

WHASS Workers' Standalone:

Technical Report

January 2006

Prepared for: **Health & Safety Executive**

Prepared by: **BMRB Social Research**

Part of BMRB Limited

BMRB/GK/NG/KW/45103690

Quality System Registered by BSI

BSI Certificate of Registration No: FS 28348

Table of Contents

1.	SUMMARY	1
1.1	Data Collection method	1
1.2	Sample	1
1.3	Respondent Selection.....	1
1.4	Worker Interview	1
2.	BACKGROUND AND INTRODUCTION	2
3.	SURVEY METHOD	3
4.	SAMPLE	4
5.	FIELDWORK PROCEDURES.....	6
5.1	Briefing of Interviewers	6
5.2	Fieldwork period.....	6
5.3	Supervision and Quality Control.....	6
5.4	Partial Interviews.....	6
6.	SURVEY RESPONSE RATES.....	8
6.1	Calculating response rates on RDD telephone surveys	8
6.2	Target response rate.....	8
6.3	Actual response rate.....	9
6.4	Adjusted response rate	10
7.	WEIGHTING PROCEDURES.....	14
7.1	Weighting.....	14
7.2	Design Effects	15

1. SUMMARY

This report provides the technical background to the **workers'** standalone survey.

In summary:

1.1 Data Collection method

The survey was a telephone survey of private households in Great Britain that contain one or more adult aged 16 or over who had worked in the previous 12 months.

1.2 Sample

Random digit dial (RDD) sampling was used to provide a sample of the residential population, with both listed and unlisted phone numbers represented, drawn from all eligible exchange codes. Interviewers then attempted to phone each issued telephone number and checked eligibility at all those that were answered

1.3 Respondent Selection

The 'Rizzo' method, a computerised method of random selection was used to select eligible respondents. This survey was the second survey ever in the UK to use this method.

1.4 Worker Interview

The mean interview length for this survey was 30 minutes. The response rate was 26%.

2. BACKGROUND AND INTRODUCTION

The Health and Safety Executive (HSE) is planning to conduct a programme of large-scale workplace surveys during the period 2005-2015, in order to study the state of health and safety in Britain's workplaces. The first survey was expected to start in late 2005 (referred to as the 'main survey'). This survey is intended to take a leading role among the range of sources used by HSE to assess the progress towards government targets set for health and safety at work.

The intended approach to the survey is to select a sample of workplaces and conduct interviews with the employer (the person responsible for health and safety) and provide self completion questionnaires to be given to a sample of their employees. Those workers responding to the survey and providing permission for recontact will then form the sample available for a longer questionnaire to be administered by telephone. In this way robust data on experiences and attitudes will be gathered from both employers and their workers, enabling thorough analysis within the workplace.

In advance of the main survey, BMRB has been commissioned to conduct a 'Dress Rehearsal' of the main survey to test this innovative and challenging approach. Separate reports (process and technical reports) have been produced describing BMRB's experience of the Dress Rehearsal. BMRB has also been commissioned to conduct standalone surveys of employers and employees in order to test the questionnaires and to provide baseline data in advance of the main survey.

This report provides the technical background to the employees' standalone survey.

3. SURVEY METHOD

This can be summarised as follows:

- The sample was produced using the Random Digit Dialling (RDD) technique
- All selected numbers were “pinged” to remove numbers not working
- Remaining numbers were checked against the electronic Yellow Pages (www.yell.com) to remove listed business numbers; the sample was then issued to the interviewers
- All numbers were rung and, where answered, interviewers checked the eligibility of the household for the survey
- CATI interview with randomly selected eligible respondent

4. SAMPLE

The survey used a random sample design. The sampling frame used was the one held by Survey Sampling International (SSI), an established company dedicated to the provision of survey samples. Random digit dial (RDD) sampling, developed by SSI, was used to provide a sample of the residential population, with both listed and unlisted phone numbers represented, drawn from all eligible exchange codes.

