Environmental Research and Consultancy Department

Civil Aviation Authority

AND

Acoustics and Vibration Group

Bureau Veritas

Department for Transport

Attitudes to Noise from Aviation Sources in England
Non SP Peer Review

CAA: 4ER/2/1/11
BV: NGGX0072/st/07/76
October 2007

Technical Report	CAA – 4ER/2/1/11 BV – NGGX0072/st/07/76 **Date:** October 2007
Title:	**Attitudes to Noise from Aviation Sources in England** **Non SP Peer Review**
Submitted to:	Department for Transport Great Minster House 76 Marsham Street London SW1P 4DR
Attention:	David Hart
Submitted by:	Environmental Research and Consultancy Department, Civil Aviation Authority, CAA House 45 – 59 Kingsway London WC2B 6TE Tel: 0207 453 6082 Fax: 0207 453 6097 AND Acoustics and Vibration Group Bureau Veritas 30 Great Guildford Street London SE1 0ES Tel: 0207 902 6176 Fax: 0207 902 6118 Mob: 07702 634 278
Prepared by:	P.G.Havelock BSc (Hons), MBA, FRSA Head of Environmental Research & Consultancy, Directorate of Airspace Policy, Civil Aviation Authority AND S.W.Turner, MA, MSc, FIOA Director of Acoustics Bureau Veritas
Approved by:	P.G.Havelock BSc (Hons), MBA, FRSA Head of Environmental Research & Consultancy, Directorate of Airspace Policy, Civil Aviation Authority AND S.W.Turner, MA, MSc, FIOA Director of Acoustics for Bureau Veritas

Document Control Sheet

Client: **Department for Transport**

Document no: **CAA – 4ER/2/1/11**
BV – NGGX0072/st/07/76

Document Status and Approval Schedule

Issue	Status	Description	Prepared by	Approved by
1	Draft	First draft for CAA / BV approval	PH and SWT	PH and SWT
2	Final	Final version July 2007	PH and SWT	PH and SWT
3	Final 30/10/2007	Following feedback from researchers and delivery of penultimate and final versions of the report – October 2007	PH and SWT	PH and SWT
4	Final 31/10/07	Final version following QA review	PH and SWT	PH and SWT

Project Quality Assurance:

This Project has been undertaken in accordance with the Bureau Veritas Quality Management System registered by BSIQA to BS EN ISO 9001: 1994. (Certificate No. FS34143 in the name of the BV Group, of which the Acoustics and Vibration Group forms part of the UK organisation).

This Project has been undertaken in accordance with the Civil Aviation Authority Directorate of Airspace Policy Quality Management System registered to BS EN ISO 9001: 1994. (Certificate No. FS36365).

Contents

1 Introduction

1.1 As part of its research study entitled 'Attitudes to Noise from Aviation Sources in England' (ANASE), the Department for Transport (DfT) formed, in May 2006, what has been known as the Non Stated-Preference Peer Review Group. (A parallel Stated Preference (SP) Peer Review Group had been formed some time earlier). The role of the Non SP peer review group was to review the outcomes of the ANASE work with a particular focus on the findings regarding the responses to aircraft noise expressed primarily in terms of annoyance.

1.2 The Non SP peer review group comprised:

- Peter Havelock, Head of Environmental Research and Consultancy, Directorate of Airspace Policy, Civil Aviation Authority; and
- Stephen Turner, Director of Acoustics, Bureau Veritas.

hereinafter described as the "reviewers".

1.3 Both these reviewers have many years experience of dealing with issues associated with the measurement, prediction and the assessment of the impact of aircraft noise.

1.4 This report has been jointly prepared by the reviewers. Section 2 sets out a chronology of the review process including the interactions with the Department for Transport project managers and the research team. Section 3 considers issues to do with the characterisation of the aircraft noise exposure of the respondents surveyed in the study; and Section 4 comments on the social survey element of the project. Section 5 provides comments on specific issues contained in the final report. Section 6 presents some thoughts about the conclusions to be drawn from the results if taken as presented. Section 7 provides overall conclusions regarding the study.

1.5 The comments expressed in this report are a true independent evaluation by the reviewers of the outcome of the project. They are based on the documentation supplied by the Department for Transport and the research team during the review process.

Background

1.6 Government policy on the impact of aircraft noise has for many years been underpinned by research carried out by the CAA in the early 1980s. This research,[1] normally referred to as the Aircraft Noise Index Study (ANIS), concluded that there was a statistically significant relationship between community annoyance and aircraft noise as measured by the L_{Aeq}[2] metric. Furthermore, ANIS concluded that there was no better metric than L_{Aeq} in terms of correlation between aircraft noise and community annoyance.

1.7 Following consultation[3] the Government decided to adopt the use of L_{Aeq} to describe noise exposure at the three designated airports[4]. Other major UK airports followed Government practice for the most part. Based on the ANIS study the Government decided that 57 dB(A) $L_{Aeq,16h}$ marks the approximate onset of significant community annoyance from aircraft noise.

[1] Civil Aviation Authority, *DR Report 8402 United Kingdom Noise Index Study: Main Report*, January 1985
[2] Equivalent continuous sound level (L_{Aeq}) is defined as the level a notional steady sound which, over a defined period, would contain the same (frequency-weighted) sound energy as the actual variable sound.
[3] JB Critchley and JB Ollerhead, *DORA Report 9023 The use of Leq as an Aircraft Noise Index*, Civil Aviation Authority, September 1990
[4] Currently Heathrow, Gatwick and Stansted are designated under section 80 of the Civil Aviation Act 1982 for the purposes of section 78 (Regulation of noise and vibration from aircraft) of that act.

This level has been used in numerous analyses and most recently for work connected with the 2003 Air Transport White Paper 'The Future of Air Transport' (ATWP)[5].

1.8 On 8 May 2001 the then aviation minister announced a major study into aircraft noise. The minister stated that:

'the new study underlines the Government's commitment to underpin our policy on aircraft noise by substantial research that commands the widest possible confidence'.

The minister went on to say that conclusions from the research carried out in the 1980s (i.e. ANIS) have:

'been broadly confirmed by other studies here and abroad, and we have no reason to doubt their validity'.

