“Everything is excellent”: methodological issues in studying refugee settlement

Dr Farida Tilbury
School of Social Sciences & Humanities
Murdoch University
f.tilbury@murdoch.edu.au

Abstract:
Cross cultural research presents unique challenges to the researcher, more so when it involves a number of different cultural groups, and when issues of marginality are at stake, as in the case of research with refugees. Using 6 bilingual research assistants to survey 150 refugees from 19 different national backgrounds and three different regions of the world, this research project examined the links between degrees of ‘visible difference’ and employment outcomes in an Australian context. Challenges faced included training bilingual assistants to do social research; sampling issues; cultural differences in sensitivity to particular types and formulations of questions, producing variable results; and how to deal with anomalous results. The paper argues that for the sake of improving the validity of refugee research, researchers should engage in reflexive consideration of methodological issues publicly in forums such as this.

This paper describes some of the methodological challenges encountered in a study of refugees in Western Australia, which affect the study’s validity, generalisability and ethics. Jacobsen and Landau (2003) argue that those engaged in research about refugees have an ethical imperative to do research which upholds high standards of academic rigour, while also generating knowledge that will be useful to protect and improve the lot of refugees. In order to ensure this, they argue two principles must be met: first, analysis and conclusions must be based on sound principles of inference and data collection practices; and second, research must be conducted in an ethical way. Therefore, they encourage researchers to be explicit about the methods used and challenges faced. Key challenges for any research about refugee issues include questions of access, language barriers, sampling difficulties, funding constraints, and a bevy of ethical issues including relevance, possibilities for doing harm, confidentiality, consent etc. However, as
Jacobsen and Landau argue, refugee studies generally demonstrate a “lack of rigorous conceptualization and research design, weak methods, and a general failure to address ethical problems related to researching vulnerable communities” (2003:2). They go on to suggest that “the considerable logistical challenges facing researchers does not justify ad hoc research design, obfuscation, or exaggerated claims” (2003:2). This paper seeks to address some of these points.

The Research

The study which forms the basis of this paper sought to ascertain the degree to which, if any, poor employment outcomes for refugees, which stand at 71% six months after arrival, and 43% eighteen months after arrival (DIMIA, 2005), could be attributed to racism and discrimination in the West Australian job market. It targeted those ‘visibly different’ to varying degrees (through accent, name, physical features, religion and culture) from the mainstream population and who were the main intake under the humanitarian program during the 1990s and early 2000s (DIMIA, 2005).

Six bilingual research assistants from African, Middle Eastern and former Yugoslav backgrounds conducted face-to-face questionnaire-based interviews using a snowball sampling technique with 25 respondents each (N=150), according to the following criteria: respondents had to be people who had sought work, and had high education or skills, and literacy levels, from one of these regions. The sample was designed to be indicative rather than representative, focussing on refugees with high levels of human capital, eliminating the lack thereof as an explanation for negative employment outcomes.

The questionnaire consisted of multiple-choice, Likert-type scale and several open-ended questions. Interviewers conducted the surveys in English, although they were able to translate where required, and recorded answers in English. The use of a survey rather than in-depth interviews, which may have been more culturally appropriate and provided more valid data, was a cost-cutting measure, as translations/interpreting, transcription and back-translations of lengthy interviews add considerably to the cost of research, making funders baulk at the size of the data collection budget. Since the sample was selected for its high human capital it was felt that language would not be a considerable barrier to obtaining valid data. The standardised questionnaire was developed by the CIs, and bilingual assistants
provided feedback on it and pilot tested it. Bilingual research assistants were provided with training in how to conduct questionnaire interviews (as per Babbie, 2005: 274-278).

The questionnaire (available on request from the author) covered demographics, job-seeking strategies and job-market experiences, expectations, experiences of service provision, perceptions of barriers to employment and a number of measures of resettlement success including life satisfaction.

The findings indicate that despite high levels of work relevant skills, refugees experience both personal and institutional discrimination in the employment market (Colic-Peisker and Tilbury, 2006a, 2006b; Tilbury and Colic-Peisker, 2006), resulting in pockets of ‘niche’ employment, low income levels, lack of opportunity, and perceptions of discrimination, and, from the Australian economy’s point of view, a massive waste of human capital.

