

Dadansoddi ar gyfer Polisi



Analysis for Policy

Ymchwil gymdeithasol

Social research

Rhif/Number: 21/2015



Llywodraeth Cymru
Welsh Government

www.cymru.gov.uk

Review of the draft questionnaire for the National Survey for Wales, 2016 onwards

A review of the draft questionnaire for the National Survey for Wales, 2016 onwards

Jo d'Ardenne, Sally McManus and Melanie Doyle

NatCen Social Research

Views expressed in this report are those of the researchers and not necessarily those of the Welsh Government.

For further information please contact:

Chris McGowan

Knowledge and Analytical Services

Welsh Government

Cathays Park

Cardiff CF10 3NQ

Tel: 029 2080 1067

Email: chris.mcgowan@wales.gsi.gov.uk

Welsh Government Social Research, 27 March 2015

ISBN 978-1-4734-3364-9

© Crown Copyright 2015



All content is available under the Open Government Licence v3.0 , except where otherwise stated.

<http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>

Contents

1	Background and methodology	4
1.1	Context	4
1.2	Methodology	5
1.2.1	Providing length estimations	5
1.2.2	Reviewing selected modules against alternative approaches	6
1.2.3	Reviewing questions for risks of mode effects	7
1.2.4	Review questionnaire content and suggesting a questionnaire order	7
1.2.5	Reviewing the questionnaire for context effects	8
1.2.6	Conducting 'in-office' test interviews	10
2	Review of the CORE questions.....	11
2.1	Module content and duration.....	11
2.2	Comments on specific modules	13
2.2.1	Start of the respondent interview	14
2.2.2	Tenure	14
2.2.3	Welsh language ability	14
2.2.4	Health demographics.....	15
2.2.5	Wellbeing.....	15
2.2.6	Demographic questions	16
2.2.7	Material deprivation	19
2.3	Missing variables of interest?.....	22
2.3.1	Coverage of the 2010 Equalities Act.....	22
2.3.2	Other classification items	23
3	Review of Welsh Health Survey (WHS) questions ...	24
3.1	Questionnaire content and duration	24
3.1.1	Are there shorter existing health questions?	26
3.2	Alternatives to the Health and Wellbeing module (the SF36).....	28
3.2.1	Measuring wellbeing	28
3.2.2	Single item measures	29
3.2.3	ONS4	29
3.2.4	Satisfaction measures	30
3.2.5	WEMWBS and SWEMWBS.....	30
3.2.6	Measuring 'low wellbeing'	31
3.2.7	Summary and recommendations	31
3.3	Adjustments for change in mode.....	33
3.3.1	Summary and recommendations	33

3.4	Comments on specific modules	34
3.4.1	Illnesses and other health problems.....	34
3.4.2	Health and wellbeing (SF-36)	36
3.4.3	Smoking	36
3.4.4	Alcohol.....	37
3.4.5	Fruit and vegetable consumption	37
3.4.6	Exercise.....	38
3.4.7	Other questions (care, BMI and pregnancy variables)	39
4	Welsh Outdoor Recreation Survey (WORS).....	40
4.1	Module content and duration.....	40
4.2	Minimising mode-effects	41
4.3	Comments on specific sections.....	43
4.4	Summary and recommendations	45
5	Length estimates of further questionnaire sections ..	47
5.1	National Survey (NS)	47
5.2	Arts in Wales Survey (AIWS)	49
5.3	Active Adult Survey (AAS)	50
5.4	Overall questionnaire length	52
5.4.1	Contribution of annual components	52
5.4.2	Contribution of rotating modules	53
6	Questionnaire order	55
6.1	Areas of overlap.....	55
6.1.1	NS and WHS sections on ‘use of health services’.....	55
6.1.2	NS and WHS questions on walking	57
6.1.3	NS and AIWS questions on ‘Arts/Museums/Heritage’	57
6.1.4	AIWS, AAS, WORS	58
6.2	Questionnaire outline	59
6.2.1	Further research on question order during the field-test.....	61
7	Reviewing the questionnaire for context effects	63
7.1	CORE questions	63
7.2	National Survey for Wales.....	65
7.3	Welsh Health Survey (WHS).....	68
7.4	Welsh Outdoor Recreation Survey (WORS)	70
7.5	Arts in Wales Survey (AIWS)	71
7.6	Active Adults Survey (AAS).....	73
8	Findings from test interviews	74
8.1	Consistency of answer options	74
8.2	Consistency of reference periods.....	75
8.3	Additional comments.....	75
	Appendix A. How WHS data is currently used.....	77
	Appendix B. HSE – Alcohol consumption.....	80

