Buenos Aires – 5 to 9 September, 2016
Acoustics for the 21st Century…

PROCEEDINGS of the 22nd International Congress on Acoustics

Acoustics in Education: ICA2016-241

Effects of aircraft noise on children's reading and quality of instruction in German primary schools: results from the NORAH study

Jan Spilski[(a)], Kirstin Bergström[(a)], Ulrich Möhler[(b)], Jochen Mayerl[(a)], Thomas Lachmann[(a)], Maria Klatte[(a)]

[(a)] University of Kaiserslautern, Germany, jan.spilski@sowi.uni-kl.de
[(b)] Moehler + Partner Consulting Engineers Munich, Germany

Abstract

Earlier studies mostly show that chronic exposure to aircraft noise may impair children's cognitive development, especially reading acquisition. In the framework of the NORAH-study, the effects of aircraft noise on reading and verbal precursors of reading acquisition were investigated in 1,090 second-graders from 29 schools in the vicinity of Frankfurt/Main Airport. Aircraft noise levels at schools and at the children´s homes were calculated based on radar data from the Flight Track and Monitoring System (FANOMOS), provided by German Air Traffic Services. Although aircraft noise levels at schools did not exceed 60 dB (LAeq 8-14) and were thus considerably lower when compared to prior studies, multilevel analyses revealed significant effects of aircraft noise on children´s reading comprehension after adjustment for individual (e.g. socioeconomic status) and class-level (e.g., road traffic noise, classroom insulation) factors. A 10 dB increase in aircraft noise was associated with a decrement of one-tenth of an *SD* on the reading test, corresponding to about one month reading delay in this test. Sensitivity analyses confirmed the robustness of the results. Teachers' reports indicate impairments of classroom instruction due to aircraft noise. This relationship could be a possible reason for the mainly negative effects of aircraft noise on reading performance.

Keywords: Aircraft Noise, Children, Reading, Instruction

Effects of aircraft noise on children's reading and quality of instruction in German primary schools: results from the NORAH study

1 Introduction

Prior studies found mostly negative effects of aircraft noise exposure on children's cognitive performance [1,2]. Across studies, exposure to aircraft noise was mainly associated with lower reading performance [2,3]. However, results are not that consistent as it is reported in some reviews [e.g. 3,4]. For example Haines et al. [6] found no decreasing reading scores with increasing aircraft score, when models are adjusted for socioeconomic status (SES). Furthermore, to the best of our knowledge, only three studies analyzed exposure-effect relationships [5,6,7].

Compared to previous studies, the average aircraft noise levels have decreased. Even the most exposed schools considered in the NORAH sample have levels comparable to schools included as unexposed controls in prior studies [1].

Existing noise research focused mainly on the direct relationships of noise on children's cognition. However, in a post-hoc analysis of a teacher survey in the RANCH study it was shown that 20-25% of aircraft noise exposed teachers reported disorders of communication, attention, and concentration of students [8]. Overall, the impacts of aircraft noise on teaching conditions have hardly been explored, although it could be a possible reason for the predominantly negative effects of aircraft noise on reading performance.

Despite the existing evidence, knowledge concerning effects of chronic aircraft noise exposure on children is still limited and, for reasons stated elsewhere [9], does not allow well-founded predictions for children from specific noise-exposed areas such as the Rhine-Main-Region surrounding Frankfurt/Main Airport in Germany. The current study was performed in the context of the NORAH (Noise-Related Annoyance, Cognition, and Health) project, an interdisciplinary research project on the effects of transportation noise on citizens of Rhine-Main. Here we summarize the NORAH-subproject concerning effects of aircraft noise exposure on children's reading and quality of instruction in primary school in the Rhine-Main region. A full report of this study is provided elsewhere [9].

2 Methods

2.1 Participants

A total of 1,243 second-graders and 84 class teachers from 29 primary schools participated in the study. From a total of 297 schools, 29 schools were selected based on extent of aircraft noise exposure. Those schools exposed to the highest amounts of aircraft noise were selected first. The remaining schools were selected using a-priori direct matching to avoid potential confounding factors (for example, SES). The schools were matched by indicators of the pupils´ socioeconomic status, migration background, and German language proficiency, according to the headmasters´ reports. For cognitive performance, complete data from the parent questionnaires and the test battery were available for 1,090 children. For quality of instruction, complete data were available for 84 teachers (78 female). Mean age for children was 8 years and 4 months (*SD* 5 months), and 60% of the children had a migration background.