However, he stated that:

'in the light of our commitment to develop a new air transport policy, of changes to traffic patterns … and the general reduction in noise levels of individual aircraft, it is now timely to commission a fresh study'.

Attitudes to Noise from Aviation Sources in England (ANASE)

1.9 It is understood that the Department for Transport commissioned a consortium led by MVA Consultancy Ltd, (hereinafter described as the "researchers") to conduct the ANASE project and this commenced in December 2001.

1.10 The ANASE study can be divided into two aspects:

- The monetary valuation of annoyance caused by aircraft noise (described as Stated Preference by the researchers); and
- The relationship between aircraft noise and annoyance.

1.11 It is understood that the study took place between December 2001 and February 2007 and examined reported annoyance from residents at 56 survey sites in the vicinity of 9 airports. The survey sites were located at positions exposed to levels of aircraft noise from 36 dB(A) - 68 dB(A) $L_{Aeq,16h}$. Social survey questionnaires were used to elicit socio-economic data on respondents as well as information on their perception, especially annoyance, with respect to aircraft noise.

[5] Department for Transport, *The Future of Air Transport,* Cm 6046, December 2003

2 The Review Process

2.1 As indicated in the introduction, the Non SP Peer Review Group was formed in May 2006 and the reviewers were supplied with the following documentation:

- ANASE Phase 2 National Survey – Initial Report (final) v 3.1 (February 2006);
- ANASE Phase 2 National Survey – Progress Report (final) (May 2006);
- Technical Note 3 - Phase 2 Sampling Design Proposals, Version 5 (August 2005);
- ANASE Airport Maps Areas for Sampling (July 2005); and
- The Appendices (May 2006).

2.2 During the following month, the reviewers were supplied with

- Attitudes to Aircraft Noise in England (ANASE): Phase 2 National Survey Technical Report (June 2006).

2.3 The reviewers focused on evaluating this document ahead of a meeting with DfT and the research team on 29th June 2006.

2.4 The reviewers identified several issues that related to the characterisation of the noise exposure of the respondents, including:

- The use of only one month's data to characterise a survey question that referred to the previous 12 months;
- The consequential issue of the impact of modal split on the results used;
- The lack of the inclusion of dispersion in any of the modelling;
- The lack of certainty (at the meeting) over the stage length assumptions employed in the model;
- The lack of detailed information about the response at individual Common Noise Areas (CNAs) to enable the reviewers to understand better the results; and
- The reviewers' concern about the details of the survey methodology.

2.5 The reviewers' concern about the lack of information was expressed at a subsequent meeting with DfT on 11th July 2006 and set out in a letter dated 12th August 2006. In particular, the reviewers were seeking a basic set of data showing site location, noise dose and the average annoyance response as a numerical ranking (0 – 10) and the percentage highly annoyed. It was also agreed that the researchers would be asked to carry out additional analysis and checks.

2.6 This additional work was carried out during autumn 2006 and the results presented in a series of Technical Notes:

- TN 15 Version 3 – Comparing ANASE Reported Annoyance with Previous Research, together with appendices;
- TN 17 Version 5 – Additional Analysis of Reported Annoyance;
- TN 18 Version 2 – Additional Material for Reviewers
- TN 19 Version 3 – Effect of 65 LAmax and SEL Sampling and Modelling Cut – off; and
- TN 20 Version 3 – Generic Airport Noise Model.

2.7 These documents were received by the reviewers in October 2006 and a meeting was held between the reviewers, DfT and the researchers on the 20th October 2006. Despite this additional information provided, the reviewers still had concerns about the characterisation of the noise exposure and were not totally clear on the methodology adopted in the derivation of the dose values. It was agreed that the researchers would present a step by step guide to the process used at a subsequent meeting which was held on 30th November 2006. In the meantime the researchers provided responses to other issues raised by the reviewers in a further set of technical notes:

- TN 21 Version 4 – Lav and Log Nav;
- TN 22 Version 4 – Lden Curves; and
- TN 23 Version 4 – Indicators of Annoyance.

2.8 At the November meeting it became clear that:

- Modelling had assumed all departures were stage length 1 rather than taking account of the different loadings of departing aircraft and hence their different flight profiles;
- No dispersion had been assumed about the flight paths, thus not reflecting properly how aircraft tend to be spread across a departure swathe rather than following a single track;
- Arithmetic averaging had been used at one step in the process where logarithmic averaging ought to have been used, thus potentially causing an underestimate of the noise exposure value;
- There was a limited range of movement data used to characterise the average dose at any one location. The movements that occurred during June 2004, and in some cases, July 2004, only had been used. There was no guarantee that these months would represent a longer term average, and, in fact, it was noted that June 2004 was atypical in this respect; and
- Assumptions about the fleet mix were not necessarily robust. As had occurred with the SP work, the range of aircraft types were aggregated into four categories. It was felt that this approach was too coarse for the purposes of determining the noise exposure at a location.

2.9 It was agreed that the researchers would look again at the various assumptions.

2.10 During January 2007, the reviewers were informed that whilst carrying out these further assessments the researchers had discovered errors in their spreadsheet calculations that had caused the results that had previously been presented to be incorrect by 3.0 dB(A) and 4.8 dB(A) depending on the airport. At the same time, the following document was received:

- TN 25 Version 3 – Aircraft Noise Modelling

which presented some results of the re-assessment of the assumptions used.

2.11 A further meeting with the research team and DfT was held on 21st March 2007, prior to which the reviewers received the following documents:

- TN28 Version 2 - Survey Methodology;
- A document showing a revised Figure 4.3 with the revised noise exposure values, together with more detailed information regarding when the social survey work was carried out in each CNA;
- TN29 Version 1b – Aircraft Noise Modelling – Audit of ECAC.CEAC 3rd Edition compliant modelling;
- Information showing the revised detailed noise exposure data for the CNAs;
- Appendices to TN 29, Version 1 – Aircraft Noise Modelling – Audit; and
- An early draft of Chapter 8 of the final report on the analysis of reported annoyance (this was received the day before the meeting).

2.12 At the meeting, TN 29 was discussed and it was confirmed that the modelling issues previously identified by the reviewers had been addressed.