All UK residential telephone numbers comprise 10 digits, excluding the leading zero. The first six of these digits are the exchange or 'block'. SSI's sampling frame is a database containing all 'blocks' allocated for all residential telephone service providers in the UK. By adding all possible combinations of the final four digits, SSI is able to create a sampling frame of all possible residential numbers within a given geographical area (for this survey it was Great Britain).

The sample selection process had a number of stages:

1. SSI selected a systematic random sample of numbers, with equal probability of selection, from the list of all possible 10 digit numbers in Great Britain exchange areas, which were initially sorted to ensure a correct geographical spread of numbers; the numbers selected were set so as to ensure that, after making due allowance for ineligible numbers and non-response, the target number of interviews could be achieved;
2. SSI then screened all selected numbers (using the "pinging" method) and removed those found to be not working;
3. SSI then checked the remaining numbers against the electronic Yellow Pages (yell.com) in order to remove listed business numbers, though it should be noted that unlisted business numbers remain in the sample; the remaining numbers were allocated to interviewers for fieldwork;
4. Interviewers then attempted to phone each issued telephone number and checked eligibility at all those that were answered (i.e. that the address was a private household in Great Britain and contained one or more adult aged 16 or over who had worked in the previous 12 months).
5. Having made contact with an eligible household, interviewers selected one adult for interview, where more than one was eligible, using the 'Rizzo' selection method.

The 'Rizzo' method is a new selection method first used on the Scottish Crime and Victimization Survey by BMRB, and this survey was the second survey ever in the UK to use this method. In most RDD surveys using random sampling, the 'last birthday' rule has been used as the method for selecting one respondent to take part in an interview where a household contains more than one eligible person. However, concerns have been raised about both the

reliability of this method and the impact that it has on co-operation rates. Firstly, there is an apparent tendency (which has been noted elsewhere¹) for the person who answers the 'phone to be selected more often than would be expected by chance. Secondly, the last birthday rule (or next birthday rule, for that matter) is cumbersome to administer on the telephone, and can appear intrusive, particularly at the start of a call, when the interviewer has not had a chance to establish rapport. It is not uncommon, therefore, for respondents to drop out at this point in the interviewing process.

The premise for the 'Rizzo' method is that the large majority of households contain only one or two eligible adults. In most scenarios, therefore, either there will be one eligible adult in the household, in which case there is no selection to be made, or there will be two. Where there are two adults, Rizzo's approach is simple: one of the two adults is selected by a random procedure (the Kish method²), which can be programmed into the CATI script. Each of the two adults has an equal (50 per cent) chance of selection. In these cases, the only information that the interviewer needs to establish in order for the selection to be made is how many eligible adults live in the household. Where there are three or more eligible adults in the household, the CATI programme initially selects the person answering the 'phone³ as the respondent (this person is allocated a probability of selection of $1/N$, where N is the number of eligible adults) or one of the other adults. If the latter, there needs to be a second stage of selection, which will require the interviewer to list the other adults (*eg* using their first name), so that the CATI programme can again select one at random⁴. Again, this approach ensures that all eligible adults have an equal probability of selection, provided that the information provided by the first respondent is accurate.

An initial sample of 97,540 numbers was provided. Because response rates were lower than had been anticipated, it was necessary to select and issue additional sample during the course of fieldwork. An additional 24,426 numbers were provided during the course of fieldwork, resulting in a total issued sample of 121,966 numbers.

¹ Rizzo L, Brick J, and Park I (2004), 'A minimally intrusive method for sampling persons in random digit dial surveys', *Public Opinion Quarterly*, 68(2), 267-274.

² Kish L (1949), 'A procedure for objective respondent selection within the household', *Journal of the American Statistical Association*, 44, 380-387.

³ Or who comes to the 'phone if it is initially answered by a child or ineligible adult.

⁴ Rizzo *et al.* suggest that the last birthday rule could be used at this stage, but this seems a less satisfactory solution, since it means that the problems associated with this approach remain to some extent.