However, it is difficult to make definitive conclusions from the research due to certain anomalies in the data collection process.

**Sampling and survey completion**

In order to ensure validity, reliability and generalisability, and ensure maximal explanatory value, the ideal method for sampling for any research is random selection of a representative sample, together with equivalent measures using a control group (Jacobsen and Landau, 2003; Babbie, 2005; De Vaus, 2003; King, Keohane and Verba, 1994). Few studies of refugees use this technique due to the difficulty of identifying a clear sampling frame i.e. a list of all members (potential participants) of the group of interest – in our case, refugees living in Perth from Middle Eastern, ex-Yugoslav and African countries; and the added cost of finding a similar control group. Accessing up-to-date data on refugee populations has ethical and practical implications which make such an approach difficult. Thus a snowball sample, commonly used in refugee research because of these difficulties, was used.

However, the result was a clearly non-representative, and non-comparable (across groups), sample. Bilingual assistants tended to snowball from people like themselves, despite instructions to seek a wide range of participants. Thus the samples were skewed.
to those similar in gender, age, marital status and country of origin to the interviewers\textsuperscript{2}. For example the two female ex-Yugoslav assistants, selected samples predominantly like themselves, middle-aged partially-employed married females from Bosnia, who had been in Australia for over 5 years. The result was that two thirds of the ex-Yugoslav sample were female, whereas two thirds of the African and Middle Eastern samples were male (perhaps reflecting assumptions about gendered breadwinning or the invisibility of women among these communities). The other groups were also younger on average, and had higher education levels, and had been in Australia fewer years. The result is difficulty of comparison of data across groups, problems of interpretation in terms of attributing differences in findings to real differences between the communities or sampling artefacts, and therefore a lack of generalisability. We can also make no conclusions about income, employment levels, levels of English competence, etc among these communities from our samples. Therefore any findings are necessarily specific to our sample. However, norms of journal and report writing, which encourage definite statements and discourage qualification and tentativeness, mean that we tend to make exaggerated claims for the generalisability of the findings from this sample to all refugees, something Jacobsen and Landau (2003) argue is common among refugee research reporting.

We had hoped to avoid differences in comprehension levels, and ensure high quality responses, by asking the research assistants to complete the survey on behalf of respondents, and to translate where respondents had difficulty understanding the questions\textsuperscript{3}. However, the approach to collecting the data also varied between assistants. When we received the responses back from one set of the Middle Eastern sample, predominantly Iraqis, many were incorrectly completed and the language level made it very difficult to understand some responses. When asked why this was the case, the assistant told us that she could not insist on completing the survey form herself, on behalf of the respondents, as instructed, because Iraqis felt patronised if not allowed to do it themselves, particularly since they felt their levels of English were adequate to the task. For her to insist on filling it in for them would have been insulting. The quality (and comparability) of the data was affected as a result. Analysis also revealed that assistants had interpreted questions slightly differently, resulting in different types of responses. For example, when asked about whether respondents worked at, above or below their
qualification level, some assistants filled in ‘below’ for respondents who were not employed, whereas others left the answer blank.

We also discovered differences in completion of the qualitative sections of the survey – Africans and Middle Easterners produced slightly fewer and shorter qualitative responses than ex-Yugoslavs, despite ex-Yugoslavs’ self-reported English being at lower levels than the other groups’. It is unclear whether this was a researcher effect, a language effect, or a cultural effect. Certainly the fact that one researcher let participants fill in the questionnaire themselves may have lowered the rate of Middle Eastern qualitative responses. In another question, a number of Africans, and a few Middle Easterners, ticked the ‘no’ option when asked whether they had experienced discrimination in the job market, but added comments to qualify this answer – saying they felt it was discrimination but didn’t have ‘evidence’ that it was.