Appendix C. Copy of coder briefing instructions	91
Appendix D. Coder recording sheet	100

1 Background and methodology

1.1 Context

Welsh Government (WG) and Welsh Government sponsored bodies currently commission a number of large-scale surveys in Wales. These surveys provide essential information on a range of policy relevant topics and the findings are widely used to provide national and local level estimates, to examine differences between population subgroups, to set and monitor targets, and to develop strategies at both national and local level. Demand from the data is from a wide range of users including the Welsh Government and sponsored bodies themselves, local government, the voluntary sector, and academic users.

The cost of carrying out large-scale surveys is relatively high. Therefore in spring and summer 2014 the Welsh Government and its partners commissioned a review of the various options for carrying out large scale surveys in Wales.¹ The purpose of this review was to identify options for collecting survey information in a more cost effective way without compromising data quality. The surveys in scope of this review were several regular large-scale surveys of adult individuals. It has now been decided to integrate five existing surveys into a single survey vehicle from 2016 onwards. The surveys that will be integrated are:

1. The National Survey for Wales 2012-15 (NS)
2. Welsh Health Survey (WHS)
3. Active Adults Survey (AAS)
4. Arts in Wales Survey (AIWS)
5. Wales Outdoor Recreation Survey (WORS)

The approach proposed for the new survey is one that is methodologically similar to the National Survey for Wales 2012-15. It will involve approximately 12,000 adult interviews each year and will consist of Computer Administered Personal Interview (CAPI) modules plus a small Computer Assisted Self-Interview (CASI) element. It will cover a range of topics from the five surveys currently in scope. Additional ad hoc questions could potentially be included in future surveys.

¹ Smith P, Nicolaas G, Sturgis P. *Options for large-scale surveys in Wales: A report for the Welsh Government*. Welsh Government Social Research, 2014.

The Welsh Government is conducting a range of work to develop the new integrated survey including:

1. A public consultation on topics for the new survey;
2. A large-scale field test;
3. Work to explore statistical discontinuities between the current surveys and the new survey; and
4. A desk review of the draft questionnaire for the large-scale test

This report documents the findings from the *fourth strand* of development work: the desk review of the draft questionnaire.

1.2 Methodology

The draft questionnaire under review includes selected questions from the five surveys in scope. The purpose of the review is to conduct a desk-based evaluation of the new questionnaire prior to field testing being carried out.

The review involved the following stages:

1. Providing length estimations for each of the proposed questionnaire modules.
2. Reviewing selected modules against alternative approaches used in other surveys.
3. Reviewing WHS and WORS questions for risks of mode-effects.
4. Reviewing the overall content to assess whether there were any areas of duplication or overlap. This stage also involved proposing an order for the integrated questionnaire modules.
5. Reviewing the reordered questionnaire for risks of context effects.
6. Conducting 'in-office' interviews to check the overall flow of the refined instrument.

The desk-review was conducted from December 2014 through to January 2015. Two interim reports, with recommendations for questionnaire cuts and alterations, were submitted to WG during this period. Amendments to the questionnaire were made iteratively based on feedback from various stakeholders. The methods used are described in further detail in the following sections.

1.2.1 Providing length estimations

As part of the desk review we provided timing estimates for each proposed question. The aim of this was to estimate how much time each individual question, group of questions or module could contribute to average (mean) questionnaire length.

Overall questionnaire length is dependent on two factors: the amount of time it takes to ask each individual question, and the proportion of respondents who will answer each question. We used a simple code to estimate how long it would take to administer each question. This code-frame is shown in Table 1.2.1 below:

Table 1.2.1 Timing estimate code-frame

Question format	Type no:	Predicted contribution to mean interview length (mins)
Simple (e.g. short yes/ no or other simple codes)	1	0.25
Complex single code (e.g. long question stem or list of answer codes)	2	0.37
Complex multi-code code (as above but multi-code)	3	0.75
Open question (verbatim or interviewer multi-code response)	4	0.66
Repeated attitude statements/ repeated answer categories	5	0.17

In cases where a question (or a set of introductions) did not fit this code-frame coders were asked to read the question out loud in order to establish a rough length estimate.

