NOTICE OF FILING

This document was filed electronically in the FEDERAL COURT OF AUSTRALIA (FCA) on 1/02/2013.

DETAILS OF FILING

Document Lodged:	Expert Report(s)
File Number:	VID969/2010
File Title:	DANIEL HAILE-MICHAEL & ORS v NICK KONSTANTINIDIS & ORS
District Registry:	VICTORIA REGISTRY - FEDERAL COURT OF AUSTRALIA

Dated: 1/02/2013   **Registrar** *Warwick Soden*

Third report of Professor Gordon, Proceeding No. VID 969 of 2010

Professor Ian Gordon   PhD, AStat, Director
Statistical Consulting Centre
The University of Melbourne

31 January 2013

Executive Summary

1. This report comments on two reports prepared by Dr John Henstridge. The reports relate to the following Proceeding: **Daniel Haile-Michael & Ors v Nick Konstantinides & Ors, Federal Court of Australia Proceeding No. VID 969 of 2010**. His first report was dated 5 December 2012 and his second report was dated 21 December 2012. For convenience, I refer to Dr Henstridge's first report as "JH1" and his second report as "JH2".

Preliminaries

2. I have previously provided two reports in relation to the Proceeding. A commentary on my reports was included in JH1. This third report of mine deals with some issues raised in JH1. I also comment on the methodology and findings contained in JH2. The relevant background context and information for this report, such as my original letter of instructions and my own background and qualifications, are dealt with in my first report. In this report I use the same terms as those in my earlier reports and assume that these terms are familiar to someone reading this report. In addition to previous material provided, noted in my previous reports, I have been provided with the following:

- A letter of instructions to Dr Henstridge dated 26 October 2012.
- A letter of instructions to Dr Henstridge dated 19 November 2012.
- A report by Dr John Henstridge dated 5 December 2012.
- A report by Professor Andrew Goldsmith dated 10 December 2012.
- A letter of instructions to Dr Henstridge dated 17 December 2012.
- A letter from Norton Rose Australia to Arnold Bloch Leibler dated 21 December 2012, enclosing the following two documents.
- A report by Dr John Henstridge dated 21 December 2012.
- Outline of evidence of Craig Darragh.
- Electronic copies of Annexures CD-1 to CD-6.
- A letter from Norton Rose Australia to Arnold Bloch Leibler dated 17 January 2013, enclosing ABS Census data.
- The ABS Census data referred to in the previous dot point.
- A letter by Dr Henstridge, dated 15 January 2013, containing corrections to his first report.

  On 21 January 2013 I received an email from Mr Peter Seidel of Arnold Bloch

Leibler, with the following request:

"We refer to the expert reports of Dr John Henstridge dated 5 December 2012 and 21 December 2012.

Would you please, in accordance with the Expert Code of Conduct, answer the following questions:

Would you please provide any comments you have on the report of Dr John Henstridge dated 5 December 2012. To the extent that you agree with any criticism Dr Henstridge has made of the methodology in your report of 11 September 2012, please update that report to address those criticisms. To the extent that you disagree with any criticism of your 11 September 2012 report, please state why you disagree.

Would you please comment on the methodology and findings contained in the report of Dr John Henstridge dated 21 December 2012."

3. I was provided with guidelines for expert witnesses providing evidence in the Federal Court. I have read, understood and complied with these guidelines in the preparation of my report.

Responses to matters in Dr Henstridge's first report

4. In JH1 Dr Henstridge has a section called "The Data Files". His findings and views here are in substantial agreement with mine, in general.

5. At paragraph 7(e) he comments that I assume that the 4917 rows of data in 9.00001.xls "refer to 4917 interactions between police and a specified male person". He is critical of this, pointing out that there are duplicate rows. I have a number of responses to this.
    a. I have not been provided with any detailed description of the files 9.00001.xls to 9.00006.xls. In particular, I have not been informed how the different files were constructed, or, in general, the nature and coding of the variables. I made this point at paragraph 17 of my first report. I was instructed to assume that the files complied with the Court Order for the provision of data. The absence of descriptions of the files and variables is unusual and undesirable. Of necessity, I relied on scrutiny of the files themselves to make sense of them.
    b. At paragraph 18, I did not state that the 4917 rows of the files contain 4917 interactions, but, rather, 4917 records of interactions. Even if there are duplicates, I assume the rows are records of interactions; this is a simple consequence of the assumption that the provision of the data complied with the Court Order.
    c. Dr Henstridge suggests that the existence of identical rows in file 9.00001.xls implies that these are duplicate records, and not separate

---

interactions. However, two records of the same individual on the same date may be different interactions if they occur at different times of the day. The time of the interaction is not recorded in file 9.00001.xls. More than one interaction on the same day at different times explicitly occurs in some of the other files, which include the time of day of the interaction.

