Original Article

Annoyance and task performance during a single high-level aircraft noise and multiple lower-level aircraft noises

David Welch[1], Daniel Shepherd[2], Kim N. Dirks[3], Jessica Ong[1,4]

[1]Audiology Section, School of Population Health, University of Auckland, Private Bag 92019, Auckland 2100, New Zealand, [2]Department of Psychology, Auckland University of Technology, Private Bag 92006, Auckland, 1142, New Zealand, [3]Department of Civil and Environmental Engineering, Faculty of Engineering, University of Auckland, Private Bag 92019, Auckland 2100, New Zealand, [4]Cochlear Limited, 238A Thomson Road #25-06, Novena Square Office Tower A, Singapore 307684, Singapore

Abstract

Objective: Measurement of noise exposure tends to focus on sound level; however, the number of noise events might also influence health and cognition. We investigated how the distribution of noise events over time influences experiences of loudness, annoyance, and task performance. **Materials and Methods:** We presented recordings of a passenger aircraft flying overhead, either as a single 15-s overflight at 80 dB $\text{LeqA}_{15\text{s}}$ or four 15-s overflights at 60 dB $\text{LeqA}_{15\text{s}}$. Levels were chosen on the basis that an increase of 10 dB doubles the perceived loudness, thus four stimuli at 60 dB might be expected to seem as loud as a single stimulus at 80 dB. Participants performed a mental arithmetic task during half the stimulus presentations and rated their perception of loudness and annoyance to every presentation, while pulse and skin conductance were monitored. **Results:** Overall, the single 80-dB-flight stimulus was perceived as louder ($F = 124.519$, $P < 0.001$) and more annoying ($F = 63.530$, $P < 0.001$) than the four 60-dB flights. Noise did not influence task performance; however, there was an interaction ($F = 36.256$, $P < 0.001$) in that while doing the task, the four 60-dB flights were perceived as louder and more annoying than without the task, whereas the single 80-dB stimulus was less loud and no more annoying than without the task. Physiological markers were consistent with the intent that the task be difficult and that the single high-sound-level stimulus was more stressful for participants. **Conclusion:** Results showed that, at the levels used, the higher-level stimulus influenced ratings of loudness and annoyance, even though the lower-level stimulus occurred four times as often. Future research exploring systematically the relationship between the number and sound level of overflights and the reactions they induce is needed.

Keywords: Noise, Loudness perception, Annoyance, Psychophysiology

KEY MESSAGES

(1) Annoyance and Loudness were greater for a single aeroplane overflight at 80 dBA than for four overflights at 60 dBA.
(2) Doing an engaging visual task causes the single overflight to be less loud and the four overflights to be louder than when not doing the task.
(3) Sound level and the number of sound events interact with task engagement to influence perceived loudness and annoyance.

INTRODUCTION

Environmental noise causes distraction, interference with tasks, and annoyance, resulting in both direct and cumulative health effects.[1] Annoyance due to noise is an important mediator in the relationship between noise, stress and health;[2-4] people who are annoyed by noise experience negative emotions, including fear and anger, accompanied by physiological arousal, leading to negative health effects.[5,6]

Two principal pathways have been proposed.[7] The first is the nonconscious pathway mediated by the limbic system.

Access this article online

Quick Response Code:

Website:
www.noiseandhealth.org

DOI:
10.4103/nah.nah_207_25

Address for correspondence: David Welch, Audiology Section,
School of Population Health, University of Auckland,
Private Bag 92019, Auckland 2100, New Zealand.
E-mail: d.welch@auckland.ac.nz

Received: 13 October 2025 **Revised:** 17 March 2026
Accepted: 20 March 2026 **Published:** 30 April 2026

For reprints contact: reprints@medknow.com

How to cite this article: Welch D, Shepherd D, Dirks KN, Ong J. Annoyance and task performance during a single high-level aircraft noise and multiple lower-level aircraft noises. Noise Health 2026;28:493-501.

The second is the conscious pathway that triggers emotional responses and subsequent cortical activation.[8]

Sounds Varying over Time

Despite significant reductions in the sound levels, aircraft noise is rated as a highly annoying environmental noise source.[9] People report being more annoyed by a given sound level of aircraft noise exposure now than in the past, which may reflect sensitization to noise.[10-12]

A previous experiment was conducted to improve understanding of how humans perceive and rate sounds that vary over time.[13] The sounds of varying numbers of aircraft overflights were presented at varying sound levels, and participants rated loudness. Some ratings were based on the loudest part of the exposure, while others integrated the overall loudness of the experience.[13,14]

Task Performance

Noise exposure may interfere with complex task performance.[15,16] This includes reduced attention, problem-solving ability and memory, as well as slower rehearsal in memory, altered selectivity in memory and slower rates of choice of strategy selection for performing a task.[17,18] A possible explanation is that noise captures attention and interrupts cognitive processes. To measure this effect, a task with a measurable outcome is needed.[19-21] The paced auditory serial addition task (PASAT) requires participants to perform mental arithmetic within a tight timeframe. It assesses concentration ability, capacity, and rate of information processing, as well as sustained and divided attention.[22] During the PASAT, cerebral blood flow has been shown to increase; however, the increase was less when the PASAT was carried out in the presence of traffic noise.[23] Interestingly, exposure to noise did not impact upon PASAT performance indices, which the authors suggested might be a learning effect. It was also suggested that future research into noise effects should measure cognition using a technique based on the visual modality rather than an auditory-modality-based assessment such as the PASAT.[23]