After assigning each question a length code an estimate was made of what proportion of respondents would answer the question in practice. These estimates were based on national statistics, what proportion of respondents had answered the question in previous surveys and assumptions made about module sub-sampling. From this information we made a prediction of how much time each question would contribute to mean interview length. For example, a 'simple question' (code 1) asked to approximately half the sample would contribute 0.125 minutes to predicted interview length.

Findings from this length predicting exercise are described in Chapters 2, 3, 4 and 5. It should be noted that all timing estimates are based on the above heuristics. They are designed to provide a rough estimate of length only. All estimates are subject to a margin of error and will need verifying as part of the forthcoming field-test.

1.2.2 Reviewing selected modules against alternative approaches

As part of the desk review we looked at alternative approaches for capturing demographic and health information based on what is collected in other UK surveys. A range of other questionnaires were reviewed including:

- Harmonised questions recommended by ONS.
- UK census questions (please note there is some overlap between these and the harmonised questions).
- The Health Survey for England (HSE).

- Labour Force Survey (LFS 2013).
- The British Social Attitude Survey (BSA 2013).
- The Crime Survey for England and Wales (2013).

Findings from this exercise are described in Chapters 2, 3 and 4. Details on what questionnaires were reviewed are reported in more detail in the relevant sections.

1.2.3 Reviewing questions for risks of mode effects

As part of the desk review we looked at the risk of mode effects occurring in the Welsh Health Survey questions (in light of the mode change from paper self-completion to computerised interviewer administration). As part of this review we considered whether WHS questions should be asked in CAPI or CASI, whether any interviewer instructions would be beneficial and whether any additional changes could be made to maximise the possibilities of computerised administration (more complex routing, etc.). Findings from this exercise are described in Chapter 3.

We also reviewed the questions from the Welsh Outdoor Recreation Survey in light of the mode change (from telephone administration to face-to-face administration). As part of our review we made suggestions on whether showcards should or shouldn't be introduced to specific questions. Findings from this exercise are described in Chapter 4.

1.2.4 Review questionnaire content and suggesting a questionnaire order

After reviewing all the proposed questionnaire modules in terms of estimated length and content we reviewed the overall topic coverage. The aim of this was to establish whether there were any areas of duplication and to recommend a questionnaire order. The rationale when suggesting an order was to:

1. Provide a logical narrative or flow (for instance by clustering sections on similar topics together).
2. To start the questionnaire with non-sensitive, universally relevant questions to ease participants into the questionnaire. More sensitive questions should come later in the instrument after the interviewer has had time to develop a rapport with the interviewee.
3. The most sensitive items should be included in a separate self-completion (CASI) module. The aim of this is to reduce the risk of inaccurate reporting and question refusals. However, we attempted to keep the number of CASI items to a minimum (to keep the interview streamlined for participants less confident using computers).
4. To have 'rotating modules' or sub-sample modules towards the end of the questionnaire. We did this to reduce the risk of question order effects.

For some questions two or more of the above considerations were mutually contradictory. For example some questions would logically sit within one section because of their subject matter but were also candidates for CASI because of their sensitivity and/or were also flagged as potential candidates for rotating modules. In these cases we have suggested an item positioning on a case-by-case basis.

Findings from this exercise are described in Chapter 6.

1.2.5 Reviewing the questionnaire for context effects

After suggesting a questionnaire order the entire questionnaire was reviewed for risks of context effects. The questionnaire is mostly made up of questions that have been asked before in one of the five surveys that are being combined. However, just because a question has been asked before (and the wording has not changed) is not a guarantee that it will be understood in the same way between surveys. When completing a questionnaire survey participants do not view questions in isolation. Instead, the order in which questions are asked, and the prior material covered, can influence how questions are answered. The same questions can elicit different response distributions if the preceding material is different. This is an example of a *context effect*.

Although there are experimental cases studies providing examples of context effects there is little literature on which questions are at risk and which questions are not.² To assess the risk of context effects we adapted a psychological model (Tourangeau's Question and Answer Model³) to specifically look for potential triggers of context effects.

The Question and Answer Model describes four cognitive processes survey participants go through when answering a survey question. These four stages are:

1. **Comprehension:** In which participants have to work out what information is being requested.
2. **Retrieval:** Whereby participants have to access the information requested (either from memory or by checking records).
3. **Judgement:** Where participants make decisions on what level of information they are willing to provide or make assumptions about what information is actually required. At this stage participants may decide to edit answers based on what is most socially desirable, what responses they have previously given and so on.
4. **Response mapping:** Where participants, having established what answer they wish to give, try to find an answer category that best fits their in-mind response from the fixed options provided. If no suitable answer option is found participants may have to revisit the above three processes.