2.13 The reviewers received a draft final report in the middle of June 2007:

- Robust Foundations – Attitudes to Noise from Aviation Sources in England (ANASE) June 2007; and
- Robust Foundations – Attitudes to Aircraft (*sic*) Noise in England (ANASE) – Technical Appendices, May 2007

2.14 The first version of this review document, dated 20th July 2007, was based on these reports.

2.15 The penultimate version of the ANASE report consisting of the following documents was received on 4 October 2007:

- ANASE Final Report – Issue 5 dated 4 October 2007;
- ANASE Executive Summary – Issue 3 dated 4 October 2007;
- Appendix A9.1 (Addressing some methodological Concerns raised by NSPRG); and
- Sign-posting of Actions in Response to Comments relating to the ANASE Report

2.15 In the sign-posting document the researchers noted that the reviewers had raised 'two fundamental concerns' with the analysis. These were:

- The characterisation of the aircraft noise exposure; and
- The social survey process.

It went on to state that ' these comments have been discussed in paragraphs 9.6.7-8 in the main report and, in detail, in Appendix A9.2'.

2.16 In the main report no paragraphs 9.6.7 or 9.6.8 could be found and it is suspected that it was to Appendix A9.1 that the researchers were referring and not appendix A9.2 as stated. However, the reviewers noted that in the main report reference was made to their comments in paragraphs 9.4.11/12 and in one of the bullet points in paragraph 9.9.

2.17 The reviewers provided feedback on this version of the report to DfT at a meeting held on 16th October 2007.

2.18 The final version of ANASE was received on 28th October 2007 consisting of:

- ANASE Final Report;
- ANASE Executive Summary; and
- Appendices.

2.19 The researchers also provided some feedback in an e-mail dated 24th October 2007 on the comments made by the reviewers to DfT at the 16th October meeting.

2.20 This version of the review document is based on the final versions of the ANASE reports.

3 Characterisation of the Aircraft Noise Exposure

3.1 It can be seen from Section 2 that the reviewers had identified several issues that could have meant that the noise exposure results for the various common noise areas (CNA) did not necessarily represent the noise environment about which the respondents were questioned during the survey. In addition to identifying some methodology issues that were not robust, the main cause for concern was that the results used in the project for some CNAs did not correspond to the reviewers' own experience.

3.2 As was mentioned in paragraph 2.11, the reviewers received geo-referenced data regarding the CNA locations. From that, the results from the published CAA/DfT London Heathrow Summer 2005 (actual) aircraft noise exposure contours were determined by the reviewers for each relevant CNA and compared with the values shown in Table 7.1 of the final report. It should be noted that noise exposure levels become progressively less accurate below 57 dB(A) $L_{Aeq,16h}$ and the reviewers believe that any modelling results below 54 dB(A) $L_{Aeq,16h}$ this level should be treated with due caution. The results are shown in Table 1 below.

Table 1
Comparison of published CAA/DfT London Heathrow Summer 2005 $L_{Aeq,16h}$ with noise exposure used in the study

Site	Location	ANASE Study	CAA/DfT Published	Difference (Published – Study)
R01	Tooting	40.9	45.8	4.9
R02	Colliers Wood	41.6	46.2	4.6
R03	South Wimbledon	43.0	44.9	1.9
R04	Hersham	47.6	48.5	0.9
R05	Walton on Thames	47.5	48.1	0.6
R06	Sunbury	46.5	51.4	4.9
R07	Kneller Hall	55.2	56.0	0.8
R08	Kempton Park	48.9	50.4	1.5
R09	Teddington	47.2	52.2	5.0
R10	Hanworth	50.4	52.6	2.2
H3C	South Ealing	46.0	50.3	4.3
H3A	West Ealing	50.4	52.8	2.4
H5E	Virginia Water	49.6	46.3	-3.3
H3B	North Ealing	50.5	53.2	2.7
H5A	Windsor Forest	50.9	53.5	2.6
H3D	Chiswick	52.7	52.8	0.1
H3E	West Brompton	53.0	54.3	1.3
H5B	Eton	56.1	58.6	2.5
H1L	Hounslow Heath	59.7	58.9	-0.8
H1P	East Sheen	54.7	57.6	2.9
H1M	South Hounslow	59.8	59.4	-0.4
H5F	Dedworth	56.2	58.3	2.1
H5D	Old Windsor	58.7	57.8	-0.9
H1J	West Hounslow	61.7	61.8	0.1
H5C	South Windsor	59.3	60	0.7
H1K	Isleworth	60.3	59.8	-0.5
H1H	Osterley	63.1	62.3	-0.8

3.3 It can be seen that there are differences between the study values and the corresponding published values of up to 5 dB(A).

3.4 Although the key methodological concerns initially expressed by the reviewers have been addressed, there is still a discrepancy between the values used in the study and the equivalent values from the annual contour production as can be seen from Table 1. As a rule of thumb, the annual contours produce values for the summer average $L_{Aeq,16h}$ to an accuracy of +/- 1 dB(A). It can be seen that for several sites at Heathrow the difference between the study values and the published values is greater.

3.5 In Appendix 9.1 supplied to the reviewers in October 2007, the researchers observed that their results for Heathrow were based on a long term modal split, as described in

paragraph 2.6.3 of the main report, (in fact, 76%W/24%E as shown in Table 6.1 in Appendix A2[6]), whereas the so-called CAA data was based on the summer actual modal split for 2005, namely 71%W/29%E. There seemed to be an implication that this different assumption explained the numerical difference in the noise exposure results.

3.6 The reviewers do not agree with this conclusion as the 5% difference in modal split would account for no more than about 1 dB(A) difference between the results. Thus there still seems to be an issue over the characterisation of the noise exposure levels at certain locations and, in particular, at the lower noise levels. The reviewers feel that this feature is unlikely to be confined to sites around Heathrow alone (for which it was possible to make a detailed comparison). Therefore, caution is needed regarding the results for the other airports included in the study. (See also paragraphs 3.15 – 3.17 below).

Lav

3.7 There appears to be uncertainty over the definition of Lav. This issue was raised at the meeting held in November 2006 with DfT, researchers and reviewers. When asked whether Lav was defined as an arithmetic or logarithmic average, the researchers appeared uncertain, but subsequently gave the impression that it was a logarithmic average.