This Research

The present research partially replicated earlier work[14] while also incorporating a visual parallel to the PASAT: the paced visual serial arithmetic task (PVSAT). Like the PASAT, the PVSAT was designed to be highly engaging and cognitively demanding. Participants were asked to rate the loudness and annoyance of the aircraft recordings presented to them while performing the PVSAT and while not. Changes in skin conductance, heart rate, and blood flow were measured as physiological correlates of the task and exposures.

Objectives of the research were as follows:

(1) To investigate the influence of the number and level of sounds on the perception of loudness and annoyance of aircraft overflights.
(2) To investigate the influence of the number and level of sounds on task performance.
(3) To help understand the mechanism underlying Objectives 1 and 2, blood volume pulse and skin conductance were recorded as physiological markers.

The research was thus designed as a first step in the exploration of the relative influences of sound level and the number of sound events on perception of annoyance and loudness, and task performance.

METHODS

Participants

Initially, a sample of 30 participants was recruited between August and December 2017. Of these, one was excluded due to incomplete data. Recruitment was through posters, social media and electronic flyers, and most of the participants were university students or staff. Informed written consent was obtained. Participants were instructed to avoid the consumption of stimulants such as caffeine (4 h), or alcohol (8 h), or any sympathomimetic and/or anticholinergic drugs prior to testing. The sample size was based on that of previous research that had shown significant effects in a similar set of tasks.[14] The research was approved by the University of Auckland Human Participants Ethics Committee (Ref: 18771).

Procedure

Paced Visual Serial Addition Task

The PVSAT is a visual analogue of the Paced Auditory Serial Addition Task and has been used in previous auditory-related experiments.[24] The visual number presentation was programmed and administered to the participant through National Instruments™ LabVIEW (v.2014). It was presented through a 20 BackLite LCD Monitor (Dell Inc., USA) and adjusted so that it was at eye level and approximately an arm’s length away from the participant. The PVSAT consisted of a random series of numbers ranging from one to nine presented visually one at a time. Participants were instructed to add consecutive pairs of numbers and to respond as quickly and accurately as possible.

Each number image was 25.4 mm high and wide, with black numbers on a grey background. The answer buttons were presented on a grid that ranged from 1 to 20, each 5 mm high and 10 mm wide. Participants were required to use their dominant hand to select the correct answer using a mouse.

Initially, an adaptive tracking procedure[25] was used to individualize the response time for each participant to perform the mental addition. A three-down, one-up procedure with ten turnarounds was used to track the 79% “threshold” (interdigit interval). In this task, for every three consecutive correct responses, the time interval between digit presentations was reduced by 10 ms, while for every incorrect response, the time interval between digit presentations was increased by 10 ms.

The 79% interdigit interval, determined by the preliminary tracking task described above, was used as a constant response time window in the PVSAT during the experimental sessions, ensuring a consistent level of difficulty for each participant and over time. Participant exposure to the PVSAT task prior to data collection also reduced vulnerability of the experiment to learning effects. As a result of the variability in time limits between participants, the number of PVSAT trials carried out during each 2-min session ranged from 24 to 55 ($M = 38$, SD = 7). Participants were told to expect aircraft overflight noises during the PVSAT and were encouraged to continue to perform as accurately as they could. The scores achieved in the PVSAT were the proportion of correct responses in each session, with a missed response being scored as incorrect.

Noise Presentations

Participants were seated in a standard sound-attenuating chamber.[26] The dimensions of the booth were 2.2 m × 2.5 m. Ambient noise levels were compliant with the standard for maximum permissible ambient noise levels in an audiometric test room.[27] RP-HT160 Stereo Headphones (Panasonic Corporation, Japan), calibrated to the presented sounds, were used.

A large passenger jet aeroplane in a steady overflight was recorded with a sampling rate of 44,100 Hz, stereo channel selection and a 16-bit depth on Audition® CC 2017 (Adobe Inc., USA). Noise reduction techniques were applied, and wind noise was removed from the recording, which was trimmed to 15 s. Onset and offset were smoothed using a spline curve. The frequency spectrum and relative levels were not altered. Peak clipping was not apparent. Spectral analysis of the noise showed a dominance of middle frequency components (750–1500 Hz). Digital copies of this recording were used for all presentations of the noise.

Each condition lasted 2 min, and the overflight sounds were presented within that period, either as a single 15-s overflight at a sound level of 80 dB L_{eq} A $_{15\,s}$ or four 15-s overflights at a sound level of 60 dB L_{eq} A $_{15\,s}$ each.

The single 80 dB L_{eq} A overflight condition consisted of: 30 s of silence, 15 s of aircraft overflight noise, followed by 75s of silence.