It is worth noting that these four stages do not necessarily occur in a linear fashion, they can run concurrently and each stage may be re-visited as part of the reflection process before a final response is elicited (see figure 1.2.5 overleaf). Based on this model we assumed that changes in question order could potentially introduce context effects via three pathways. These were by changes in question order creating:

1. *Possible changes in comprehension.*

² This issue is described in Smith P, Nicolaas G & Sturgis P *Options for carrying out large-scale surveys in Wales: A report for the Welsh Government*. Welsh Government Social Research, 2014.

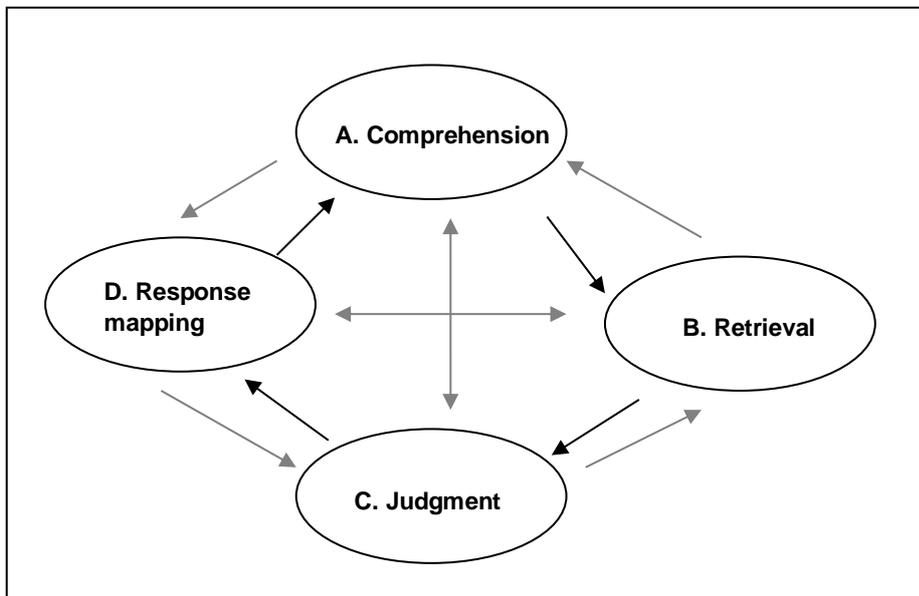
³ Tourangeau R 'Cognitive science and survey methods: A cognitive perspective'. In T Jabine, M Straf, J Tanur, R Tourangeau. (eds) *Cognitive Aspects of Survey Design: Building a bridge between the disciplines*. Washington DC; National Academy Press, 1984: 73-100.

2. Possible changes in retrieval.

3. Possible changes in judgment.

We assumed that the response mapping phase should not be directly affected by a change in question order if answer options have remain consistent.

Figure 1.2.5: Elaborated Question and Answer Model



In order to assess risk of context effects, each section of the new questionnaire was reviewed. This review was conducted by a team of questionnaire coders. Coding involved:

1. Reading a questionnaire section within the new questionnaire (e.g. WHS-Smoking).
2. Comparing the section against the equivalent section in the most recent original questionnaire.
3. Reading the directly preceding sections in both the new questionnaire and the original questionnaire.
4. Deciding whether any of the questions in the section were at risk of context effects. This was defined by a change in the preceding material having the potential to influence comprehension, retrieval or judgment.
5. Recording the variable names of the 'at risk' questions, the potential cause of the risk (change in comprehension, retrieval or judgement) and an open-ended description of the potential impact.

All coders attended a face-to-face briefing on how to identify risk factors that could contribute to context effects. Examples of 'at risk' questions were provided at this briefing. Briefing instructions are included in Appendix A.

Two coders reviewed each questionnaire section and recorded their findings independently to one another. All findings were entered into a standardised pro-forma (see Appendix B). Once all sections had been reviewed the completed forms were reviewed by a third coder. In the case of disagreement between coders the third coder made a final assessment on whether a questionnaire section was at risk of context effects or not.

Recommendations based on the context effect review are summarised in Chapter 7. In addition to recording risks of context effects, coders were asked to flag any questions which they felt could be done with further refinement or that they did not understand. These additional comments are described alongside the main findings of the context effect review as part of Chapter 7.