2.2 Assessment of Noise Exposure

Aircraft noise levels (LAeq 08-14) at school and at the children´s home address for the past 12 months before data acquisition were calculated based on radar data from the Flight Track and Aircraft Noise Monitoring System (FANOMOS), provided by German Air Traffic Services. For more details see [10]. Road traffic and railway noise were calculated using a combination of information (e.g., traffic flow data, street types, proportion of heavy traffic; quantity of train runs, speed and length of the trains) provided by local authorities. Classroom reverberation and insulation were assessed through screening procedures.

2.3 Tasks and Materials

Reading and verbal precursors of reading acquisition (e.g. speech perception, short-term memory, phonological awareness) were assessed through standardized paper-and-pencil tests administered in groups of whole classes. All materials are age-appropriate. Quality of instruction was assessed with a teacher questionnaire. One factor with three statements focused on impairments of school lessons through aircraft noise (table 1). Potential confounding factors such as socioeconomic status, migration background, non-verbal abilities, classroom insulation and exposure to road traffic and railway noise were also assessed. More details on tests and questionnaire are provided in [9].

Table 1: Factor: impairments of school lessons through aircraft noise

Item
1. Due to aircraft noise, I have to interrupt my talk/the discourse for a short time
2. During the lessons, the children are distracted by aircraft noise
3. During the lessons, aircraft noise is audible even when the windows are closed

Note. We used a 5-point rating scale (never, seldom, sometimes, often, and very often)

2.4 Procedure

The tests were performed in groups of whole classes. The speech materials were presented via wireless headphones, in order to ensure perfect signal quality at each working place and thus eliminate acute effects of classroom reverberation and noise from outside. Each task was carefully explained to the children and practiced with examples. All statements were read aloud by the experimenter and a combination of picture and number represented the statement. All in all, the testing session in the classroom took about three lessons. The children were given a questionnaire to complete at home by their parents, which contained questions concerning socioeconomic status and other potential confounding factors (e.g., main language spoken at home, parental support in school work). The teacher questionnaire was given to the teachers before the children's testing session and collected immediately after it.

2.5 Statistical Analyses

2.5.1 Preliminary Analyses

Missing analyses and imputation

Excluding cases with single missing values can reduce the statistical power and bias the parameters of analysis. Multiple data imputation is one of the best techniques to maintain statistical power and reduce potential bias in estimates of parameters [11].

Before conducting the imputation, we tested for a possible aircraft-related bias in patterns of missing values. We found no correlations or systematic patterns. Due to the relatively low amount of missing values (average < 2%) that was also distributed over a relatively high number of items and participants, missing values can be assumed to be unsystematic and missing at random [11]. Therefore multiple imputation is an adequate technique of missing value treatment. We use the "two-way imputation with normally distributed errors" [12] for which SPSS macros are available.

Psychometrics

In order to test the construct validity and reliability of the questionnaire scales, principal axis factoring (PAF), confirmatory factor analysis (CFA), and calculation of Cronbach's alpha were conducted. Calculations were related to the thematic blocks (e.g., home and school) and were carried out separately for the children's and teacher's judgments.

We tested the psychometric quality using a stepwise process with the following order: First, we used explorative factor analysis with oblique rotation in SPSS to examine the factor structure. As the criterion to determine the number of factors, we employed three extracting methods (eigenvalue > 1, screeplot analysis, and parallel analysis according to Horn). The methods lead to different models with differences in the number of factors. Second, to derive the model with the best fit of empirical data to theoretical assumptions, we used CFA. We calculated a parsimony criterion (AIC) and a chi-square difference test to identify the most appropriate model. Third, we evaluated the factors´ discriminant validity using Fornell and Larckers test [13], which revealed that each factors´ average variance explained (AVE) is greater than its squared correlation with any other factor. Forth, internal consistency was assessed by calculating Cronbach's alpha values. Table 2 shows psychometric adequacy of potential negative factors on school lessons.

The factor *impairments of school lessons through aircraft noise* (factor 1), for example, has only item loadings above .6 with no higher cross-loadings on another factor (items were shown in table 1). The AVE was .73 and therefore greater than the Fornell/Larcker criterion (table 2). The composite reliabilities (CR) of scales were excellent (CR =.93) as was the internal consistency (α = .93).