The four 60 dB L_{eq} A overflights condition consisted of: 15 s of aircraft overflight noise, 20 s of silence, 15 s of aircraft overflight noise, 15 s of silence, 15 s of aircraft overflight noise, 25 s of silence, and finally 15 s of aircraft overflight noise. The silent intervals were varied to imitate the unpredictability of consecutive aircraft overflights, separated by a noise-free period.

The third condition was silent, with no overflights.

Each 2-min condition was presented twice, with participants either occupied doing the PVSAT or inactive. The order was pseudo-randomized across participants such that the conditions with the PVSAT task alternated with the inactive conditions, to reduce boredom. Each condition was assigned a number from one to six for randomizing purposes. The even-numbered conditions included the PVSAT, while odd numbered conditions did not. Conditions one and two had no aeroplane noise, three and four had the four 60 dB L_{eq} A overflights, and five and six had one 80 dB L_{eq} A overflight. Pseudorandomized sequences of these numbered conditions were made, where each had alternating odd and even-numbered conditions, and thus the PVSAT occurred in alternating conditions. The presentation order was generated through an online randomizer (Research Randomizer) for each participant. There were, therefore, a total of six conditions, and each was repeated three times for a total of 18 periods. For presentation, the means across repeats of each condition were used. A 2-min break was taken between conditions.

The sound levels and numbers of stimuli were chosen on the basis that the unit, one sone, is the modulus of a relative scale used to denote loudness, being the loudness experienced in response to a 40 dB SPL tone presented at 1 kHz. On average, a 1-kHz sound at 60 dB has a loudness of 4 sones, and at 80 dB, of 16 sones. Other frequencies follow the equal-loudness contours that are driven by the receptivity of the human auditory system. The sound levels in the present study were set so that the overall loudness of the four 60-dB overflights would match the loudness of the single 80-dB overflight if participants integrated the perceptions in a linear manner.

Rating of Loudness and Annoyance

Immediately after each overflight stimulus presentation condition, participants provided responses on the perceived loudness and annoyance scales. We used different rating scales (a 9-point scale for loudness and an 11-point scale for annoyance), in part to discourage participants from indicating the same number for both scales without due consideration. The 9-point loudness scale ranged from (1) Soft to (9) Loud.[14] The 11-point numerical scale, developed by the International Commission on the Biological Effect of Noise (ICBEN), was used to assess annoyance; it ranged from (0) Not at all to (10) Extremely annoyed.[28]

Psychophysiological Measures

Participants sat with the nondominant hand placed palm down on the table, and the elbow comfortably rested. Blood volume, pulse, and skin conductance were simultaneously recorded via a NeXus-10 MKII (Mind Media, The Netherlands) integrated biofeedback and neurofeedback multimodal system. Real-time central and peripheral responses were recorded using BioTrace+ version 2014 Software for NeXus-10 (Mind Media, The Netherlands).

The blood volume pulse amplitude (mV) and heart rate (beats per minute: bpm) of each participant were derived by photoplethysmography. The blood volume pulse sensor was clipped on the second digit of the nondominant hand. Participants were instructed to rest the hand in a static position to reduce artifacts during recording.

Finger perspiration was monitored by a skin conductance sensor (Mind Media, The Netherlands) with a resolution of 0.001 microSiemens (μS). Two Type 1700-050Cleartrace^{TM} Adult ECG Ag/AgCl Electrodes (Conmed, USA) were positioned onto the palmar surfaces of the distal phalanx on the third and fourth digit of the nondominant hand.

Data Treatment and Analysis

The same group of participants underwent all of the conditions and measurements. As such, repeated-measures analyses were conducted. Mauchley's test was used to assess the sphericity assumption of the data, and was not significant, implying that the distribution was satisfactory for the analysis approach used. The loudness and annoyance ratings of the sounds were compared using separate two-way ANOVAs where the stimulus type (four 60-dB or one 80-dB overflight) and the PVSAT task (present or absent) were the factors. The psychophysiological measures (skin conductance, blood volume, pulse amplitude, and heart rate) were measured during the same conditions as loudness and annoyance and were also measured during conditions where there was no overflight stimulus presented. Separate three (stimuli) by two (PVSAT task) repeated-measures ANOVAs were run for each measure. Preliminary models were generated that included age and sex; however, neither the main effects of these variables nor interactions involving them were observed. Their inclusion did not materially alter the main findings either, so analyses without these variables were presented. An exploration of the simple effects of the interactions between task and stimulus type was conducted using paired *t*-tests. SPSS Version 30 (IBM Corporation, USA) was used for analyses.

RESULTS

Participant Characteristics

Ultimately, 29 (11 male; 18 female) adults aged 21–67 years ($M = 29.4$, SD = 12.2) participated in the study. The hearing thresholds of all participants were no poorer than 25dB HL (250–8000 Hz) bilaterally.[29] Participants reported no history of cardiovascular or neurological issues.

PVSAT Task

The mean proportion correct of answers on the PVSAT during the no-overflight condition (proportion correct = 0.892), one 80 dB-overflight (proportion correct = 0.894) and four 60-dB overflights (proportion correct = 0.890) did not differ ($F = 0.127$, $P = 0.881$). The overall mean proportion correct in the experimental sessions was 0.89, which was greater than the mean proportion correct of 0.79 tracked during the preliminary phase.