1.2.6 Conducting 'in-office' test interviews

After the context effect review a small number of test interviews were conducted in the office. The purpose of this exercise was to double-check the overall flow of the questionnaire once further order changes had been made.

In total four test interviews were conducted by NatCen staff. A paper version of the questionnaire was read out loud in order to mimic (as far as possible) a CAPI/CASI interview.

A different route through the questionnaire was trialled in each of the four test interviews. This was to reflect the fact that different groups will be asked different questions due to:

- Questionnaire routing (e.g. some questions are only asked of specific groups e.g. Welsh speakers or parents)
- Within-year sub-sampling of modules (e.g. where some questions are only asked of a sub-sample of respondents each year in order to reduce overall questionnaire length)
- Rotating modules between years (e.g. some modules are currently planned for inclusion every two years or every four years).

After completing the test interview, interviewees were asked to rate the new questionnaire in terms of its overall structure and flow. Interviewers were asked to give feedback on questionnaire administration. Recommendations based on the test interviews are summarised in Chapter 8.

2 Review of the CORE questions

The new integrated questionnaire will contain a set of core questions. These will be asked every time the survey is conducted and will be used primarily for cross-analysing other topics. The first part of the desk review involved our reviewing the proposed items for inclusion in the CORE questionnaire. The aim of the review was:

- To provide length estimations on administration time.
- To identify alternative approaches used in other surveys (in particular 'shorter' ways of collecting the same information).
- To identify whether any classifications of interest are currently missing.

Findings on each of these areas are summarised in the following sections.

2.1 Module content and duration

The original draft of the CORE module contained the following sections:

1. **Start of the respondent interview** (checking of household composition, sex, marital status, relationships to other household members, whether in full-time education).
2. **Tenure**
3. **Welsh language ability**
4. **Health demographic questions** (self-rated general health, whether the respondent has a limiting long-standing illness)
5. **Wellbeing** (the four questions on personal wellbeing developed by ONS)
6. **Demographics** (nationality, national identity, ethnicity, sexual orientation, religion, qualifications, whether any employment in the last 7 days, whether ever had any employment, whether any household member in employment).
7. **Household material deprivation**
8. **Child material deprivation**
9. **Permission to re-contact and data linkage**

The first interim report predicted that in total these sections would contribute 14 minutes to mean interview length. As result of the interim report a number of CORE sections were amended. Alterations include:

- The question on tenure has been replaced with a simpler item from the WHS, and an item on 'type of landlord' (HhLndLrd2) has been dropped.

- An item on 'country of birth' (CountBirth) has been dropped.
- The number of questions on employment status has been reduced. These questions have been replaced with two items (adapted from WHS and the census) on overall employment status in the last week and number of hours worked per week. In keeping with feedback from stakeholders the answer categories used mimicked Census categories as far as possible to assist in the calculation of small area estimates.
- Type of illness items were added to the section on health demographics.

In addition some of the original length estimates from the interim report have been refined in keeping with timestamp data available from 2012-13 and 2014-15 National Survey fieldwork. This mainly applies to the introductory section (start of the respondent interview) and the data linkage section, where our timing assumptions appeared to overestimate the amount of time required to introduce non-standard sections.

Based on these changes our new timing estimates for the CORE section are as follows:

Table 2.1: Predicted module length of CORE section

Module section	Contribution to predicted mean length (mins)
Start of the respondent interview	2.1
Tenure	0.3
Welsh language ability	0.7
Health demographic questions	1.0
Wellbeing	0.8
Demographics (includes nationality, ethnicity, qualifications, economic activity, sexual orientation and religion)	1.9
Household material deprivation (based on a reduced set of around 5 questions)	1.1
Child material deprivation (based on a reduced set of around 5 questions)	0.3
Permission to re-contact and data linkage	2.0
TOTAL	10.2

Currently the aim is for the CORE questions to last approximately 8.5 minutes and the current length estimate is around 10 minutes. However it should be noted that our length predictions slightly *overestimated* some other National Survey sections based on timestamp data reviewed from the 2012-2013 and 2014-2015 interviews. As explained in section 1.2.1 the timing assumptions we use follow a heuristic based on the 'question type' and predictions of what proportion of the sample will be routed to the question. These estimations are subject to a margin of error. Therefore fully accurate measures of length should be collected via the forthcoming field test. Therefore we recommend that no further cuts are made to the CORE questions prior to the field test being completed.