Table 2: Factors' psychometric adequacy "impairments of school lessons through noise sources"

Factor	CR	AVE	α	1	2	3	4	5
1. Aircraft noise	.93	.84	.94	1.00				
2. road traffic noise	.93	.83	.93	.10	1.00			
3. Adjacent room noise	.87	.69	.87	.36*	.32*	1.00		
4. railway noise	.78	.55	.72	.18	.44*	.05	1.00	
5. Schoolyard noise	.87	.69	.86	.33*	.34*	.60***	.19	1.00

Note. CR = composite reliabilities; *AVE* = average variance extracted*;* α = Cronbach's alpha; the remaining values indicate correlations between factors. * $p < .05$, *** $p < .001$.

2.5.2 Multilevel Analyses

To analyze the impact of aircraft noise exposure on children´s abilities and quality of instructions, multilevel analyses (MLA) were realized using Mplus 7 [14]. Multilevel modeling is necessary in studies with a hierarchical structure of the data (children grouped within classes) in order to avoid misspecifications of parameters, e.g. underestimation of standard errors [15]. We used two-level random intercept models and adjusted for confounding variables. The final (fully adjusted) model was adjusted for Level 2 and Level 1 variables [9]. Factors (see 2.5.1 psychometrics) from the teacher questionnaire with respect to the classroom activity (e.g. disruptions of instructions, distractions) were included as dependent variable while adjusting for Level 2 confounders. Following all main analyses, we conducted post-hoc-matching methods to test the robustness of estimates.

2.5.3 Post-hoc Analyses of robustness

There are different matching approaches to control for a possible selection bias post-hoc: e.g., one-to-many matching, full-matching, or nearest-neighbor-matching. We used the propensity score-matching (PSM) as a form of the nearest-neighbor-matching with calipers of width equal to *SD* = .20 of the logit of the propensity score [16].

3 Results

3.1 Reading

Aircraft noise levels at school ranged from 39 to 59 dB (LAeq, 8-14), with mean *M* = 50 and *SD* = 6. Aircraft noise exposure at school was significantly associated with a decrease in children's reading after full adjustment. A 20 dB increase of aircraft noise at school was associated with a decrease in children's global reading scores by one fifth of an *SD*, corresponding to a two-months reading delay in this test. For reading outcome variables, the associations between aircraft noise was described best with a linear function. Multilevel model statistics and exposure-effect curves are provided in [9]. We conducted post-hoc analyses of robustness to control for a possible selection bias. Due to matching, the final sample sizes lay in a range of *n* = 708 to *n* = 1006. For the matched samples, MLA were calculated and it was checked whether the estimation results remain stable (table 3). We found large overlaps of propensity scores in all subsamples and no extreme values of propensity scores. The comparison of multilevel models between the original sample and the matched samples showed no differences in the direction and strength of relationships (*b*) and intercepts (see table 3). Therefore, although the sample was reduced, we found support for successful a-priori direct-matching and the robustness of the results without significant distortion.

Table 3: Fully Adjusted Multilevel Model for Effects of Aircraft Noise on Children's Reading Abilities (Global Score), for the Whole Sample, Median and Extreme Groups Propensity Score Matching (PSM)

	Final Model (*N* = 1090)		Median (PSM) Match (*N* = 1006)		Extreme groups (PSM) Match (*N* = 708)	
	b (SE)	*p*	*b (SE)*	*p*	*b (SE)*	*p*
Global reading score						
(*ICC* = .081/.035/.052)						
Intercept	45.94 (0.534)		45.74 (0.539)		45.63 (0.644)	
Aircraft noise school – Level 2	-0.097 (0.050)	.027	-0.119 (0.049)	.007	-0.107 (0.056)	.029
Word reading						
(*ICC* = .091/.057/.087)						
Intercept	46.45 (0.640)		46.27 (0.646)		46.06 (0.792)	
Aircraft noise school – Level 2	-0.105 (0.064)	.049	-0.128 (0.062)	.020	-0.130 (0.072)	.035
Sentence reading						
(*ICC* = .090/.036/.045)						
Intercept	45.03 (0.543)		44.80 (0.541)		44.80 (0.644)	
Aircraft noise school – Level 2	-0.064 (0.056)	.125	-0.091 (0.056)	.053	-0.065 (0.058)	.130
Text reading						
(*ICC* = .062/.018/.028)						
Intercept	46.33 (0.570)		46.14 (0.587)		45.79 (0.668)	
Aircraft noise school – Level 2	-0.118 (0.045)	.005	-0.133 (0.044)	.002	-0.124 (0.051)	.008

Note. ICC = intra-class correlation; *SE* = Standard error. Adjusted for age, gender, SES, number of children's books, non-verbal abilities, story comprehension, phonological awareness, access to phonological representations, class SES, class size, percentage of children with a migration background, parental involvement in school affairs, classroom insulation, road-traffic noise and railroad noise at school.