Loudness

Overall, participants rated the 2-min interval with the single 80 dB overflight as louder than the interval with four 60-dB overflights ($F = 124.519$, $P < 0.001$; Figure 1). The rating of the loudness depended on an interaction between the number/sound-level of the overflights and whether or not the participants were performing the PVSAT task ($F = 36.256$, $P < 0.001$). As can be seen in Figure 1, loudness ratings were generally higher when performing the PVSAT task for the four 60-dB overflights ($M = 4.230$, SD = 1.222) than when not performing the task ($M = 3.839$, SD = 1.272). On the other hand, loudness ratings for the 80-dB overflight tended to be higher when participants were not performing the task ($M = 6.735$, SD = 1.505) than when performing the task ($M = 6.046$, SD = 1.704).

Differences between the task and no-task conditions were significant with the four 60-dB flights stimulus ($t = -2.745$, $P = 0.010$) and the single 80-dB flight stimulus ($t = 3.823$, $P = 0.001$). The flights seemed louder when doing the PVSAT task under the four 60-dB flights stimulus, but louder when doing no task under the single 80-dB flight stimulus [Figure 1].

Annoyance

Results for annoyance largely paralleled those for loudness, and the single 80 dB overflight was more annoying overall than the four 60-dB overflights ($F = 63.530$, $P < 0.001$; Figure 2). As with loudness, annoyance depended on an interaction between stimulus type and task ($F = 21.032$, $P < 0.001$). However, in this case, the pattern of the interaction effect was slightly different in that the influence of the task appeared to be less for the single 80-dB overflight [Figure 2].

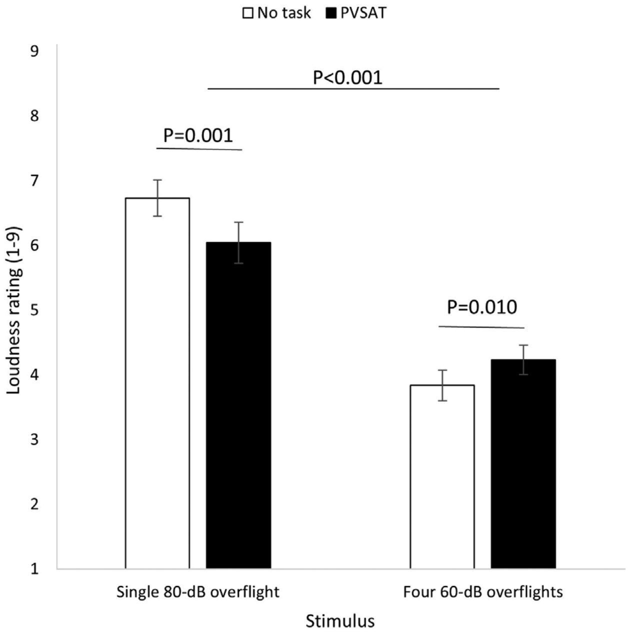

Figure 1: Loudness ratings in response to the two overflight stimuli with and without the PVSAT task. Error bars represent one standard error of the mean.

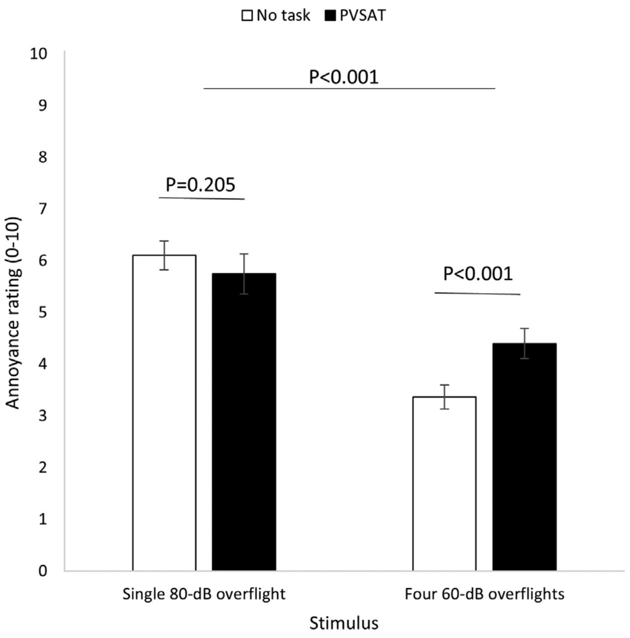

Figure 2: Annoyance ratings in response to the two overflight stimuli with and without the PVSAT task. Error bars represent one standard error of the mean.

Annoyance ratings were generally higher when performing the PVSAT task for the four 60-dB overflights ($M = 4.402$, SD = 1.562) than when not performing the task ($M = 3.369$, SD = 1.818). On the other hand, annoyance ratings for the 80-dB overflight were similar when participants were not performing the task ($M = 6.103$, SD = 2.017) than when performing the task (M = 5.747, SD = 2.092).