A final issue with the sampling was the use of ‘insider’ research assistants. The use of insiders is argued to improve validity due to assumptions about ‘insider knowledge’ and automatic rapport (see Wilkinson and Kitzinger, 1996, for an extended analysis). From a community development perspective, we had also argued its positive effects in providing short term employment and training (human capital development) for members of marginalised refugee communities. However, using insiders can also cause confidentiality problems in small communities, and produce biased results as members of the same community ‘face manage’ their responses eg. about income levels, well-being, employment experiences, job seeking strategies, qualifications etc. (Jacobsen and Landau, 2003; Wilkinson and Kitzinger, 1996). Certainly we believe some of the missing data was due to assistants feeling uncomfortable about asking some of the questions, as well as respondents feeling uncomfortable about answering them. It may also be that due to political (or social, religious or economic) sensitivities between different nations, and indeed ethnic communities, from within the three regions, that the assistants only recruited participants from groups with whom they felt comfortable, meaning some groups were excluded from participation (see endnote ii). It may be that an outsider, with no assumed political, national or ethnic affiliation, would have produced a more representative sample and more valid data. One might legitimately ask whether the value of trying to use community members as a sign of respect and to give something back to
the community is worth these resulting concerns about the quality and representativeness of the data.

**Anomalous findings and questions of validity**

The rest of the paper explores two interesting findings and their possible relationship to methodological issues. One illustrates the differences between qualitative and quantitative data and questions of validity, the other is concerned with the effects of culture on response patterns, which also goes to issue of validity.

The author was prompted to undertake this project as a result of earlier qualitative research into ‘mental health’ among the same refugee populations, where racism and lack of employment were identified as two of the most important factors associated with negative well-being and mental illness such as ‘depression’ (see Tilbury, under review). However, in this project we found that quantitative measures of well-being were relatively high – despite only 43% being employed full-time (and 28% unemployed), and despite almost half the respondents indicating that they had been discriminated against generally and also in the employment market specifically⁵, respondents reported generally positive quality of life (see Tilbury and Torezani, under review, for an extended discussion of these findings). On a four point scale, over three quarters reported being entirely or mostly satisfied with their lives (76.7%). Our respondents also generally reported that life was back to normal, indicating a return to well-being for those in resettlement situations, with 72.6% agreeing entirely or mostly. To produce a ‘refugee satisfaction score’, the two items (general satisfaction and life being back to normal) were aggregated. On a 4 point scale the mean hovered around the ‘mostly’ mark (2.94).

Multivariate analysis produced further surprising results. Very few life satisfaction factors were predicted by experiences of discrimination (from the question ‘have you been treated in an unpleasant way as a refugee’). Discrimination had no statistically significant effect on perceptions that Australia was a fair country (.026), that it was difficult to be a refugee in Australia (.146), that Australia was ‘home’ (.034), regret in coming to Australia (.130), and most surprisingly ‘life satisfaction’ (.088)⁶. Elsewhere we have suggested possible reasons for the lack of correlation between discrimination and well-being (Tilbury and Torezani, under review) and between employment and well-
being (no correlation for the African and Middle Eastern samples, and only slight correlation for the ex-Yugoslavs, see Colic-Peisker, under review), but these arguments assume the result is ‘real’ i.e. based on valid findings from statistics derived from a representative sample. Here I would like to challenge the rationality of undertaking such statistical manipulations on such a sample.

Qualitative data from the research clearly connects well-being with both employment (positively) and discrimination (negatively). A feature of the questionnaires completed by the participants was the individual stories of discrimination provided in response to the open-ended questions. A simple content analysis revealed that of 150 respondents 78 commented that they had been discriminated against in the workplace and/or while looking for work (spread evenly across the three groups). Stories of work and personal discrimination included examples of personal abuse, rejection, stereotyping, and various forms of injustice. These experiences were framed as having a negative impact on their lives, causing frustration, feelings of alienation, and loss of hope. While ex-Yugoslavs rationalised their difficulties, Africans were far more likely to attribute difficulties to discrimination. The Middle Eastern sample generally were much more cautious about attributing their difficulties to racism – they tended to dress it up, using the sorts of rationales that they had received in rejection letters – lack of work experience and local referees, the ‘competitive market’ etc.