6. At paragraph 7(h) in JH1 Dr Henstridge comments on the non-uniqueness of records in file 9.00003.xls, stating that there are 1916 rows of data, but only 930 combinations of <idno> and date, suggesting, implicitly, a significant problem of duplicate records in file 9.00003.xls. In Dr Henstridge's letter of 15th January 2013, he notes that this is an error, and that there are in fact 1825 unique combinations of <idno> and date. However, file 9.00003.xls contains other information about the interaction. In fact, all 1916 rows of data in file 9.00003.xls have unique combinations of <idno>, date, time and contact ID. "Contact ID", as I understand it, refers to the field contact number. Accounting for it demonstrates that the records are all different, which may mean that all 1916 rows are different interactions, contrary to Dr Henstridge's inference.

7. This point applies generally; even allowing for Dr Henstridge's corrections in his letter dated 15th January, it is not, in my opinion, clear that there are duplicate records to the extent that he suggests, for the reason explained in paragraph 6 above: non-uniqueness of <idno> and date is not sufficient to claim duplication of records.

8. In other respects Dr Henstridge and I are in agreement in the matters he discussed in JH1 in the section called "The Data Files".

9. In JH1 there is then a section called "Census Data". There are three matters I would like to respond to in this section.

10. At paragraph 14a of JH1 Dr Henstridge points out a typographical error in my paragraph 104, namely, that the equation has the figure of 9.6% instead of 4.3%. As he notes, I used the correct figure in the calculation.

11. In paragraph 13d, Dr. Henstridge notes that I have assumed that there were no African/Middle Eastern individuals in the "Not Stated" categories of the 2006 Census data that I used to estimate the percentage of relevant males in the suburbs of Flemington and North Melbourne as 14.5%.
    a. I agree with Dr. Henstridge that my approach is likely to underestimate the percentage of African/Middle Eastern males in the relevant area.
    b. I agree it is appropriate to consider how the estimate of the percentage of relevant males in the suburbs of Flemington and North Melbourne will vary depending on the assumptions made about how missing information about ancestry arises for the group of interest. I concur with Dr. Henstridge's estimate of the weighted percentage of relevant males in the suburbs of Flemington and North Melbourne of 18%, assuming ancestry information is missing at random.

    c. Dr. Henstridge uses two values, 18% and 25%, for the percentage of African/Middle Eastern males in the suburbs of Flemington and North Melbourne. His choice of the value of 25% appears to be arbitrary.

    d. Dr. Henstridge suggests that there is evidence for an association between the percentage of the population that is African/Middle Eastern and the proportion of "Not Stated" responses on the 2006 Census. He does not quantify the magnitude of this association, nor does he provide enough detail of his work to allow an evaluation of his claim.

    e. The uncertainty about the mechanisms underlying missing ancestry information in the group of interest means there is a great deal of uncertainty about the true value of the percentage of interest.

    f. The most appropriate estimate, having regard to Dr Henstridge's initial point about the missing data, is 18.0%, in my opinion.

    g. Finally, the difference between my initial figure of 14.5% and the figure I now believe to be the appropriate estimate, namely, 18.0%, is not material, in the sense that the comparisons with this figure are not materially altered. I explore this more below, in paragraph 15.

12. In paragraph 17(c) in JH1 Dr Henstridge makes comments about racial profiling. I was not asked about racial profiling and my report did not attempt to answer questions about this.

13. In the section "Question 1" in JH1, Dr Henstridge and I are in substantial agreement, apart from matters I now discuss.