Differences between the task and no-task conditions were significant with the four 60-dB flights stimulus ($t = -4.145$, $P < 0.001$) but not with the single 80-dB flight stimulus ($t = 1.297$, $P = 0.205$). This implies that the flights were more annoying when doing the PVSAT task under the four 60-dB flights stimulus, but that the annoyance induced by the single 80-dB flight stimulus was not influenced by doing the task [Figure 2].

Psychophysiological Measures

Heart Rate

For heart rate, there was no effect of task or interaction between task and stimulus condition, but there was a marginal effect of the number/sound-level of overflights, wherein the mean heart rate appeared to be slightly lower in conditions with four 60-dB overflights ($M = 72.881$, SD = 9.242) than in the quiet ($M = 73.486$, SD = 9.237) or in the single 80-dB overflight ($M = 73.285$, SD = 9.385) conditions ($F = 2.950$, $P = 0.061$; Figure 3).

Blood Volume Pulse Amplitude

There was a significant main effect of task whereby blood volume pulse amplitude was generally lower when performing the task ($F = 11.130$, $P = 0.002$). However, there was also a significant interaction between the stimulus type and the presence of the task ($F = 4.598$, $P = 0.014$). This appears to be driven by the mean blood volume pulse amplitude when participants were not performing the task; the lines appear approximately parallel during the overflight conditions, but approach each other more closely in the no-overflight conditions [Figure 4]. The means and standard deviations were: without flights, with no task $M = 25.136$, SD = 13.640, and with the task $M = 22.940$, SD = 12.977.

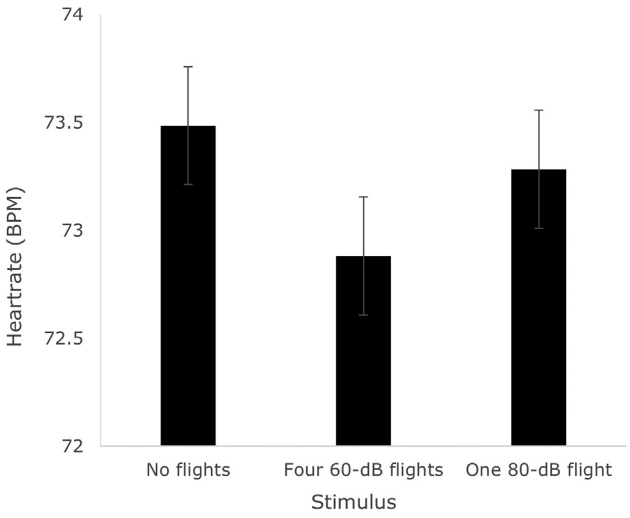

Figure 3: Mean heart rate in beats per minute (BPM) for each of the three types of stimulus conditions. The error bars are one standard error of the mean difference between adjacent conditions.

Skin Conductance

Skin conductance was influenced by the number/sound-level of flights ($F = 4.830$, $P = 0.012$; Figure 5). It was also

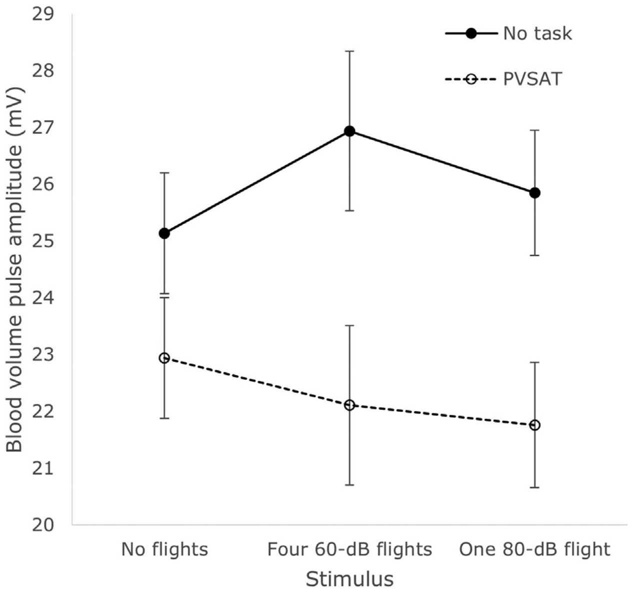

Figure 4: Blood volume pulse amplitude (mV) during each stimulus condition and when performing the task or not. The error bars represent one standard error of the difference between the task and no-task conditions under each stimulus condition.

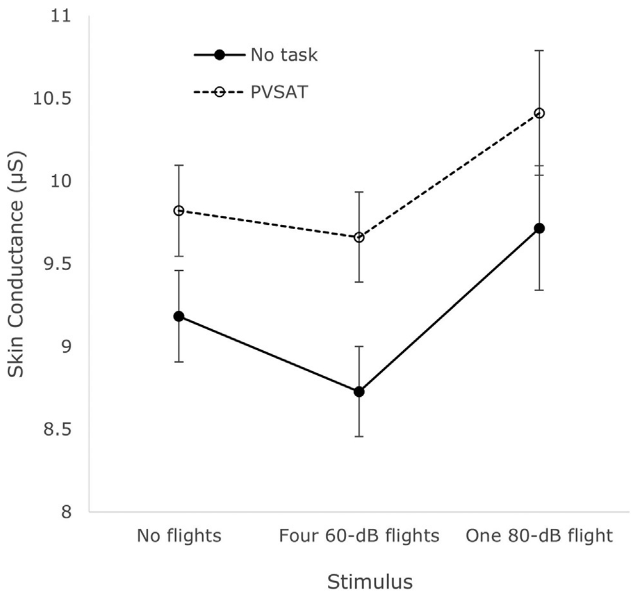

Figure 5: Skin conductance (μS) during each stimulus condition and when performing the task or not. The error bars represent one standard error of the difference between the task and no-task conditions under each stimulus condition.

influenced by the presence of the PVSAT task ($F = 16.105$, $P < 0.001$; Figure 5). Participants had higher skin conductance in sessions with the single 80-dB overflight and when performing the PVSAT task.