Thus we have a contradiction between the quantitative data, and the statistical relationships (or lack thereof) found in it, and the qualitative data which reflects respondents’ perceptions of their condition and its causes. Certainly the qualitative results indicate a fairly high level of bitterness among refugees.

A second feature was the apparent cultural differences in responses. For all groups it is likely that there was what is known as a ‘social desirability response’ effect (Babbie, 2005) due to the face-to-face nature of the interview process, compounded by the use of ‘insider’ researchers, and cultural imperatives to present a positive face. This would have affected questions about levels of English language proficiency, which were self-assessed (objective measures would have increased the cost and time of the research, and its viability); financial status (fortnightly income) which may have been underestimated, to
not appear too successful, or overestimated, to not appear unsuccessful; and the well-being measures.

Likewise, for the entire sample, the gratitude factor which sees refugees grateful to the host country for their refuge (Richardson et al., 2004; Tilbury, 2006) may have produced more positive responses.

The Middle Eastern respondents appear to have provided more positive responses to certain questions than the others, and to have selected the superlative category eg. ‘excellent’ rather than ‘good’, ‘always’ rather than ‘mostly’ etc (an effect which is lost if the top two categories are collapsed when reporting). It appears they brought cultural rules of politeness to the research process. For example, when asked about their experience of the job-seeking support services (Job Network), respondents provided massively different appraisals (See Table 1).

**Table 1. RATINGS OF EMPLOYMENT SERVICES USED**

<table>
<thead>
<tr>
<th></th>
<th>Yugoslav</th>
<th>African</th>
<th>MidEast</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>2.0</td>
<td>2.0</td>
<td>36.0</td>
<td>13.3</td>
</tr>
<tr>
<td>Good</td>
<td>6.0</td>
<td>8.0</td>
<td>16.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Acceptable</td>
<td>10.0</td>
<td>12.0</td>
<td>8.0</td>
<td>9.3</td>
</tr>
<tr>
<td>Poor</td>
<td>32.0</td>
<td>20.0</td>
<td>8.0</td>
<td>20.0</td>
</tr>
<tr>
<td>No answer/n/a</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

There were also massive differences in responses to other questions about service delivery. Only a small proportion of such differences could be explained by actual differences in services (given that participants were using the same service providers), therefore these findings indicate the propensity of those from the Middle East to rate their experiences, particularly where those experiences involve an evaluation of others, more highly than those from Africa or the former Yugoslavia.

We have further evidence of this cultural ‘politeness imperative’ which sees those from some cultures not openly wanting to criticise others. One of the Iraqi assistants wrote the following comment on a survey “This participant feels strongly that there is kind of invisible discrimination, but he didn’t want to mark or even comment”. In the debrief session another of the Middle Eastern assistants, who works with Iraqi settlement, gave the following insight:
I did not even mention the word ‘research’. They have to answer everything positively, that’s what they are used to: ‘everything is excellent’. They are afraid if they criticise Centrelink they will somehow find out. Iraqis do not want to criticise anyone. They think if you’re seen as positive, it keeps the relation alright. Negative means [having] a critical position. …If I ask them about employment but they talk about general things. And they would never say anything against the government.

**Conclusion and Policy implications**

So far it appears I have heaped a litany of criticism on a project with which I have been involved. I have done this for the purpose of what Robson (2002) refers to as ‘establishing trustworthiness’ – out of a desire not to claim more for the findings than can or should be claimed, and to identify some interesting methodological effects in the data, which may be worthy of further exploration. As such I do not wish to dismiss the research undertaken, but to clarify methodological shortcomings, which make broad statements of generalised findings inappropriate. Journal articles do not encourage tentative statements, qualification, or criticism as part of the reporting process, so I have taken this opportunity to draw out some of the challenges of this particular project.