5. FIELDWORK PROCEDURES

All fieldwork was carried out by fully-trained interviewers from the telephone division of The Operations Centre, which provides operational services to the Kantar Group of research companies in the UK, of which BMRB is part. Fieldwork was conducted in the TOC telephone centre in Ealing.

5.1 Briefing of Interviewers

Before starting work on the survey all interviewers received a personal briefing from the research team, and received a set of written instructions. The briefings covered:

- The background to the survey, the role of the Health and Safety Executive and how they plan to use the survey findings;
- Information about sampling and screening procedures and the importance of high response rates;
- How to introduce the survey and deal with sensitivities that might arise;
- A run-through of the questionnaire in its entirety, including detailed explanation of certain key questions and topics covered;
- Any questions the interviewers or supervisors had for the researchers.

5.2 Fieldwork period

Main stage fieldwork took place between 23rd August 2005 and 13th December 2005. Originally fieldwork was due to be completed by 5th November. However, because the response rate on the survey was lower than anticipated, it was decided to extend the fieldwork. By the end of fieldwork a total 10,016 interviews had been achieved, at a response rate of 26%. The average interview length was approximately 30 minutes and 20 seconds.

5.3 Supervision and Quality Control

Throughout the fieldwork, measures were taken to ensure that the standard of interviewing was appropriate. At the start interviews were monitored by the researchers as well as a representative from the HSE. In addition, researchers monitored interviews at various points during the main stage fieldwork to check for potential problems and ensure that the standard of interviewing remained high. Telephone unit supervisors regularly monitored the interviewers and relayed any issues or concerns back to the researchers. Our quality control and supervision standards exceed those stipulated by the market research industry's Interviewer Quality Control Scheme (IQCS).

5.4 Partial Interviews

It was recognised that the anticipated interview length of 30 minutes is very long for a telephone survey of the general public. Therefore provision was made in the survey for "partial interviews" to count towards the total number of achieved interviews. The cut-off for a "partial

interview” was set at a point in the questionnaire where the most valuable data had been gathered and most of the questions had been answered. In this survey the cut-off for a “partial interview” was after the climate tool questions (after qwcl). In total out of the 10,016 achieved interviews, 9,168 were full interviews (92%).

6. SURVEY RESPONSE RATES

6.1 Calculating response rates on RDD telephone surveys

The calculation of response rates for RDD telephone surveys is not as straightforward as for face to face surveys, or telephone surveys with named respondents, because the denominator for the calculation – the so-called ‘in scope’ sample – cannot be verified. Numbers that remain unanswered after many calls may or may not be valid residential numbers – for example, they may be business lines not staffed during the survey fieldwork times, public telephone boxes, second homes which are only occupied for short periods, and so on. Although there are no accepted estimates for the proportion of unanswered numbers that will be of this type, methodological research on a small scale has shown that a high proportion are ineligible.

Response rates for RDD surveys are usually based, therefore, on telephone numbers that have answered and been positively identified as eligible private households, with non-contacts excluded as out of scope. This is the basis on which response rates have been calculated for this study.

It should be noted that some uncertainty inevitably remains as to whether *all* persistently non-answering lines are in fact ineligible. If they are not, ‘true’ response rates would be lower than those reported.

6.2 Target response rate

It is well documented that response rates for RDD telephone surveys are generally considerably lower than those that can be expected from random face-to-face surveys. Respondents are, in effect, being ‘cold-called’ and, unlike face-to-face surveys, have no immediate means of checking on the credentials of the caller. It is also much easier, culturally, for people to refuse over the ‘phone, than it is to do so to an interviewer on the doorstep. An instant refusal in the first few seconds of telephone contact is often the result, and refusal rates of 30-50% are typical for RDD telephone surveys among the general public, even for surveys introduced as being on behalf of government clients. A further 3-4% of interviews are also commonly abandoned before completion, even with quite short interviews.