DISCUSSION

The single, higher-sound-level overflight was perceived as louder and more annoying than the four lower-sound-level overflights. This was the case whether participants were performing the PVSAT task or not. Doing the PVSAT task led to the perception of the four lower-sound-level overflights being louder and more annoying than when not performing the task. Similarly, the single higher-sound-level overflight was perceived as quieter (though not significantly less annoying) when doing the task. In other words, the difference between the single higher-sound-level and multiple lower-sound-level overflights was reduced while performing the task. Blood volume pulse amplitude was lower when performing the PVSAT task than when not. Skin conductance was greater when doing the PVSAT task, no matter what noise condition, and was higher in the conditions with the single, higher sound-level overflight.

Time-Varying Noise Stimuli

Part of our intent in conducting this research was to confirm and extend earlier research by Genuit and Fiebig[14] showing that different groups of participants adopted different approaches to estimating the loudness of aircraft overflights that varied over time. In almost all cases, our participants found the single, higher-sound-level overflight to be louder and more annoying than the four lower-sound-level overflights. One difference between the earlier findings and the current study was the timeframe of the sound presentations: the previous research presented the overflights within a 6-min timeframe, and each overflight lasted 1 min, whereas the current study used 15-s presentations. This may have altered the experience for participants, making the salience of the single, higher-sound-level overflight stimulus greater in the current research. Another possible explanation for the difference in findings is the relative sound level of the different overflight stimuli used in the two studies; our sound levels were sufficiently distinct that the single, higher-sound-level overflight was almost always experienced as the louder and more annoying.

We had designed the experiment on the basis that a 10-dB change in stimulus level tends to double the rated loudness (in sones). Under that assumption, the loudness experienced from four 60-dB overflights would be the same as the single 80-dB overflight, since the 20-dB difference would cause two doublings in loudness. Since almost all participants rated the single 80-dB overflight as louder (and more annoying); however, the implication is that our participants did not integrate their perception of loudness across events, but rather focused on peak loudness.

The perceived loudness and annoyance associated with single versus many sound events must be based on either the peak levels or an integration of energy over time.[14] Jeffress[30] argued that the primary auditory stimulus is the displacement of the basilar membrane, with the amplitude of displacement transformed into a neural count. If the peak mechanism is strictly true, then had the four overflights been presented at the same sound level as the single overflight, they would have been rated as having the same loudness/annoyance as the single overflight. This seems counterintuitive.

The second possibility, that the overall sound exposure is integrated, predicts that the four overflights would be rated louder and more annoying than the single overflight if all individual flights were presented at the same sound level. If that is the case, then there must be a point at which the loudness/annoyance functions intersect: not at the equivalent sones as we had predicted, but at some intermediary sound level. This suggests research in which varying the sound-level of the multiple flights systematically would allow functions to be traced that would describe the perceived relative loudness and annoyance.

PVSAT Task

We set the time interval for each participant by tracking the 0.79 proportion correct with a 3-down–1-up adaptive procedure in the preliminary phase. This provided a time interval tailored to each participant's ability at mental arithmetic, but the overall proportion correct during the main experiment was 0.89. This may imply that participants continued to learn, or that when the response interval became stable in the main experiment, participants found the task easier. We do not believe this discrepancy impacted the findings because the PVSAT still provided an engaging task for the participants during exposure to noise.

The proportion correct on the PVSAT did not vary between the silent condition and the two different overflight conditions. This implies that the differences in perceived loudness and annoyance do not influence the ability to perform an engaging cognitive task. The finding is in line with previous research that used an auditory version of the task.[23] Since the task was presented via the visual modality in the current study, this rules out possible explanations due to auditory masking effects.

Psychophysiological Measures

Heart rate did not vary significantly, but was marginally lower during the four 60-dB overflights than in the other conditions. We assume this was simply natural variation, as strict adherence to the statistical significance levels would imply.

Blood volume pulse amplitude was lower when conducting the PVSAT task, a finding that is in keeping with the intent that the task would be difficult and stressful for participants. There was significant variation according to the overflight stimulus, but the interaction appeared to be driven mostly by a low blood volume pulse amplitude during the silent condition when not performing the task. Again, this may represent random variability.