One of the main issues is quantitative versus qualitative approaches. Here we have used a sampling technique appropriate for qualitative research, to collect quantitative data, which was then analysed using statistical techniques. The result is often an attempt to make the data do more than it can do. Another is the assumption that blunt 4 point scales, for questions which contain sophisticated language and concepts, will produce valid results across cultures and languages. The result is uncertainty as to whether findings reflect real differences or are cultural artefacts. It is also likely that responses were selected by respondents through a process of ‘satisficing’ rather than optimizing, known to occur where the research task is difficult in terms of ‘meaning making’; and where respondents’ ability and motivation to answer are low (Krosnick, 1999). Certainly, qualitative data indicate a relationship between variables which quantitative statistics did not identify. It would be easy to suggest that these issues should have become evident and been dealt with in the preliminary meetings and pilot phase, but they either did not arise, and were thus overlooked, or pragmatic (or financial) decisions were made to ignore them, to the detriment of the quality of the data. Even Jacobsen and Landau (2003) who are quite scathing in their critique of the sloppiness of refugee research, found the need
for compromise due to the practicalities of on-the-ground circumstances, in their own research with refugees in Johannesburg. However, they conclude that “from an academic and advocacy perspective, the benefits of rigorous methods in refugee research outweigh the costs” (2003:18) because research well done can be a powerful tool for policy-makers - which leads me to the final issue of concern. Some of the quantitative findings may lend support to arguments that since neither lack of employment opportunities nor experiences of discrimination have a strongly negative impact on general well-being of refugees, no policy action should be taken to improve either of these social realities. For policy-makers to come to such a conclusion, given the shortcomings outlined, would be a serious tragedy.

Footnotes

1 The research was funded by an Australian Research Council Discovery Grant. Chief investigators on the project, entitled ‘Refugees and employment: the effect of visible difference on discrimination’, are Val Colic-Peisker, Farida Tilbury and Nonja Peters. The three year project runs from 2004-2006, and data was collected in 2004. Two other papers from the project (Torezani et al, Colic-Peisker) are presented in these proceedings.

2 Consequently, the samples varied as follows: Age (ex-Yugoslav average age 44 yrs; African and Middle-Eastern (37 and 38 yrs respectively); Gender (ex-Yugoslavs female 58%; Africans and Middle-Easterners 28% each); Education (32% ex-Yugoslavs university-educated; Africans and Middle-Easterners 68% and 66% respectively university-educated, often with postgraduate degrees – 26% and 20%); self-assessed written English language proficiency (ex-Yugoslavs 66% very good or fluent; Africans 68%, Middle Easterners 88%) and self-assessed English good enough to do the same job as before coming to Australia (Ex-Yugoslavs 52%; Africans 84% and Middle Easterners 80%). Thus the African and Middle Eastern samples were far more similar, more educated, and with better self-assessed English, than the ex-Yugoslavs. Two thirds of the ex-Yugoslav sample were from Bosnia, over three quarters of the Middle East sample from Iraq, over half the African sample were from Somalia.

3 Initially we planned to translate the questionnaires. This would have added significantly to the costs of the research (adding perhaps $20,000) due to the need for translation into a number of different languages (there are no standard languages for each of the three regions targeted), back translation, and then translation of responses into English. Thus the pragmatic decision was made to rely on the bilingual assistants to do ‘on-the-spot’ interpreting where needed. This most likely limited the sample as assistants would only select participants with whom they shared a language, and may have affected the quality of the data collected due to interpretation issues.

4 The question followed one which asked ‘what is your job (if employed)’ and by implication, should not have been filled in if participants were not employed.

5 Discrimination in the job-market was reportedly due to differences of language ability, accent (more common for the ex-Yugoslav sample), religious customs, name and appearance (more common among the African and Middle Eastern samples).

6 The figures are Pearson correlation scores (r) indicating relationships between general discrimination and related variables. Note: * = significance at p ≤ .05 level; ** = significance at p ≤ .01 level; *** = significance at p = .000 level

7 For ex-Yugoslavs the discrimination was framed as being mainly related to lack of English and local qualifications, for Africans it was framed more often as religious, as well as about lack of local qualifications, and for Middle Easterners it was framed mainly the lack of local qualifications and
experience (despite them reporting a much higher rate of full qualification recognition at 62%, compared to 2% for Africans, and 44% for ex-Yugoslavs) but secondarily as a result of religion or Middle-Eastern appearance.

References


Tilbury, F and Torezani, S, “Discrimination and well-being: perceptions of refugees in Western Australia”, under review for *International Migration Review*.