BMRB has carried out many telephone surveys with RDD samples that require response rates to be maximised. Our experience with these surveys is that a response rate of as high as 50% (after excluding ‘no replies’ from the base) requires considerable persistence, skilled interviewing and a number of special methodological techniques. The surveys we have conducted include several that were introduced as being about crime (a high interest subject) and with relatively short interview lengths of 15-20 minutes.

Surveys of less potential interest than crime to respondents generally achieve lower response rates. In estimating the response rate that might be achievable in this survey, a number of

factors needed to be taken into account, including the sponsor and subject matter of the survey, the interview length and the time available for fieldwork.

It was felt that the survey subject matter would be hard to 'sell', certainly when compared with a subject such as crime. The concept of 'health and safety at work' is much less salient in people's minds than subjects such as crime – only 10% of the working population, according to the Labour Force Survey have suffered a work-related illness or injury in the last year. In addition to this, the average interview length was anticipated to be around 25 minutes, which is considered long for a telephone survey.

Taking into account these factors and experience on other surveys, BMRB estimated that a response rate of 30% should be achievable, but with the caveat that additional sample be held in reserve in the event that this target could not be realised.

6.3 Actual response rate

In practice, the survey proved considerably more difficult to sell to respondents. A number of steps were taken in an effort to improve the situation. These included reducing the size of the interviewing team, so that interviewing was conducted by a relatively small number of high performing, experienced interviewers; increasing the fieldwork period by a number of weeks; making changes to the survey introduction to reflect interviewer feedback on what was and was not working in introducing the survey; and setting up a 'refusal survey', whereby experienced interviewers re-contact 'soft' refusals and attempt to convert them to interviews.

With this range of measures in place, the response rate at the time fieldwork was closed on 13th December rose to 26%, a significant improvement on the response rate in the early stages of fieldwork, but still lower than the rate originally anticipated (30%). Although non-response weights can be applied to the survey data to correct for biases in the demographic profile of the achieved sample, the likelihood of other unmeasurable biases occurring as a result of non-response increases with the level of non-response. The low response rate therefore needs to be borne in mind when interpreting the survey data.

A detailed analysis of field outcomes is overleaf:

Sample issued		121,966		
Out of scope		80,697	66%	
Deadwood		65,469		
	Business number	29,048		
	Fax/Computer Line	4,046		
	Unobtainable	1,745		
	Duplicate number	198		
	Dialler – bad number syntax / incomplete	4,879		
	Dialler – site out of service	580		
	Dialler – incomplete number	60		
	Dialler – unknown error	0		
	Dialler – modem tone	148		
	Dialler – other dialler errors	0		
	No contact after 20+ calls	24,765		
Ineligible		15,228		
	Ineligible – no one in household worked in last 12 months	14,569		
	Ineligible – not a private household	659		
In scope		41,269	34%	100%
	Refusals:	26,737	22%	65%
	Abandoned interview	1,183		
	Proxy refusal	588		
	Soft refusal	17,940		
	Hard refusal	7,026		
	Other non-productive outcomes	4,516	4%	11%
	Appointment	15		
	General call back	520		
	No answer/answering machine	640		
	Engaged	18		
	Dialler hangup	0		
	Stopped interview	19		
	Respondent has hearing difficulties	357		
	Respondent has language difficulties etc.	446		
	Respondent incapable of interview	1,000		
	Respondent unavailable during fieldwork	82		
	Max level of unsuccessful calls (contact made)	1,419		
	Interviews	10,016	8%	24%
	Full	9,168		
	Partial	848		

6.4 Adjusted response rate

In this survey, the survey population is those households where at least one member aged 16 or over has worked in the last twelve months. Hence the simple response rate of 24% above is not the real response rate, as it does not take into account that not everyone in the population is eligible to take part.

Hence it is necessary to make an adjustment to try and estimate the true response rate.