2.2 Comments on specific modules

As part of reviewing the CORE module we looked at existing survey questions looking for 'shorter' ways of collecting the same information. When reviewing alternatives from other surveys we focussed on:

- Harmonised questions recommended by ONS.
- UK Census questions (please note there is some overlap between these and the harmonised questions).
- Questions alternatives from WHS, WORS, AIWS and AAS.
- Recent surveys that are administered in CAPI. Questionnaires included in the review so far have included:
 - The Annual Population Survey / Labour Force Survey (LFS 2013)
 - The British Social Attitude Survey (BSA 2013)
 - The Community Life survey (2013)
 - The Crime Survey for England and Wales (2013)

Other sources of health questions were also reviewed, for instance the Health Survey for England and the Scottish Health Survey (see Chapter 3). Other questionnaires on outdoor recreation have also been reviewed (see Chapter 4).

There were some variations in the classification questions used in all the surveys reviewed. Generally speaking surveys that exist entirely to describe the characteristics of the population (such as the LFS questions) are relatively complex and go into a lot of detail. Therefore, although the questions are highly nuanced they are more time-consuming to administer. In contrast, multi-topic omnibus style collect less detailed information. When deciding classification questions to use there is a trade-off between overall length and the depth of data collected.

Generally speaking we agree with WG's existing approach for using harmonised questions where possible (including UK Census questions). As well as helping ensure

the questions are good quality, this will allow for comparisons to be made between other government surveys and allow for more consistency when making comparisons across time.

2.2.1 Start of the respondent interview

The first part of the respondent questionnaire involves collecting information on sex, age, marital status and relationships to other household members. We estimate that this section will contribute just less than three minutes to average interview length. We recommend that all these questions are retained as some of them are required to monitor protected characteristics (see 2.3). Others are required for subsequent routing (for instance if respondents are living with a partner or child).

Comments and suggested amendments

- We recommend that these questions are retained in their current format.

2.2.2 Tenure

The CORE questions originally contained two questions on tenure. As part of the review, WG queried whether these items are necessary. It is our view that tenure may be useful for analysis (e.g. for looking at how health, well-being or material deprivation is related to household ownership and type of tenure).

Comments and suggested amendments

- Following discussions with WG the main question on tenure (Tenure2) has been replaced with a simpler question from the WHS, and the follow-up question on type of landlord (HhLndLrd2) has been dropped.

2.2.3 Welsh language ability

The CORE questions currently contain six questions on Welsh Language ability. We estimate that this section will contribute less than a minute to average interview length. Whether or not these questions should be asked every year (or rotated) could be discussed with data-users. Whether or not rotation is appropriate is dependent on how regularly data-users require estimates on Welsh language and whether Welsh language is regularly used as a cross-tabulation classification.

Comments and suggested amendments

- Following discussions with WG these questions have been retained in their current format. One item on frequency of Welsh speaking (WelFrqSpk) has been rated as potentially less useful than the other items. This question is a candidate for dropping if the overall questionnaire is identified as too long after the field test.

2.2.4 Health demographics

The CORE questionnaire originally contained three questions described as 'Health demographics.' The questions are on self-rated general health and whether the respondent has any limiting long-term health conditions. We recommend the current questions (which are harmonised) are retained in their current format and that the follow-up question on type of long-standing illness is moved to come directly after these questions (Llticd1, previously in the WHS section of questions).

In addition to this we recommend that multiple health conditions should be captured at this question, following a similar approach to that used on HSE. This should have very little impact on mean questionnaire length as only a small minority of respondents will report having multiple health conditions (18% of participants report two health conditions in HSE and 9% report three health conditions in HSE).

It should be noted that the harmonised questions on limiting health conditions are not the only proposed questions on health conditions. More detailed questions on treated conditions are also being proposed (see section 3.4.1 for further details). We do not consider these two approaches to be mutually exclusive as both serve slightly different functions. The inclusion of both approaches (limiting conditions and treated conditions) allows participants to report conditions that limit their activities that they are not receiving treatment for and report conditions that they are receiving treatment for that they do not consider limiting (e.g. fully managed diabetes, asthma and so on). The questions about treated illness invite a yes/no and could prompt participants to include conditions that they do not currently experience as severe.

Comments and suggested amendments

- We recommend that the NS 'Health Demographic' questions are retained in their current format. Llticd1 (an open question on type of condition- currently in the WHS module) should be included alongside these questions. We also recommend that participants should be able to report multiple types of health condition at Llticd1. We propose capturing up to six types of health condition that limit day-to-day activities.
- In addition to the Llticd the CASI module contains further questions on what health conditions respondents are currently receiving treatment for (see section 3.4). This means that the proposed questionnaire will capture data on types of health conditions even if participants don't consider their condition to limit their activities.