14. In paragraph 22 of JH1, Dr Henstridge notes that I did not give the details of the statistical test I carried out. My reason for not doing so is that it is somewhat technical, but I now provide it here. Dr Henstridge and I agree that the test needs to account for the dependencies in the data, and we further agree that it is difficult to do so, given that individuals have a varying number of interactions, and different individuals can be involved in the same interaction. I used the following approach; this is not a complete description and can be expanded further if necessary. I consider that a "large sample" normal approximation is reasonable. In each case, I therefore sought to obtain the standard error of the estimated fraction of interactions that were African/Middle Eastern. Define the number of interactions relating to African/Middle Eastern males to be X, and the number of interactions relating Other males to be Y. The fraction of interest is equal to $X/(X+Y)$. Note that X and Y can each be expressed as sum of interactions, summing the numbers of interactions of individuals in each group separately. Assuming that the number of interactions per individual is independent between individuals, we can estimate the mean and variance of X and Y. With this background, I used a Taylor series approximation to obtain an estimate of the variance of $X/(X+Y)$. While this method is approximate, and only accommodates some of the dependence in the data, it gives a standard error that is considerably larger than would be obtained by assuming complete independence. In my opinion it is a reasonable approach.

15. The comparisons made using this approach were provided at paragraph 110 of my first report. The file percentages were 33.1%, 29.5%, 45.6%, 25.2% and 26.1%. In my first report I used the test to compare with the percentage of 14.5%. As outlined above, I now consider that the appropriate percentage to compare them with is 18.0%. The P-values for each of the five comparisons (33.1% with 18.0%, 29.5% with 18.0% and so on) are accordingly different than they were in my first report. They are now:
P < 0.001 (9.00001.xls, 33.1% compared to 18.0%);
P < 0.001 (9.00002.xls, 29.5% compared to 18.0%);
P < 0.001 (9.00003.xls, 45.6% compared to 18.0%);
P = 0.01 (9.00005.xls, 25.2% compared to 18.0%);
P = 0.006 (9.00006.xls, 26.1% compared to 18.0%).

My custom generally is to report any P-value smaller than 0.001 as P < 0.001 (rather than P = 0.0000001 for example). The P-value for the 9.00003.xls (containing field contacts, including person checks), in particular, is – as calculated – very small, so that reporting it as P < 0.001 is quite conservative. All these P-values are small, and most of them are much smaller than the commonly-used threshold for statistical significance, namely, 0.05. Therefore the conclusions I drew at paragraph 110 are not materially altered by substituting the revised Census value of 18.0% for the one I originally used (14.5%).

16. In paragraph 27 of JH1, Dr Henstridge is critical of my use of the t-test to carry out the comparison of the average number of offences in file 9.00005.xls, comparing African/Middle Eastern males with Other males, confined to males with one or more "OFFENDER" record in file 5. First, the key point is that the average number of offences among the Other males (namely, 12.3) was greater than the average number of offences among African/Middle Eastern males (namely, 7.8). This is reported at paragraph 120 in my first report. The importance of this is that, in interpreting the profile of the two groups with respect to field contacts and person checks in file 3, it is instructive to examine whether they are otherwise similar, with respect to file 5 information. For example, a logically possible explanation of a higher rate of file 3 person checks among the African/Middle Eastern "OFFENDER" males from file 9.00005.xls (which is what was found) would be that they had a greater rate of offending in 9.00005.xls. However, the opposite is the case: they had a lower rate of offending.

17. It is a further and somewhat secondary question as to whether the difference is *statistically significant*. The average for African/Middle Eastern males was lower than the average for Other males.

18. Dr Henstridge suggests that the data are not normally distributed and that therefore a two-sample t-test is not appropriate. I did not and would not assert that the data are normally distributed; my basis for suggesting that the t-test was a reasonable approximation was that the sample sizes are large (50 and 118) and that therefore a Central Limit Theorem approximation was likely to be applicable. Dr Henstridge suggests that the data may be too skew for this

approximation to be adequate, and investigates other tests, including the Mann-Whitney test and three other tests based on transformations of the data, obtaining P-values of 0.149 (Mann-Whitney), 0.042 (cube root transformation) and 0.054 (fourth root transformation). The P-value for the square root transformation, examined but not reported by Dr Henstridge, as he preferred the others, is 0.023. I replicated the P-values obtained by Dr Henstridge. In the light of these further tests, I now consider that the statistical significance is less strong than claimed by me in the first report; as noted above at paragraph 17, the statistical significance (or otherwise) is a rather secondary issue in my opinion.