3.8 This view was reinforced when the reviewers received early versions of various report chapters, dated April / May 2007 where Lav was defined in words and mathematically, as the logarithmic average of a set of LAmax levels (paragraph 3.2.5 of that document). This definition was also to be found in the July 2007 version of the final report (paragraph 3.2.5).

3.9 In the penultimate version of the ANASE report presented in October 2007 Lav is now defined, in words and mathematically, as the arithmetic average of a set of LAmax levels. It is very concerning to find that the researchers appear to be unclear about such a fundamental element of their research and that such a change has occurred very late in the process. It is also disappointing that the researchers did not draw this change to the attention of DfT or the review groups.

3.10 The two definitions of Lav are conceptually and practically different. How Lav is defined has important implications for the interpretation of the study results and its conclusions. It must be remembered that the noise component in Leq / NNI is a logarithmic average. If Lav is defined as an arithmetic average then comparison with Leq / NNI and the ANASE regressions, requires a different interpretation from the same comparisons as presented in the July 2007 version. This is particularly the case with the comparisons with ANIS that are made in paragraphs 9.6 *et seq*. It must be debatable whether comparing arithmetic averaged indicators with logarithmic averaged indicators has validity, particularly in the case of the conclusions being drawn regarding the weighting that might be applied to the number of aircraft (section 9.8 of the final report).

3.11 In commenting on this change, the researchers stated in their 24th October e-mail that the previous references to logarithmic average were a “drafting error”.

3.12 It is noteworthy that in their consideration of NNI (paragraph 9.7.15 of the penultimate October 2007 version and paragraph 9.8.15 of the final version) there has been a change of wording compared with the July 2007 version (paragraph 10.3.9). In the July 2007 version, it stated:

[6] Reference is made in Table 6.1 of Appendix A2 to ‘DfT 2005 REP’ which is assumed to refer to ERCD report 0601. The long term modal split in this report is 76%W/24%E but refers to a 20 year average and not a 10 year average as implied in Para 2.6.3 of the main report. Although in Para 6.15 of Appendix A2, it suggests that the basis for the long term modal split for the designated airports is > 10 years.

To investigate this change in importance, combined metrics based on Lav and log Nav, rather than LAeq, were modelled. These are similar to the Noise and Number Index, NNI, used prior to the 1982 study. The Noise and Number Index was defined as:

whereas now, the corresponding quotation (paragraph 9.8.5) is:

The Noise and Number Index was defined as…:

i.e. with no reference to the previously mentioned similarity between L in the NNI definition and Lav. As Lav is now an arithmetic average there is of course far less similarity between the two indicators.

3.13 The reviewers feel that when drafting this section in July 2007, Lav was perceived by the researchers to be a logarithmic average and that the wording has now been changed to reflect the change in definition. Therefore to suggest that the change in definition of Lav is a 'drafting error', does not seem plausible. Given that this issue was raised by the reviewers almost one year ago, it is very disappointing that the researchers did not take the opportunity at that stage to clarify the situation.

Lav and Nav cut-off

3.14 The reviewers also have concerns that the use of cut-off values for Lav (and as a consequence) Nav were inappropriate when trying to secure a robust noise exposure value for the various CNAs. A Technical Note[7] from the researchers on this issue appeared to indicate that the calculated LAeq was quite sensitive to the cut-off values selected. Therefore, the reviewers are not convinced that the selection of cut-off values had no influence on the ANASE findings.

Aircraft Noise Modelling at low noise levels

3.15 Modelling aircraft noise below 57 dB(A) $L_{Aeq,16h}$ becomes increasingly inaccurate as the level decreases primarily because of variability in atmospheric conditions along the propagation path. It is for this reason that noise exposure contours are not routinely produced below 57 dB(A) $L_{Aeq,16h}$ and only exceptionally at 54 dB(A) $L_{Aeq,16h}$. The researchers originally appeared to have taken no account of the relative inaccuracy of noise exposure modelling at low noise levels. It is possible that inherent inaccuracy in noise modelling at low noise levels may have influenced the ANASE findings.

3.16 The researchers originally acknowledged that noise measurement becomes progressively less accurate below 57 dB(A). In the context of ANASE it is the fact that noise modelling becomes progressively less accurate below 57 dB(A) that is important. Either way the researchers do not indicate how they took this into account in planning their research. They point out that the difference between the regression lines for mean annoyance using ANASE and CAA noise data for Heathrow is not large. They have ignored the possibility that both sets of data may be inaccurate at low levels. It should be noted that CAA data are based upon the collection of over 800,000 noise data points each year and that even this amount of data collection is not sufficient for accurate noise modelling below 54 dB(A). By contrast, the ANASE study noise modelling has used a maximum of sixteen days of data in order to calibrate its results.

[7] TN 19, Effect of 65 LAmax and SEL sampling and Modelling Cut – off

3.17 In the final version of the report, the researchers do acknowledge this point at paragraph 4.2.18, although they appear not to have taken any specific account of it in the subsequent analysis.

4 Issues regarding the Social Survey Process

4.1 Social research using questionnaires can be subject to a number of effects that might bias the outcome. The context in which the questionnaire is administered can be a relevant factor influencing how a particular respondent answers questions. Also, the order in which the questions are asked is known to be a possible source of bias in social surveys. Such issues appear not to have been fully addressed during the course of the ANASE study.

4.2 The reviewers have concerns about the conduct of the social surveys that were used to elicit a measurement of annoyance. It is understood that the social surveys were administered in respondents' homes by qualified personnel. The researchers report that noise playback equipment was installed and calibrated in the respondents' homes some twenty minutes or so before the questioning commenced. The use of noise playback equipment was a requirement for the monetary valuation aspects of the study. The researchers stated that the calibration procedures were undertaken using white noise rather than examples of aircraft noise, and consequently the respondents were not overtly given prior knowledge that the survey was related to aircraft noise.

4.3 The act of setting up and calibrating equipment would almost certainly have enabled respondents to deduce that the study was about attitudes to noise. Furthermore, the fact that the social survey sites selected were located away from other sources of noise may have enabled some respondents to conclude that the study pertained to aircraft noise. In addition, it is noted that noise monitoring took place at some sites shortly before or during the social survey. It is thought possible that the act of placing and maintaining monitors during the period of social survey may have affected the survey responses from some respondents. It is possible that the noise monitoring may have been the subject of local gossip and this may have caused a more widespread effect than just for respondents in the immediate locality of noise monitors. Consequently, the reviewers believe that there is a risk that the social survey results may have been contaminated by respondent bias. That is, respondents may have used the opportunity to voice their opinion on the Government's aviation policy and may have either deliberately or sub-consciously exaggerated their reaction to aircraft noise in the way they answered the question.