Skin conductance was higher when performing the PVSAT task, which again supports the idea that the task was difficult for participants. Whether doing the task or not, skin conductance was greater during conditions with the single 80-dB overflight than in either the silent condition or the condition with the four 60-dB overflights. This finding aligns well with the perceived loudness and annoyance of the overflight conditions. It is interesting that the skin conductance during the four 60-dB overflight conditions was no greater than in the silent conditions, given that the sound was rated as being loud and annoying, even though it was less than in the single 80-dB overflight conditions.

Influence of a Task on Perceptions of Loudness and Annoyance

The presence, sound level, and number of overflights did not influence the PVSAT score. However, performing the PVSAT influenced the rated loudness and annoyance of the overflight sounds: during the task, the multiple, lower-sound-level sounds were perceived as louder and more annoying, while the single, higher-sound-level overflight was perceived as less loud than when not performing the task. Combined with the finding that the single high-sound-level stimulus was consistently rated as louder and more annoying than the stimulus with multiple lower-sound-level overflights, there is a possible explanation. Since lower-sound-level flights were generally less annoying than higher-sound-level flights, it may be that each flight was distracting from the task, and so each flight provided an opportunity to get distracted. Performance did not deteriorate during the overflight situations, but perhaps participants were only able to overcome the distraction with effort. It is possible that this effort was annoying, hence the increased ratings of loudness and annoyance during the task for the multiple-flight stimulus.

But why did the perceived loudness of the single, higher-sound-level overflight decrease when doing the task? One difference between the stimuli was that the final overflight of the four lower-sound-level overflights stimulus finished right at the end of the experimental session and so immediately prior to the annoyance and loudness ratings, whereas the single overflight finished 75 s before the end of the session. This longer time interval may have reduced the salience for participants when performing the engaging task more than when they were inactive. This may have allowed them to put the perception of the overflight out of their minds.

Recommendations and Limitations

The findings suggest that, for this particular comparison, the maximum sound level was more salient than the overall sound exposure, implying that the integration of multiple events is less important to listeners than the peak sound level. We have reported on laboratory presentations of recorded overflights presented against a background of silence, so the experience of people in the real world may differ. While we pseudorandomized the order of presentation of the conditions, and conditions with and without the PVSAT task were alternated, we did not record the specific order of conditions for each participant, meaning that it was not possible to investigate order effects. Furthermore, the participants in this research tended to be highly educated and varied in age, which may have influenced performance on the PVSAT task and potentially contributed to the findings of annoyance. Results should therefore be treated with caution, given the small sample size and the lack of controls outlined above.This research suggests several directions for future work: first, since the sound levels differed by such a large amount (20 dB), the finding that the single higher-level condition was louder may seem unsurprising. Future research should explore the functions relating sound level and number of events to perceived loudness/annoyance, and investigate where the functions cross. Another area that would be interesting to explore is the role of noise sensitivity in these associations; potentially, noise sensitivity may be associated with a different pattern of responses. We had originally intended to include a measure of noise sensitivity in this research; however, in preliminary research, we discovered that noise sensitivity is a more complicated phenomenon than we had supposed.[31] More research into noise sensitivity is ongoing[32], and this may ultimately provide a suitable measure.

CONCLUSION

In summary, a single overflight at 80 dB L_{eq} $A_{15\,s}$ was perceived as louder and more annoying than were four overflights each presented at 60 dB L_{eq} A $_{15\,s}$. Future research may explore and map the relative roles of sound level and the number of events to describe the function relating these two contributors to noise annoyance. The loudness and annoyance of the two stimuli were more similar when participants were engaged in a task. That the multiple-flights stimulus was louder and more annoying when trying to concentrate on a task is intuitive and may explain this. Physiological findings support the idea that the PVSAT task was highly engaging and that the single higher-sound-level stimulus caused greater stress in participants. Overall, this research supports the idea that the peak sound level is more salient to people than the number of noise events under the specific conditions used here; however, research should explore this further to enable more general comments to be made.

Availability of Data and Materials

Data are not available due to the requirements of the ethical approval.

Author Contributions

David Welch developed the concept for the research, analyzed the data, and wrote the first draft of this paper. Daniel Shepherd wrote the software for the PVSAT task and contributed to the interpretation of the findings and the writing of the paper. Kim Dirks contributed to the interpretation of the findings and the writing of the paper. Jessica Ong collected the data, contributed to the design of the research and data analysis, and was involved in writing and interpretation of the findings.

Ethics Approval and Consent to Participate

The research was approved by the University of Auckland Human Participants Ethics Committee (Ref: 18771). Participants gave informed consent. This study was conducted following the ethical principles of the World Medical Association Declaration of Helsinki.

Acknowledgment

Thanks to the participants for their time and contribution to this study.

Financial Support and Sponsorship

No funding was received for this research.

Conflicts of Interest

This manuscript is based on work that was previously presented as an article at the NOISE-CON Congress (reference DOI: https://doi.org/10.3397/IN_2024_3494).