19. In paragraph 31d, Dr. Henstridge discusses the limitations of using Fisher's exact test in the analysis of presence of one or more instances of particular phrases (listed in paragraph 30) in field contacts with groups of one or more individuals. The data are from file 9.00004.xls. Groups were classified as "All African/Middle Eastern", "Mixed" or "No African/Middle Eastern". I agree with that there is a limitation of Fisher's exact test in this context; it assumes the contacts are independent. This is unlikely to be true, for some of the field contacts. The structure of the data here is quite complex: field contacts can involve one or more males, the males in a single field contact with multiple individuals may be all African/Middle Eastern, all Other, or mixed, and any single male may be in several field contacts.

    a. I carried out an additional analysis based only on the field contacts where a single individual was involved. Based on the information in file 9.00003.xls, there were 1563 field contacts, of which 1326 (that is, 85%) involved a single male. So the great majority of field contacts involved just one male. For these field contacts, there is only one type of dependence in the data, which can be accounted for in the analysis.

    b. There were 262 males corresponding to these 1326 field contacts. The males in these contacts are either African/Middle Eastern or Other. There were 90 African/Middle Eastern males and 172 Other males.

    c. The number of field contacts per male ranged between 1 and 71.

    d. I calculated the percentage of field contacts, per male, that had any of the nominated phrases. The average percentage among the African/Middle Eastern males was 23.0%, and the average among the Other males was 8.7%.

    e. I used logistic regression to examine whether the proportions of field contacts with the nominated phrases were significantly different in the two groups, allowing for extra-binomial dispersion. (There was little evidence of over-dispersion, in fact: the residual deviance was 260.5 on 260 degrees of freedom.)

    f. The result was that the difference between the two groups was highly statistically significant, with a P-value of $P = 0.001$.

20. This result is consistent with the result in my first report, and addresses Dr Henstridge's concern about independence. It is not surprising that the results are consistent, since the great majority of field contacts involve just one person, as

---

explained above.

21. I have no comment on the matters covered in the sections "Question 2", "Question 3" and "Question 4" of JH1.

22. Regarding the section "Question 5" in JH1: I agree with the percentage of 29.8% obtained in paragraph 45. The same distribution is given in my first report at paragraph 124 (although the percentage is not explicitly mentioned).

23. Regarding the section "Question 6" in JH1: while I agree with the results of the statistical tests, I consider that the choice of 25% as a comparison population percentage to be arbitrary, as discussed above at paragraph 16.

24. I now turn to the section "Question 7". There is an error in the answer Dr Henstridge has provided to this question. The question put to him, repeated in his report, is restricted to "specified male persons who have at least one involvement identified as "offender" in File 5", that is, file 9.00005.xls. However, Dr Henstridge's provides an answer that it is not so restricted, and includes all specified male persons in File 5. This can be seen by reference to his report. Question 5, paragraphs 44 to 46 in JH1, is about the same group, namely, the specified males "with Offender involvements" in file 5. I agree with the data obtained at paragraph 45. However, for the same group, at paragraph 50 in JH1, different frequencies are shown. These are not the correct frequencies for the specified male persons who have at least one involvement identified as "offender" in File 5; they are the frequencies for all specified male persons in file 5. The frequencies under the column "Offenders in File 5" in the table at paragraph 50 should agree with the frequencies under the column "Number of individuals" in the table at paragraph 45, since they are the same defined group. But they do not agree: this is an error.

25. The consequence of remedying this error would be that Dr Henstridge's answer to Question 7 would be based on the same data as I obtained in response to Question Sept-03.3, when the specified male persons were restricted to those with at least one "offender" involvement in file 5, that is, 9.00005.xls, and the essential conclusions would be correspondingly consistent. These results, shown in my first report at paragraphs 125 and 127, are that among specified males with at least one "offender" involvement in file 5, those with African/Middle Eastern ethnic appearance had significantly higher rates of at least one field contact and at least one person check. The corresponding P-values were P = 0.005 for field contacts and P = 0.005 for person checks.

26. At paragraph 28 in JH1 Dr Henstridge indicates that he agrees with my approach and conclusions in relation to Question Sept-03.3.