4.4 In the first version of Appendix A9.1 dated 4 October 2007 the researchers suggested that the presence of response bias implies that respondents colluded in their responses. The reviewers never suggested any form of collusion was present. Furthermore, the reviewers believe that collusion is not a necessary pre-condition for response bias.

4.5 Appendix A5.2 shows the survey material and it can be seen that noise is mentioned in question 6 and aircraft noise specifically in question 7. (Questions 1 – 5 are not shown in this appendix). By contrast, the questionnaire for ANIS began with 8 questions about general perception of the local area and enquired about the respondent's intentions and motivations before asking about transport and aircraft noise. The reviewers consider that the ANASE questionnaire design precluded any opportunity to test statistically whether respondent bias is likely. The researchers have asserted throughout the review process that respondent bias is not an issue and that their pilot studies showed no propensity for respondents to amend their responses during the course of the survey. A recent technical note on the subject now acknowledges the possibility that ANASE might have suffered from more bias than ANIS[8]. (See paragraphs 4.26 – 4.31).

4.6 It has to be acknowledged that the ANASE study has taken place against a background of some public antagonism to the Government's aviation policy and extensive media interest. It

[8] TN 28, Survey Methodology, paragraph 3.2

is possible that this has had an influence on the ANASE findings. The fact that results from Phase 1 of the study which was undertaken before publication of the ATWP indicated reported annoyance consistent with the ANIS findings[9] suggests the possibility that the differences now found in the dose-response relationship may be due, at least in part, to the public interest in aviation environmental matters following the publication of the ATWP.

Site selection

4.7 The rationale for using random site selection is understood. However, it is unfortunate that the outcome was that only one site at each of Stansted, Gatwick and Luton were selected as part of the random site selection process. There is an indication that some adjustments were made to the site selection in order to avoid Heathrow related sites from dominating the study[10]. The reviewers feel that an opportunity was missed to have applied similar adjustments so that the overall study would have been seen more obviously to have applicability to other major international airports in England.

Study Scope

4.8 As a reasonable proportion of sites included in the study were related to Heathrow, the reviewers feel that an opportunity was missed to examine the attitudes to the use of segregated mode operation as a means of noise mitigation when the airport is operating on westerlies. It is fairly well known that there will soon be a consultation to consider runway usage options at Heathrow and attitudinal information about this subject would have been helpful.

Study Timing

4.9 During the review process it became apparent that the social surveys were administered between autumn of 2005 and the early part of 2006 in order to elicit views on aircraft noise during the previous 12 months. As indicated in paragraph 2.8 above, the movement data used to characterise the noise exposure was originally based on June/July 2004. The reviewers are pleased that the summer period of 2005 has now been used as a basis for determining the aircraft noise exposure (but see paragraph 5.3 below).

Mean Annoyance

4.10 Paragraph 7.2.4 of the final report states that the calculation of mean annoyance was carried out in line with other research transforming all annoyance scales to run from 0 to 100. However, it should be noted that the method employed by the researchers precludes either individual or average annoyance scores below 10 or above 90 because they have ascribed numerical values to the five point qualitative scale.

4.11 Paragraph 9.4.11 states that annoyance was assessed using the 10 point ICBEN scale, whereas the reviewers understood that mean annoyance as used in the analysis is based on the process described in paragraph 7.2.4 and Table 9.3 of the final report. The researchers through question 7 (Appendix A6.1) did seek respondents' views on their level of annoyance through what was, in fact, an eleven point scale. Apart from showing a relatively high correlation with the outcome of ascribing numerical values to the qualitative comments, no further use seemed to have been made of these data. The reviewers believe that the researchers could have made more use of the two different annoyance scales in their study in order to assess the reliability of their results. For example, the postulation in paragraph 9.5.10 would become redundant if more use had been made of the responses to the eleven point scale question.

[9] Para 5.3.16 of the main report dated July 2007 (no longer in the final version).
[10] Para 5.2.17 in the main report.

Degrees of Annoyance

4.12 In the July 2007 version of the ANASE report (paragraph 8.3.12) it was stated that 'at least 40% of respondents were at least very annoyed for all except one of the areas with LAeq greater than 57 dB(A)'. In the equivalent paragraph in both the penultimate and final versions of the report (paragraph 7.3.2) produced in October 2007 it was stated that 'for all except one of the areas with LAeq greater than 57, more than 60% of respondents were at least very annoyed'.

4.13 The latter statement does correspond to the data shown in Figure 7.2. However, closer scrutiny of that figure shows that the data presented do not correspond to the data given in Table 7.1. For example, Figure 7.2 shows that there were three sites where the percentage of respondents at least very annoyed exceeded 90%, whereas the highest value for this category in Table 7.1 was 80%.

4.14 The reviewers then examined the corresponding figure in the executive summary (final version, Figure 1) and the executive summary (July 2007 version). Two points were noted. Figure 1 in the final executive summary is identical to Figure 7.2 of the final main report, but is not compatible with Table 7.1 of the final main report. Figure 1 in the July 2007 executive summary was different from the corresponding figure in the July 2007 main report, but did appear to correspond to the data given in Table 7.1 of that version of the main report.

4.15 If Figure 1 of the July 2007 executive summary is used, then the original conclusion (as stated in the July 2007 version of the main report), namely that 'at least 40% of respondents were at least very annoyed for all except one of the areas with LAeq greater than 57 dB(A)' would seem to correspond to the data presented in Table 7.1 of the main report.

4.16 It is understood that the inconsistency between the main report and the executive summary in the July 2007 versions was mentioned at the Steering Group meeting held in July 2007. The figures in the executive summary and the main report are now consistent but do not correspond with Table 7.1 in the main report.

4.17 It is disappointing that this issue has not been resolved in the final report and instead the researchers seem to have just altered their findings (quite considerably). Again, this change has occurred, apparently, without it being highlighted to the DfT, steering group or the review groups.