3.2 Quality of instructions

Aircraft noise exposure at school was significantly associated with an increase in impairments in school lessons (factor) after full adjustment (b = 0.144, SE = .013, p < .001, 95%CI: 0.118, 0.168). The model explains 63% of the variance of the factor *impairments of school lessons through aircraft noise.* A 10 dB in aircraft noise was associated with an increase of 1.44 points on the 5-point rating scale, corresponding to an increase of 1.37 *SD*. A subsequently calculated exposure-effect curve is shown in Figure 1. The approach is similar to the one described in [9]. A linear function showed the best fit for this association. Figure 1 shows a very strong linear effect; however, the adjusted mean was higher than a rating of three only in the highest 5 dB band (< 55 ≤ 60).

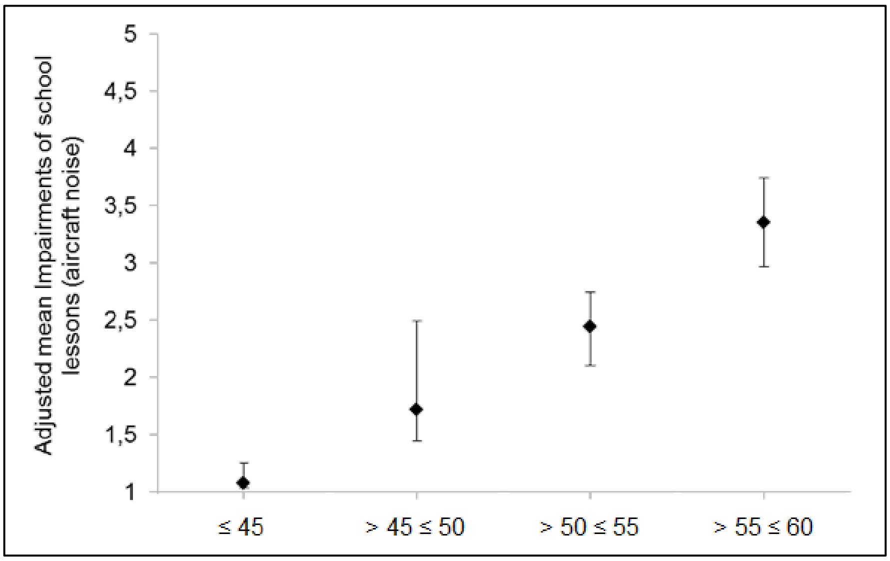

Figure 1: Exposure-effect relationship for the factor *impairments of school lessons through aircraft noise*. *Note.* Adjusted mean factor impairments (Index), with 95% bias corrected bootstrapping *CI* for 5dB bands of aircraft noise at school [LAeq 08-14 (dB)]. *CI* = confidence interval. Adjusted for class SES, class size, parental involvement in school affairs, classroom insulation, road-traffic noise and railroad noise at school.

Inspections of single items are conducted to define more clearly the effect we mentioned above. Out of 21 teachers from highly exposed schools, 11 (52%) reported frequent interruptions of classroom discourse and observable distractions of the children due to aircraft noise. In the less exposed groups, frequent interruptions of discourse and distractions of the children were reported by just one out of 63 teachers (2%). For each item, the group differences in response frequencies were significant (p < .001).

4 Discussion and Conclusion

In the current study, harmful effects of aircraft noise exposure were found for children´s reading and quality of instructions, although the exposure levels were considerably lower compared to prior studies. Exposure levels in NORAH did not exceed 60 dB, whereas in the

RANCH-study [7] aircraft noise levels at school reached 77 dB (LAeq 06-23). Extensive preliminary analyses with respect to missing values and psychometric quality were conducted to ensure validity of results. To avoid potential biases, special care was taken to rule out potential confounders such as socioeconomic status and children´s proficiency in the language of instruction. We found our results to be robust. The sampling procedure by an a priori direct matching was efficient and useful. The effect on reading found in the current study is comparable to that found in the RANCH study.