References

1. Basner M, Babisch W, Davis A, *et al.* Auditory and non-auditory effects of noise on health. Lancet 2014; 383: 1325–32.
2. Shepherd D, Welch D, Dirks KN, Mathews R. Exploring the relationship between noise sensitivity, annoyance and health-related quality of life in a sample of adults exposed to environmental noise. Int J Environ Res Public Health 2010; 7: 3579–94.
3. World Health Organisation. Burden of disease from environmental noise: Quantification of healthy life years lost in Europe. Copenhagen; 2011.
4. Park SH, Lee PJ. Effects of noise on health-related quality of life: the roles of outdoor noise, indoor noise, and noise sensitivity. J Expo Sci Environ Epidemiol 2026; 36: 398–408.
5. Vastfjall D. Influences of current mood and noise sensitivity on judgments of noise annoyance. J Psychol 2002; 136: 357–70.
6. Münzel T, Kuntic M, Daiber A, Sorensen M. Transportation noise and the cardiometabolic risk. Atherosclerosis 2025; 403: 119148.
7. Babisch W. The noise/stress concept, risk assessment and research needs. Noise Health 2002; 4: 1–11.
8. Babisch W, Pershagen G, Selander J, *et al.* Noise annoyance – a modifier of the association between noise level and cardiovascular health? Sci Total Environ 2013; 452-3:50–7.
9. Babisch W, Houthuijs D, Pershagen G, *et al.* Annoyance due to aircraft noise has increased over the years-Results of the HYENA study. Environ Int 2009; 35: 1169–76.
10. Gelderblom FB, Gjestland T, Fidell S, Berry B. On the stability of community tolerance for aircraft noise. Acta Acust United Acust 2017; 103: 17–27.
11. Guski R. How to forecast community annoyance in planning noisy facilities. Noise Health 2004; 6: 59–64.
12. Janssen SA, Vos H, van Kempen EEMM, Breugelmans ORP, Miedema HME. Trends in aircraft noise annoyance: the role of study and sample characteristics. J Acoust Soc Am 2011; 129: 1953–62.
13. Genuit K. The need for transdisciplinary actions – psychoacoustics, sound quality, soundscape and environmental noise. Internoise; Innsbruck 2013.
14. Fiebig A, Human strategies to provide overall assessments of unsteady sound episodes. Internoise; 2016; Hamburg.
15. Cohen S, Evans GW, Krantz DS, Stokols D., Physiological, motivational, and cognitive effects of aircraft noise on children: Moving from the laboratory to the field. Am Psychol 1980; 35: 231–43.
16. Lee PJ, Jeon JY. Relating traffic, construction, and ventilation noise to cognitive performances and subjective perceptions. J Acoust Soc Am 2013; 134: 2765–72.
17. Smith AP, Broadbent DE. Non-auditory effects of noise at work: a review of the literature. United Kingdom; 1991. Report No.: HSE Contract research report 30/1991. Contract No.: 30/1991.
18. Wang L, Ma J, Li H, Yam MCH, Zhao ZY, Yu CY. Environmental sounds on office worker performance: Non-invasive monitoring of physiological responses in laboratory settings. Build Environ 2026; 290: 114137.
19. Hughes RW, Hurlstone MJ, Marsh JE, Jones DM, Vachon F. Cognitive control of auditory distraction: impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. J Exp Psychol Hum Percept Perform 2013; 39: 539–53.
20. Sörqvist P, Stenfelt S, Rönnberg J. Working memory capacity and visual-verbal cognitive load modulate auditory-sensory gating in the brainstem: toward a unified view of attention. J Cogn Neurosci 2012; 24: 2147–54.
21. Sörqvist P, Rönnberg J. Individual differences in distractibility: an update and a model. Psych J 2014; 3: 42–57.
22. Gronwall DMA. Paced auditory serial-addition task – measure of recovery from concussion. Percept Mot Skills 1977; 44: 367–73.
23. Nowakowska-Kotas M, Pokryszko-Dragan A, Brodowski M, Szydlo M, Podemski R. Effects of noise and mental task performance upon changes in cerebral blood flow parameters. Noise Health 2015; 17: 422–8.
24. Royan J, Tombaugh TN, Rees L, Francis M. The adjusting-paced serial addition test (Adjusting-PSAT): thresholds for speed of information processing as a function of stimulus modality and problem complexity. Arch Clin Neuropsychol 2004; 19: 131–43.

25. Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am 1971; 49: 467–77.
26. ISO 8253–1. Acoustics – Audiometric test methods – Part 1: Pure-tone air and bone conduction audiometry. 2010.
27. ISO 8253–2. Acoustics – Audiometric test methods – Part 2: Sound field audiometry with pure-tone and narrow-band test signals. 2009.
28. Fields JM, De Jong RG, Gjestland T, *et al.* Standardized general-purpose noise reaction questions for community noise surveys: Research and a recommendation. J Sound Vib 2001; 242: 641–79.
29. Carhart R, Jerger JF. Preferred method for clinical determination of pure-tone thresholds. J Speech Hear Disord 1959; 24: 330–45.
30. Jeffress LA. Pedestal experiments: amplitude versus power. J Acoust Soc Am 1979; 65: 1034–8.
31. Welch D, Dirks KN, Shepherd D, Ong J. What is noise sensitivity? Noise Health 2022; 24: 158–65.
32. Wu JA, Dirks KN, Shepherd D, Welch D. A revised model of noise sensitivity. Environ Res Health 2025; 3: 025004.