Low level noise effects

4.18 In paragraph 2.3.18 of the original main report provided in June 2007 and paragraph 5.3.16 of the July 2007 report the researchers observed in regard to the Phase 1 study:

The results were consistent with findings from ANIS that only about 5% are 'very much annoyed' or 'highly annoyed' below 57 dBA Leq noise levels. Thus the initial assessment had not found any new evidence that annoyance due to aircraft noise in low aircraft noise areas had become any more significant than it has in the past.

4.19 The reviewers passed comment on this statement which had led them to question whether there was a systematic difference between the results of Phase 1 and Phase 2 of the study and, in particular, asked whether this was evidence of response bias in Phase 2.

4.20 In the penultimate version of the ANASE report (October 2007) the quotation set out in paragraph 4.18 above was not present. The reviewers did not understand why this finding had been removed from that version, and, indeed, the final version. The researchers, in their e-mail of 24th October simply stated:

The statements made in previous drafts which stated that only about 5% were only highly annoyed in low noise areas referred to pilot study data which found low levels of annoyance in areas where the sound levels were far below 57 LAeq and because of this, actual sound levels were too low to be determined with any accuracy. These pilot study data are not inconsistent with the higher levels of annoyance found in the main study in areas with higher sound levels which are still below 57 LAeq. We dropped this in the re-drafting simply because it wasn't appropriate to the argument at that point.

4.21 There are three points to note here.

- The recognition by the researchers of the difficulties, originally identified by the reviewers, of characterising noise exposure at low levels with any reliable accuracy;
- It is unclear at what noise level the phrase 'far below 57 dB(A)' refers; and
- The meaning of the last sentence is far from clear.

4.22 The reviewers still believe that the original statement provides an indication that the possibility of response bias in the main study cannot be ruled out.

ISO standard question on annoyance

4.23 The reviewers accept that the ANASE study employed guidance issued by the International Organization for Standardization (ISO). However, recent research[11] has highlighted factors not included in ISO guidance that can affect surveys, including response bias. This bias is discussed in terms of 'misfeasance', or the perception that people do not believe that *bona fide* efforts are being made to control noise. The research points out that it is difficult to quantify such effects but suggests that this bias can range from a 5 dB penalty to a 5 dB bonus dependent on the quality of the relationship between the noise-maker and the community.

4.24 It should be noted that the ISO guidance cautions that

'*compliance with the recommendations of this technical specification does not guarantee the collection of accurate, precise or reliable information about the prevalence of noise-induced annoyance'.*

4.25 It goes on to say that

'other aspects of study design, as well as uncertainties of estimation and measurement of noise exposure, can influence the interpretability of survey findings to a great extent'.

The reviewers think that both of these factors – study design and noise estimation, have not been adequately considered in ANASE.

Response Bias

4.26 In paragraph 4.6 of Appendix A9.1 the researchers suggest that the hypothesis of response bias can be dismissed because only one respondent out of 185 changed their response after considerable probing. The reviewers do not believe that this provides sufficient grounds for rejecting the response bias hypothesis. The willingness to amend an opinion, or otherwise, is

[11] Schomer P, *Criteria for assessment of noise annoyance,* Noise Control Engineering Journal **53** (4), Jul-Aug 2005

not relevant to the original hypothesis of whether the way in which the survey was conducted had any effect on the results obtained.

4.27 The researchers carried out an analysis comparing full and restricted sites to test the response bias hypothesis. Respondents at full sites were exposed to playback equipment and calibration in advance of being questioned, whereas respondents at restricted sites simply had questions to answer with no equipment being present. The hypothesis to be tested was that respondents at restricted sites (i.e. without the playback equipment) would have less opportunity to deduce the nature of the study and would, therefore, be less susceptible to response bias.

4.28 In order to reject the response bias hypothesis it would be necessary to demonstrate that any difference between respondents at the two types of site was statistically insignificant. The researchers have demonstrated that the difference between the two sites is statistically insignificant when using the ANASE noise data. However, carrying out the same regression using CAA noise data (for Heathrow alone) showed that the difference between the two types of site was statistically significant (t-statistic of 2.2). This demonstrated that the response bias hypothesis cannot be rejected.

4.29 The researchers originally stated that such assumptions were 'extreme' (paragraph 4.3 of the 4th October version of Appendix A9.1) and 'very severe' (paragraph 5.5 of the same document). The reviewers rejected these descriptions and it is noted that in the final version of Appendix A9.1 these descriptions have been withdrawn.

4.30 Far from failing to provide statistical support, the regression at Model A9.6 of the final version of Appendix A9.1 demonstrates a statistically significant difference between sites where recording playback equipment was present prior to the interview and those sites where it was not. In effect, mean annoyance is statistically significantly greater for respondents at sites where noise playback equipment was used. Far from diminishing the reviewers' concerns about response bias this analysis has reinforced them.

4.31 Given that the 'equipment effect' is but one aspect of potential response bias, the reviewers remain unconvinced that other aspects of response bias can be ruled out. It is unfortunate that the methodology employed by the researchers does not permit further analysis of these issues. Indeed, it would seem that the researchers are not entirely convinced by their own arguments on the insignificance of the 'equipment effect' because it is mentioned in their recommendations for further research in paragraph 11.5.1 of the final report.

4.32 In Appendix A9, paragraphs 4.10 – 4.16, the researchers examine the "plausibility of a consistently exaggerated response to the annoyance question". Their examples seem to be predicated on trying to see what changes in response would be needed to achieve the ANIS outcome. Unfortunately, what has not been tested is the size of change that would start to alter the conclusions that might otherwise be drawn from the results. It is the reviewers' opinion that it would take only a relatively small proportion of the respondents to have exaggerated their response to produce a different outcome.

Other changes between July 2007 and October 2007 versions of the main report

4.33 It was noted that the intercept coefficient in the regressions in Model 9.1 of the July 2007 version of the main report and its equivalent (Model 8.1) in the October 2007 version were different.

4.34 It was also noted that the fourth bullet point in paragraph 9.10 of the final report has been altered from its equivalent in the July 2007 version and the penultimate version of October 2007. The earlier versions stated that:

For a given mean annoyance score, the 16-hour LAeq value in the ANASE survey is higher by between 4 dB (for a mean annoyance score of 10 or not at all annoyed) and 11 dB (for a mean annoyance score of 90 or extremely annoyed).