The nature of a CATI survey using a Random Digit Dialling technique is that the majority of those refusing to take part in the survey will do so in the first few seconds of contact, before eligibility can be established by the interviewer. Thus it is necessary to apply some kind of estimate of eligibility to those outcomes where eligibility has not been established. From this an estimated “in scope” figure can be derived, and the response rate calculated.

The first stage is to disregard deadwood sample, which is excluded from response rate calculations. Those respondents who are ineligible because they are not in a private household are also excluded. This is because the eligibility calculations are based on data from the Labour Force Survey (LFS) which is a survey of private households.

The decision on who to include in the non-response sample is important as it affects the numbers in the in-scope sample and therefore the final response rate. There are several categories of respondent where the decision as to whether to include them is debatable:

- Respondents with hearing difficulties;
- Respondents with language difficulties, or mental or physical disabilities;
- Respondents otherwise incapable of interview;
- Numbers where the maximum number of calls has been made, no contact has been established, but the number has on occasion been engaged, indicating that the line is occasionally in use.

The first three categories of respondent are effectively excluded only by virtue of the survey design and decision to use a CATI data collection method – it is conceivable that had we used different data collection instruments e.g. a paper self-completion questionnaire or in a face-to-face survey they could have been included. Therefore we have included these respondents in the non-response sample.

For numbers where we have not been able to establish contact but the number has on occasion been engaged, we propose not to include them in the non-response survey. Many of these numbers in reality are probably public phone boxes, fax or modem numbers. Given the procedures adopted in this survey (staggering calls over all days of the week and at different times, with a maximum of 15 calls made to each number where there is no contact with a person) we feel if no contact is made on any number then a deadwood classification is appropriate.

The second stage is to divide the non-responding sample into three parts: eligible, ineligible and unknown eligibility. The table below shows how we have done this:

<i>Non-response sample</i>				
Eligible		11,806		
	Stopped interview	19		
	Abandoned interview	1,183		
	Proxy refusal	588		
	Interviews	10,016		
Ineligible		14,569		
	Ineligible – no one in household worked in last 12 months	14,569		
Unknown eligibility		29,463		
	Appointment	15		
	General call back	520		
	No answer/answering machine	640		
	Engaged	18		
	Dialler hangup	0		
	Respondent has hearing difficulties	357		
	Respondent has language difficulties etc.	446		
	Respondent incapable of interview	1,000		
	Respondent unavailable during fieldwork	82		
	Max level of unsuccessful calls (contact made)	1,419		
	Soft refusal	17,940		
	Hard refusal	7,026		

There are 55,838 outcomes in our non-response sample. LFS 2003 data shows that 68.4% of households contain one or more adults aged 16+ who have worked in the last 12 months. Therefore 38,193 of our non-response sample can justifiably be thought to be eligible to take part in the survey. Of the non-response sample 11,806 have been established as eligible to take part in the survey. Therefore we can estimate that of the 29,463 non-response outcomes where we were not able to establish eligibility, 26,387 must have been eligible (89.56%).

Taking this estimated eligibility rate and applying it to non-response categories where eligibility is unknown allows us to calculate the in-scope sample as follows:

<i>Non-response sample</i>			<i>Eligibility</i>	<i>Eligible in - scope</i>
Eligible		11,806		11,806
	Stopped interview	19	100%	19
	Abandoned interview	1,183	100%	1,183
	Proxy refusal	588	100%	588
	Interviews	10,016	100%	10,016
Unknown eligibility		29,463		26,387
	Appointment	15	89.56%	13
	General call back	520	89.56%	466
	No answer/answering machine	640	89.56%	573
	Engaged	18	89.56%	16
	Dialler hangup	0	89.56%	0
	Respondent has hearing difficulties	357	89.56%	320
	Respondent has language difficulties etc.	446	89.56%	399
	Respondent incapable of interview	1,000	89.56%	896
	Respondent unavailable during fieldwork	82	89.56%	73
	Max level of unsuccessful calls (contact made)	1,419	89.56%	1,271
	Soft refusal	17,940	89.56%	16,067
	Hard refusal	7,026	89.56%	6,293

Therefore there are 38,193 outcomes which are “in scope” out of which 10,016 interviews were achieved, producing an estimated final response rate of 26.22%.