2.2.5 Wellbeing

The CORE questionnaire currently contains the four ONS questions on personal wellbeing. We estimate that this section will contribute less than a minute to average interview length.

We recommend the current questions (which are harmonised) are retained in their current format. These can then be asked every year even if more detailed wellbeing are dropped or asked on rotation. We recommend that these questions are included as

part of the CASI to encourage honest reporting (see section 3.2). It should be noted that a move to CASI may bring about systematic changes in reporting.

Comments and suggested amendments

- We recommend that these questions are retained in their current format and moved to the CASI section.

2.2.6 Demographic questions

The CORE questionnaire originally contained a number of questions described as 'demographics'. This section included items on country of birth, national identity, ethnicity, sexual orientation, religion, qualifications, whether any employment for the last 7 days, whether ever had any employment, and whether any household member is in employment. Our views on each question area are described overleaf.

'Nationality' and 'National identity'

The demographics section contained two questions on 'country of birth' and 'national identity.' Whether or not both these questions should be dropped (or rotated) could be discussed with data-users. Whether or not dropping (or rotating) is appropriate is dependent on how regularly data-users require estimates on these domains and whether these domains are regularly used in cross-tabulation classifications.

Comments and suggested amendments

- After consultation with WG the question on 'country of birth' has been dropped.

'Ethnicity'

The demographics section contains one question on 'ethnicity'. This involves a showcard containing 18 separate classifications.

WG have queried whether or not this question could be simplified by providing a reduced list of answer options. Reducing the list of options will not have a big impact on how long it takes to administer this question. This is because the most prevalent classification (White-Welsh) is listed at the top of the showcard. Therefore the majority of respondents will not need to read all the classifications in order to pick the most suitable answer. In contrast reducing the list of answer options could annoy people who are no longer able to locate their preferred description.

Comments and suggested amendments

- We recommend the current question wording is retained.

'Sexual orientation'

We recommend the sexual orientation question is moved to the CASI section of the questionnaire.

'Religion'

We recommend the religion question is moved to the CASI section of the questionnaire.

'Qualifications'

The demographics section originally contained one question on 'qualifications'. The current approach is taken from the UK Census 2011. Participants are required to look at a showcard listing 13 types of qualification (including no qualifications) and to select which ones they hold.

WG think the current approach could take too long to administer and a more simple approach would be beneficial. An alternative approach would be to only collect information on participants' 'Highest qualification' rather than each type of qualification held. An example of a highest qualification question is shown in the box below:

Example of highest qualification questions

Taken from the Crime Survey for England and Wales (CSE 2013). Similar approaches are used in other surveys

EDUCINT [ASK ALL]

SHOWCARD

Do you have any of the qualifications listed on this card?

1. Yes
2. No

EDUCAT2 [ASK IF EDUCINT = YES]

SHOWCARD

Starting from the top of the card, please look down the list of qualifications and tell me the number of the first one you come to that you have passed.

1. Higher degree/postgraduate qualifications
2. First degree (including B. Ed.)
Postgraduate diplomas/Certificates (inc. PGCE) Professional qualifications at degree level (e.g. chartered accountant/surveyor)
NVQ/SVQ Level 4 or 5
3. Diplomas in higher education/other H.E. qualifications, HNC/HND/BTEC Higher Teaching qualifications for schools/further education (below degree level)
Nursing/other medical qualifications (below degree level) RSA Higher Diploma
4. A/AS levels/SCE Higher/Scottish Certificate 6th Year Studies NVQ/SVQ/GSVQ level 3/GNVQ Advanced ONC/OND/BTEC National
City and Guilds Advanced Craft/Final level/ Part III/RSA Advanced Diploma
5. Trade Apprenticeships
6. O Level/GCSE grades A-C/SCE Standard/Ordinary grades 1-3 CSE grade 1
NVQ/SVQ/GSVQ level 2/GNVQ intermediate
BTEC/SCOTVEC first/General diploma
City and Guilds Craft/Ordinary level/Part II/RSA Diploma
7. O Level/GCSE grades D-G/SCE Standard/Ordinary below grade 3 CSE grades 2-5

NVQ/SVQ/GSVQ level 1/GNVQ foundation BTEC/SCOTVEC first/General Certificate City and Guilds part 1/RSA Stage I-III SCOTVEC modules/Junior certificate

8. Other qualifications (including overseas)

We recommend that a highest qualification approach could be adopted without having a major negative impact on analysis. However, it should be noted that this change will only equate to a very small reduction in average interview length (seconds not minutes).