The equivalent bullet point in the final version states that for a given mean annoyance score the LAeq value in the ANASE study is higher by 3 dB (for a mean annoyance score of 10 or not at all annoyed).

4.34 In a similar vein, the penultimate version of the final report stated in paragraph 9.9 that

Measured on a scale of mean annoyance, the level of annoyance is consistently 16 points greater in ANASE than it was 23 years ago (where a difference of one category on the ANASE annoyance scale is allocated 20 points)

The difference of 16 points was also noted in the July version (paragraph 10.4). However, in the final version (paragraph 9.10) the difference is stated to be 14 points.

4.35 The points in the preceding paragraphs are relatively minor, and presumably are typographical issues. Nonetheless it would have been helpful for these changes to have been highlighted to the DfT and the reviewers.

5 Specific Issues

5.1 This section identifies various specific points originally identified with respect to the July 2007 report but which appear to remain outstanding.

Paragraphs 2.6.3 and 2.8.2

5.2 The approaches to the modelling described here are not invalid. The reviewers are surprised, however, that more robust data regarding these features were not obtained.

Paragraphs 2.6.3

5.3 The meaning of 'summer' regarding the movement data is not clear. Presumably it is the 92 day (16 June to 15 September) period.

Paragraphs 4.2.5 – 4.27

5.4 The argument in this paragraph is not convincing. Further light would be shed on the issue if the $L_{Aeq,\ 8\ hrs}$ contours were compared between the two airports.

5.5 It is noted that the researchers recommend that more work might be carried out in this area (paragraph 11.5.1 of the final report).

Paragraph 5.2.13

5.6 The categorisation in Tables 5.2 and 5.3 of 'High' noise does not precisely coincide with the definition in paragraph 5.2.13.

Paragraphs 6.2.6

5.7 Further clarification of the methodology for audio presentation would have been helpful.

Paragraph 6.5.3

5.8 It would have been helpful to have a little more explanation about this incident. Was it related to the study or to something else? Might this indicate a bias in the results in this area?

Table 7.1

5.9 The reviewers only saw this very helpful table in the middle of June 2007 when the draft report was received. (It was not in the draft Chapter 8 received just before the March 2007 meeting – see paragraph 2.11 above). Comments are provided in paragraphs 4.11 – 4.16 of this review.

Paragraph 7.3.19

5.10 Clarification would have been helpful regarding what exactly is the L_{Amax} that is being used here.

Paragraph 8.3.1

5.11 Might another relevant factor influencing reported annoyance have been any expansion proposals for the airport in question (e.g. those set out in the ATWP)?

Paragraph 8.3.3

5.12 First bullet – given the possibility of development at Heathrow, it is a pity that this finding has not been pursued further. The researchers seemed to have simply focused on the site level conclusion, whereas it appears that at an individual level, there is a difference between those affected at Heathrow compared with the other airports.

6 Conclusions that could be drawn from this Study

6.1 The reviewers felt that it might be helpful to highlight some headline conclusions that might be drawn from the study results were they to be taken at face value. These are set out below.

Dose – Response relationship

6.2 In Section 4, Appendix B, Annex 4, Table 3 of CAP 725[12], the current relationship between $L_{Aeq,16h}$ and % highly annoyed is shown. It is based on ANIS, but modified in the light of limited follow-up studies. For ease of reference, Table 2 below reproduces that data:

Table 2
Percentage of People Highly Annoyed

Contour Band	% Highly Annoyed
54 – 57	6.6
57 – 60	11.1
60 – 63	18.0
63 – 66	28.0
66 – 69	40.7
69 – 72	54.9
72 - 75	68.2

6.3 A logistic expression is provided that suggests that the % highly annoyed at 57 dB(A) $L_{Aeq,16h}$ is 8.5%.

6.4 If it is presumed that the ANASE term 'at least very annoyed' is equivalent to 'Highly annoyed' in Table 2 above, then from Figure 7.2 in the main report, the level of 8.5% 'highly annoyed' is now occurring at exposures of around 44 dB(A), $L_{Aeq,16h}$ (assessed by eye). [However, if this Figure is incorrect (see paragraph 4.13 above) the value would be 47/48 dB(A), $L_{Aeq,16h}$ (assessed by eye)]

6.5 Assuming that this level of annoyance (i.e. 8.5% highly annoyed) is still to be described as the 'onset of significant community annoyance', then the research is suggesting that that this onset is occurring at a level some 10 – 13 dB(A) lower than before. The percentage 'at least very annoyed' at an exposure of 57 dB(A), $L_{Aeq,16h}$ is, according to ANASE, now around 40% – 65%, depending on whether Figure 1 in the executive summary(July 2007 version) or Figure 7.2 (final main report) is used.

6.6 ANASE includes a direct comparison with ANIS at Figure 9.2 in terms of mean annoyance. Here it can be seen that the level of mean annoyance that occurred at 57 dB(A), $L_{Aeq,16h}$ with ANIS is now occurring with ANASE at a level of just over 50 dB(A) $L_{Aeq,16h}$. Again, therefore, ANASE is suggesting that the onset of significant community annoyance is occurring at lower levels of exposure to aircraft noise.

[12] CAP 725 – CAA Guidance on the Application of Airspace Change Process (March 2007), Civil Aviation Authority

Impact of the change in the number of movements

6.7 At paragraph 9.10 in the final report, it is suggested that giving the number of aircraft movements a weighting more akin to NNI (i.e. 15) rather than 10 might be more appropriate for predicting annoyance from aircraft noise. Thus, assuming the fleet mix is unchanged, a doubling of movements would mean an increase of 4.5 dB(A) at a given location, rather than 3 dB(A).

6.8 On the basis of this result, there would be two main consequences:

- A different approach would be required for any noise impact assessment of proposed changes in an airport's activity;
- The current quota count system would not necessarily properly reflect the relationship between noise exposure and number of movements

Conclusion

6.9 It can be seen that acceptance of the results of ANASE would affect the current understanding of the impact of aircraft noise and the effect of changes in the exposure to aircraft noise.