27. The section headed "Question 8" is in relation to Professor Cunneen's evidence.

28. The section headed "Question 9" is also about Professor Cunneen's evidence. However, in Dr Henstridge's commentary at paragraph 55 he returns to an issue discussed at paragraph 31(d) in JH1. I have addressed the criticism he made there in this report, at paragraphs 18 and 19. The response there also applies here.

29. I note, in passing, as minor issue, that there is a typographical error in paragraph 55 in JH1: the number of field contacts that I analysed is said by Dr Henstridge to be "15,313". This is incorrect. The total number of field contacts of 1,916. The number of field contacts corresponding to the data Dr Henstridge is referring to in paragraph 55 is in fact 1,858.

30. The section "Question 10" in JH1 is about Professor Cunneen's evidence.

31. The section "Question 11" concerns analyses in my second report. I have a number of comments about this.

32. Dr Henstridge points out in paragraph 62 (and reiterates a number of times, as relevant) that in these analyses I included four (4) males of "Indian" ethnic appearance in the African/Middle Eastern group, and that this is inconsistent with my previous analyses. He is right about this: I did not intend to do this, the problem arose because of a typographical error in some code.

33. The consequence of this for the analyses considered here is very small, as noted by Dr Henstridge. In fact, remedying this error makes the statistical conclusions very slightly stronger.

34. In JH1 there is a lengthy technical discussion concerning the use of a generalised linear mixed model for these data. Dr Henstridge's eventual conclusion is that the model may not be applicable. I am unconvinced by his reasons. I note that at paragraph 69(j)ii he fits a different statistical model (a reasonable one in my view) and obtains a similar P-value to that obtained by my approach, namely, P = 0.00004. I regard that as supportive of the P-value and conclusion obtained using my approach. The conclusion is that the percentage of involvements that were "OFFENDER" was lower in the African/Middle Eastern (LEAP) group than in the "Other" group, and the difference is strongly statistically significant.

35. At paragraph 69(j)iii Dr Henstridge fits a different model, which he describes as a "highly robust model", that considers whether an individual had any "OFFENDER" involvements or not. The P-value for this model was 0.065. It is not clear why this model was be preferred over Dr. Henstridge's second model; the "robust model" loses some information, namely, the number of "OFFENDER" involvements per individual. It would be expected that the "robust model" would be less sensitive to differences between the African/Middle Eastern group and others, given this information loss. This model is based on different data: it is based simply on whether or not the male person had any offender involvements. While it is a valid test as far as it goes, it is not

what I sought to examine in paragraph 18 of my second report. For example, it treats the record of a male who had 1 offender involvement out of 10 in the same way as a male who had 9 offender involvements out of 10. This entails a loss of information, so it is not surprising that the test is less sensitive than that applied to more comprehensive information.

36. For all of these reasons I do not agree with Dr Henstridge's criticism of my analysis of this question; my view remains that the lower rate of offender involvements among the African/Middle Eastern males in file 9.00005.xls is strongly statistically significant. As shown in Table 3 of my second report, the percentage of involvements that were "offender" involvements were 58.2% among African/Middle Eastern males and 73.3% among Other males.

37. I do not have comments on the section headed "Question 12" in JH1.

Responses to matters in Dr Henstridge's second report

38. In Dr Henstridge's second report, dated 21 December 2012, which I refer to here as JH2, comparisons are made between groups of males defined in various ways, and the males in file 9.00003.xls.

39. As a preliminary matter, I note the point that Dr Henstridge makes concerning the identification of individuals, at paragraph 13(c). He notes that a different set of identifiers are used in the two sets of files: the files 9.00001.xls to 9.00006.xls (on the one hand) and the Darragh annexures, CD-1 to CD-6 (on the other hand).

40. I requested that I be provided with a file containing the correspondence between the two sets of identifiers; at the time of writing this report I have not been provided with such a file.

41. The fact that different identifiers are used in the two sets of files is a serious deficiency for the purposes of a proper analysis of these data, since it means, as Dr Henstridge says at paragraph 13(d), that "with the information available it is not possible to carry out a test that is strictly correct".

42. I would go further than that and say that the statistical test carried out by Dr Henstridge is definitely not correct, that is, not applicable, because the groups he compares are likely to overlap a lot: some of the males in the Darragh annexures are likely to be in file 9.00003.xls, rendering the independence assumption palpably incorrect. A male who committed a robbery or assault in Flemington or North Melbourne in the relevant period, satisfied the age and residence (CD-1, CD-2, CD-5) criteria, and who also experienced a field contact in the relevant period, will be in both files. Dr Henstridge's test implicitly proceeds on the assumption that there are no such males. He acknowledges that this is a problem; I think it is a serious problem.