The study also showed a strong relationship between noise exposure and the factor *impairments of school lessons through aircraft noise.* Efficient utilization of the lessons time is an important criterion for instructional quality. The teachers´ reports indicate that, under conditions of aircraft noise, part of the lessons time is lost. Noise-induced loss of instructional quality might contribute to the adverse effects of aircraft noise on reading found in NORAH and other studies.

The current study provides further evidence for negative effects of aircraft noise on children and also on quality of instruction. The findings are of relevance for policy of environmental noise and child development.

Acknowledgments

This study is part of the NORAH research project. NORAH is commissioned by the Environment & Community Center / Forum Airport & Region, Kelsterbach, Germany.

References

[1] Hygge, S., Evans, G. W., & Bullinger, M. (2002). A prospective study of some effects of aircraft noise on cognitive performance in school children. *Psychological Science*, *13*(5), 469–474. doi:10.1111/1467-9280.00483

[2] Clark, C., Martin, R., van Kempen, E., Alfred, T., Head, J., Davies, H. W.,…Stansfeld, S. A. (2006). Exposure-effect relations between aircraft and road traffic noise exposure at school and reading comprehension: The RANCH Project. *American Journal of Epidemiology*, *163*(1), 27–37. doi:10.1093/aje/kwj001

[3] Clark, C., & Sörqvist, P. (2012). A 3 year update on the influence of noise on performance and behavior. *Noise and Health*, *14*(61), 292–296. doi:10.4103/1463-1741.104896

[4] Evans, G. W., & Hygge, S. (2007). Noise and cognitive performance in children and adults. In L. M. Luxon & D. Prasher (Eds.), *Noise and its effects* (pp. 549–566). Chichester, England, Hoboken, NJ: Wiley.Green, K. B., Pasternack, B. S., & Shore, R. E. (1982).

[5] Green, K. B., Pasternack, B. S., & Shore, R. E. (1982). Effects of aircraft noise on reading ability of school-age children. *Archives of Environmental Health: An International Journal*, *37*, 141-145. doi:10.1080/00039896.1982.10667528

[6] Haines, M. M., Stansfeld, S. A., Head, J., & Job, R. (2002). Multilevel modelling of aircraft noise on performance tests in schools around Heathrow Airport London. *Journal of Epidemiology & Community Health*, *56*, 139-144. doi:10.1136/jech.56.2.139

[7] Stansfeld, S. A., Berglund, B., Clark, C., Lopez-Barrio, I., Fischer, P., Öhrström, E., …Berry, B. F. (2005). Aircraft and road traffic noise and children's cognition and health: A cross-national study. *The Lancet*, *365*(9475), 1942–1949. doi:10.1016/S0140-6736(05)66660-3

[8] Clark, C., Lopez Barrio, I., van Kamp, I., van Kempen, E., & Stansfeld, S. A. (2014, June). *Teachers' reactions to environmental noise at school in the RANCH project: A potential mechanism for noise effects on children's cognition?* Proceedings of 11th International Conference on Noise as a Public Health Problem, ICBEN, Nara, Japan.

[9] Klatte, M., Spilski, J., Mayerl, J., Möhler, U., Lachmann, T. & Bergström, K. (2016). Effects of Aircraft Noise on Reading and Quality of Life in Primary School Children in Germany: Results from the NORAH Study. *Environment and Behavior,* first published on April 13 as doi:10.1177/0013916516642580.

[10] Mohler, U., Liepert, M., Muhlbacher, M., Beronius, A., Nunberger, M., Braunstein, G., … Bartel, R. (2014). *Verkehrslärmwirkungen im Flughafenumfeld: Erfassung der Verkehrsgeräuschexposition* [Effects of traffic noise in areas surrounding airports: Report of methods of noise exposure assessment]. Retrieved from http://www.laermstudie.de/ergebnisse/basismodul-akustik/

[11] Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys* (Vol. 81). John Wiley & Sons.

[12] Bernaards, C. A., & Sijtsma, K. (2000). Influence of simple imputation and EM methods on factor analysis when item nonresponse in questionnaire data is nonignorable. Multivariate Behavioral Research, 35(3), 321-364.

[13] Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, (18), 39–50.

[14] Muthen, L., & Muthen, O. (2014). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

[15] Hox, J. J. (2010). *Quantitative Methodology Series: Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.

[16] Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine, 28*(25), 3083–3107. doi:10.1002/sim.3697