7 Overall Conclusions

7.1 The reviewers recognise the very challenging task that faced the researchers when undertaking this project. Issues concerning aircraft noise have rarely been out of the media's attention for long and to secure a completely unbiased set of responses was probably almost impossible.

7.2 The original need for the study was identified after the Heathrow Terminal 5 Inquiry and the Inspector's comments. Concern was expressed at that time that the dose-response derived from ANIS may have become uncalibrated, and even during the Inquiry, there was speculation that the percentage highly annoyed at 57 dB(A), $L_{Aeq,16h}$ was rising.

7.3 The reviewers have noted that the ANASE study also embraced the issue of Stated Preference / Willingness to Pay. In the original Press Notice for the study[13], it was stated that:

It will focus on how people perceive the relationship between noise levels and annoyance, or sleep disturbance at night, and how they would ***value*** *lower noise levels relative to other environmental factors.*

The research will look afresh at the major study completed in 1985, the Aircraft Noise Index Study, and on the findings of three recent Government-sponsored studies on sleep disturbance.

7.4 The fact that the Non SP Review Group was only formed in May 2006 suggests that the study focused rather more on the monetary valuation element than on the annoyance element. Furthermore, there is little in the study report that sheds further light on the issue of sleep disturbance.

7.5 It was disappointing that the reviewers found a range of issues regarding the characterisation of the noise exposure when they first looked at the reports. It took over six months for the various points to be addressed, in which time a spreadsheet error was also uncovered. The value of the $L_{Aeq,16h}$ used to characterise the noise exposure of the CNAs has changed by typically 2 -3 dB(A) from those originally presented in June 2006. Even now there still seems to be a discrepancy when comparing the ANASE values with the equivalent results from the published contours. (paragraph 3.2 above).

7.6 Of more concern is the impact on the responses of the social survey methodology as described in Section 4 above. Whilst during the review process various Technical Notes were presented to demonstrate that no bias or distortion of the results had occurred, the reviewers are not convinced that none exists. Their concerns have been borne out by the recent sensitivity tests that have shown possible bias (Appendix A9.1). It must be remembered that in determining the mean annoyance from the social survey responses, a difference of one category is worth 20 points (paragraph 7.2.4 in the main report). Figure 9.2 shows that the difference in mean annoyance at 57 dB(A) $L_{Aeq,16h}$, from 23 years ago is about 14 points. This change is smaller than the difference between two adjacent annoyance categories, thus demonstrating the sensitivity of the results to any potential response bias.

7.7 The reviewers always recognised the potential conclusions that could be drawn from the results of ANASE (as outlined in Section 6 above). During the review, the researchers made reference to several other recent studies that had found similar results to ANASE, i.e. the noise level was reducing at which a certain annoyance response was found. Consequently, it was being argued by the researchers that the ANASE results were not unusual compared with what was happening elsewhere.

[13] DTLR NEWS RELEASE - 2001/0269, 08 May 2001, Ainsworth unveils new study into aircraft noise

7.8 Reviewing in detail the other studies mentioned was not part of the reviewers' scope, but from what has been seen of the other studies the relatively larger adverse response for a given noise exposure could, in part, be attributed to local circumstances.

7.9 Given the above, the reviewers feel that the results of ANASE study are inconclusive and therefore should be treated with caution. Although the issue of the noise exposure characterisation of the CNAs could, in theory, be resolved by using the published values, the issues raised regarding the social survey cannot be addressed without repeating the survey using a modified approach that minimises the risk of bias.

7.10 The researchers seemed to have been under the impression that the reviewers were taking the position that they did not believe the results from ANASE should be different from ANIS. This is not the case. The reviewers have held a neutral stance throughout, being concerned solely that the results of the research are based on robust and reliable science. The reviewers concerns about the reliability of the characterisation of the noise exposure (especially at the lower sound levels) and more importantly the potential for response bias have been borne out.

7.11 There are three other issues where the study could have benefited from more work. These are:

1. A more detailed investigation into the possible differences in response for sites close to Heathrow compared with the other airports (paragraph 5.14 above);

2. An evaluation of the percentage at least very much annoyed with the indicator Lden, given its use in the Environmental Noise (England) Regulations 2006; and

3. An examination into the attitudes to the use of segregated mode as a means of noise mitigation when Heathrow airport is operating on westerlies (paragraph 4.8 above).

7.10 The response to the reviewers concerns presented in Appendix A9.1 and the analysis therein does not, as the researchers suggest, provide the assurance that the ANASE methodology is sound or that the possibility of response bias can be ruled out. Rather, the analysis shows a statistically significant difference between sites where playback equipment was present in advance of the interview and those sites where it was not. The results of this statistical analysis has provided support for the reviewers' concerns about the ANASE methodology.

7.11 The reviewers are concerned about the overall robustness of the ANASE study. Throughout the review process, it can be seen that the reviewers were presented with

- changing data(paragraph 2.9 above);
- the discovery of errors (solely it seems as a result of the reviewers enquiries) (paragraph 2.10 above);
- a revised definition of an important variable (paragraphs 3.7 – 3.13 above);
- findings that did exist and now do not (paragraph 4.20);
- changing conclusions (paragraphs 4.33 – 4.34 and paragraph 5.12 above); and
- even now final versions that contain inconsistencies (paragraphs 4.12 – 4.17).

These issues have caused the reviewers to lose confidence in the work of the researchers.

7.12 In the first version of this review it was stated that there were sufficient technical and methodological uncertainties still remaining with the study to mean that reliance on the detailed outcome of ANASE would be misplaced. In view of developments since the review of the July 2007 version of the ANASE main report, the reviewers are even more convinced that their concerns are fully justified.

7.13 Given the issues described above, the reviewers would counsel against using the detailed results and conclusions from ANASE in the development of Government policy. It is further suggested that the Department for Transport should issue guidance to this effect and also to clarify what reliance, in their view, should be placed on the ANASE finding by planning authorities, regulators, practitioners and other interested parties. Without such guidance, there is a strong risk that there will be many different and conflicting interpretations of the information contained within the ANASE reports.

7.14 It has already been noted that the ANASE study arose out of a Ministerial commitment six years ago. Interest in the progress and outcome of this study has been considerable. Despite the comments made above, the reviewers believe that the ANASE study should now be published without delay.