43. While the use of the different identifiers is a serious problem for analyses, the field "sub-incident", assuming it means the same thing in both sets of files, is potentially useful. I made attempts to reconcile some of the information in the Darragh annexures with the information in file 9.00005.xls. I found that in some cases there were sub-incidents in the Darragh annexures which I was unable to find in file 9.00005.xls. Examples (not a comprehensive list) are sub-incident numbers 50217961 and 50248584. These are in CD-5 but not in 9.00005.xls. Since the selection of records in the Darragh annexures (as I understand it) was more restrictive than file 9.0005.xls, it was not clear to me why this would be so.

44. In investigating this question, I noted that the initial set of files appears to have been selected on the basis of a "proc_date" (which I assume to mean "processed date") between 1 January 2005 and 31 December 2008, while the Darragh annexures appear to have been selected on the basis of a "commit date" in the same calendar period; the "commit date" is the date on which an incident is recorded as having reportedly taken place, according to the Darragh affidavit. This is an apparent difference between the sets of files which makes any comparison difficult.

45. A further and important difference between the sets of files is that in the Darragh annexures the robbery, assault or armed robbery was recorded as taking place in Flemington or North Melbourne; there was no such restriction on the location of incidents in files 9.00001.xls to 9.00006.xls.

46. For all these reasons, the similarity between the percentages of African/Middle Eastern males in the Darragh annexures (on the one hand) and the percentage of African/Middle Eastern males who had at least one person check in file 9.00003.xls (on the other hand) is, in my view, difficult to interpret statistically.

47. The Darragh annexures all had the selection criterion that the male concerned was in some way involved in a robbery, assault or armed robbery. I note that there is no indication in the questions asked of Dr Henstridge, or his answers, as to the basis for this selection.

48. It is possible to make a comparison of the sort carried out by Dr Henstridge in his second report, using the information in files 9.00005.xls and 9.00003.xls. At an overall level, this has already been done: of the males in file 9.00005.xls with at least one "offender" involvement, 50/168, or 29.8% were African/Middle Eastern, while in file 9.00003.xls, of the males with at least one person check, 116/258, or 45.0%, were African/Middle Eastern. These groups are not independent. In other respects, however they are comparable, insofar as they come from the same set of files.

49. The file 9.00005.xls contains an offence description. If the males in file 9.00005.xls with at least one offender involvement are further selected on the basis of having at least one offence description of a particular type, and the same comparison is made, then it is to be expected that there will be variation in the percentage of

African/Middle Eastern males; sometimes it will be lower than the overall percentage of 29.8%, and sometimes higher. For illustrative purposes, I examined the three most frequent offence descriptions in file 9.00005.xls. These were "burglary", "criminal damage (intent damage/destroy)" and "theft from shop (shopsteal)". The percentage of African/Middle Eastern males in 9.00005.xls with at least one offence of each of these offences descriptions were, respectively, 14/51 (27.5%) for "burglary", 13/63 (20.6%) for "criminal damage (intent damage/destroy)" and 12/51 (23.5%) for "theft from shop (shopsteal)". As it happens, these three percentages are all lower than the overall percentage of 29.8%. In each case, they are all substantially lower than the percentage of African/Middle Eastern males with at least one person check in file 9.00003.xls, namely, 45.0%. As mentioned above, there will be other offence descriptions for which the corresponding percentage will be higher than the overall figure of 29.8%: there will be variation.

50. Assessing the statistical significance of these differences is difficult, because of the lack of independence. A relevant and valid comparison is the one I previously obtained in paragraphs 125 to 127 of my first report, where it is shown that the rate of any person checks in file 3 among African/Middle Eastern "offenders" from file 5 (96%) is markedly higher than the rate among "Other" males (79%), and the difference is strongly statistically significant ($P = 0.005$). Dr Henstridge agreed with this result, at paragraphs 28 and 29 of JH1.

51. I have made all the inquiries which I believe are desirable and appropriate and no matters of significance which I regard as relevant have, to my knowledge, been withheld from the Court.