7. WEIGHTING PROCEDURES

7.1 Weighting

The weighting procedure was done in stages – first a design weight was calculated and then a non-response weight.

Design weights were applied to compensate for unequal probabilities of selection. The variables compensated for were number of eligible adults in household and number of telephone lines. For any “don’t know” or missing values the imputed values used were 2 for number of eligible adults, and 1 for number of telephone lines.

Non-response weights were applied to compensate for differential response rates by key demographic variables. The non-response weights used were Government Office Region, Age, and Gender within Working Status. For any “don’t know” or missing values imputed values were used. These were: working status (imputed to full-time), age (imputed to 35-44), and gender (men). The weighting was done using the “rim weighting” technique and the weighting targets, derived from the June – August 2005 personal LFS dataset, were as follows:

<i>Region</i>	
London	12.60%
South East	14.90%
South West	8.90%
East of England	9.85%
East Midlands	7.39%
West Midlands	8.93%
North East	4.05%
North West (including Merseyside)	11.28%
Yorkshire and The Humber	8.48%
Wales	4.80%
Scotland	8.82%

<i>Age</i>	
16-24	15.00%
25-34	21.30%
35-44	25.60%
45-54	21.50%
55-64	14.20%
65+	2.40%

<i>Gender within Working Status</i>	
Male Full-time	47.00%
Male Part-time	6.00%
Female Full-time	27.00%
Female Part-time	20.00%

The sample profile was then inspected to assess the feasibility of weighting by industry and/or occupation. The profile achieved for the variables SIC, SOC and NS-SEC was compared with the national profile to see if weighting was necessary. The results are below. It was decided that the sample profile by NS-SEC and SOC were sufficiently close to the national profile to make weighting unnecessary. As 9% of SIC codes were missing (due to inadequate information given by respondents) it was decided not to weight by SIC.

<i>SOC</i>	<i>GB</i>	<i>Valid Weighted sample</i>
Managers and senior officials	14.9%	13.9%
Professional occupations	12.6%	13.2%
Associate professional and technical	14.0%	15.5%
Administrative and secretarial	12.6%	12.9%
Skilled trades occupations	11.2%	10.6%
Personal service occupations	7.7%	8.5%
Sales and customer services occupations	7.8%	7.9%
Process plant and machine operatives	7.5%	7.0%
Elementary occupations	11.5%	10.5%

<i>NS-SEC</i>	<i>England and Wales</i>	<i>Valid Weighted sample</i>
Higher/lower managerial and professional occupations	38%	43%
Intermediate occupations/small employers and own account workers/lower supervisory and technical occupations	33%	33%
Semi-routine and routine occupations	29%	24%

<i>SIC</i>	<i>GB</i>	<i>Valid Weighted sample</i>
Agriculture, hunting, forestry and fishing	1%	1%
Energy and water	1%	1%
Manufacturing	12%	12%
Construction	5%	7%
Distribution, hotels and restaurants	24%	18%
Transport, storage and communications	6%	7%
Banking, finance and insurance	20%	15%
Public administration, education and health	26%	35%
Other services	5%	5%

The weights were then scaled so that the weighted number of cases was the same as the unweighted number.

7.2 Design Effects

Design effects have been calculated for three of the survey's most important variables and are shown below:

<i>Question</i>	<i>Qwsick</i>	<i>Qwill</i>	<i>Qwinj</i>
Proportion saying "yes"	40%	9.8%	6.6%
Standard Error	0.006	0.003	0.003
Lower 95% Confidence Interval	38.9%	9.2%	6.0%
Upper 95% Confidence Interval	41.1%	10.5%	7.2%
Design Effect (variance multiplier)	1.30	1.22	1.38
Design Factor (Standard Error multiplier)	1.14	1.11	1.17