S., Boman, E., & Enmarker, I. The effects of road traffic noise and meaningful irrelevant speech on different memory systems. *Scandinavian Journal of Psychology*, 44(1). 2003. 13-21

[5] Cain, K. *Reading Development and Difficulties*. Chichester :BPS Blackwell. 2010.

[6] Jones, D. M., Madden, C., & Miles, C. Privileged access by irrelevant speech: the role of changing state. *Quarterly Journal of Experimental Psychology*, 44A. 1992. 645–669.

[7] Tremblay, S., & Jones, D. M. Change of intensity fails to produce an irrelevant sound effect: Implications for the representation of unattended sound. *Journal of Experimental Psychology-Human Perception and Performance*, 25(4). 1999. 1005-101

[8] Banbury, S., & Berry, D. C. Habituation and dishabituation to speech and office noise. *Journal of Experimental Psychology-Applied*, 3(3). 1997. 181-195

[9] Banbury, S., & Berry, D. C. Disruption of office-related tasks by speech and office noise. *British Journal of Psychology*, 89(3). 1998. 499-517

[10] Banbury, S., & Berry, D. C. Office noise and employee concentration: Identifying causes of disruption and potential improvements. Ergonomics. 2005. 48:1, 25 – 37

[11] Beaman, C.P. Auditory Distraction from Low-Intensity Noise: A Review of the Consequences for Learning and Workplace Environments. Applied Cognitive Psychology. 19. 2005. 1041–1064

[12] Ehri, L. C. Grapheme–phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), Word recognition in beginning literacy. 1998. pp. 3–40. Mahwah, NJ: Erlbaum

[13] Fuchs, L. S., Fuchs, D., & Maxwell, L. The validity of informal measures of reading comprehension. Remedial and Special Education, 9. 1998. 20–28.

[14] Jackson, M.D. & McClelland, J.L. Sensory and cognitive determinants of reading speed. Journal of Verbal Learning and Verbal Behaviour, 14. 1975. 565-74

[15] Shield, B. M., Conetta, R., Dockrell, R., Connolly, D., Cox, T. & Mydlarz, C. (2015) A survey of acoustic conditions and noise levels in secondary school classrooms in England. J. Acoust. Soc. Am. 137 (1). 2015. 177-188