Comments and suggested amendments

- After discussion WG agreed they are happy to consider an alternative question to collect information on qualifications. We have suggested an approach similar to the CSE question show above but with a slightly simplified question stem. The list of qualifications has also been updated to include the Welsh Baccalaureate.

Economic activity

The demographics section contains eight questions on economic activity. Most of these are harmonised questions included in the UK Census 2011. Participants are first required to say whether they were doing any work in the last seven days. Participants who were not working are asked follow-up questions on whether they are looking for work, whether they have ever worked and how long ago they last worked. An additional question is asked to establish whether any other household members are working.

WG think the current economic activity questions could take too long to administer and a more simple approach would be beneficial, given that the information is only collected for cross-analysis purposes. An alternative approach would be to ask a single question on working status from which participants can select their 'main activity' or the first activity to apply. Examples of these are shown below:

Examples of alternative questions on economic activity

Taken from WHS

Which of these descriptions applies to what you were doing last week? Tick first to apply.

1. Going to school or college full-time (including on vacation)
2. In paid employment or self-employment (or away temporarily)
3. On a Government scheme for employment training
4. Doing unpaid work for a business that you own, or that a relative owns
5. Waiting to take up paid work already obtained
6. Looking for paid work or a Government training scheme
7. Intending to look for work but prevented by temporary sickness or injury (sick or injured for 28 days or less)
8. Permanently unable to work because of long-term sickness or disability
9. Retired from paid work
10. Looking after the home or family
11. Doing something else

Taken from WORS (TNS omnibus demographic questions)
<p>What is your current working status?</p> <ol style="list-style-type: none"> 1. Working full-time (30+ hours per week) 2. Working part-time (9 to 29 hours per week) 3. Unemployed – less than 12 months 4. Unemployed (long term) – more than 12 months 5. Not working – retired 6. Not working – looking after house/children 7. Not working – long term sick or disabled 8. Student – in full-time education 9. Student – in part-time education 10. Other
Taken from Active Adults (Ipsos-Mori demographic questions)
<ol style="list-style-type: none"> 1. Working full time (30 hours or more) 2. Working part time (9-29 hours) 3. Unemployed/seeking work 4. Full-time student 5. Wholly retired 6. Other permanent not working

Whether or not changing the approach is desirable is dependent on current data-users. Do data-users make use of the more nuanced data collected by the harmonised questions? Is it important for data-users to establish whether respondents have ever worked? Is it important for them to know how long ago participants last worked? Some savings on questionnaire length could be made by altering the questions but this would be at the expense of detail and comparability with other government data sources.

Comments and suggested amendments

- After discussion with WG we have recommended a new economic activity question based on the item currently used in the WHS. The rationale for this was it was considered important to have answer options that matched on to Census classifications as far as possible, to support future small area estimation modelling based on the survey results. A question on hours worked per week was also added on the request of data users.

2.2.7 Material deprivation

The CORE items contain two modules on material deprivation. One module aims to measure Household Deprivation. The second module aims to measure child deprivation.

It is however not intended to ask the full set of material deprivation questions of all respondents. WG is developing a subset of the questions designed to classify most respondents in the same way as the full set (possibly asking the full set of questions of respondents on the 'deprived / non-deprived' borderline). In a similar approach, both the English Longitudinal Study of Ageing (ELSA) and the Adult Psychiatric Morbidity Survey 2015 have used (or will be using) shortened versions of the material deprivation items. The items used in both these surveys are displayed in the box overleaf.

Examples of shortened material deprivation modules

Taken from the English Longitudinal Study of Ageing (ELSA)

EXRele

Does having too little money stop you from doing any of the following things?

SHOWCARD. MULTICODE

1. Buy your first choices of food items
2. Have family and friends round for a drink or meal,
3. Have an outfit to wear for social or family occasions,
4. Keep your home in a reasonable state of decoration
5. Replace or repair broken electrical goods
6. Pay for fares or other transport costs to get to places you want to go
7. Buy presents for friends or family once a year
8. Take the sorts of holidays you want
9. Treat yourself from time to time

SPONTANEOUS None of these

Taken from Adult Psychiatric Morbidity Survey (2015)

MATDEPINTRO

Next I have some questions about the sorts of things that some families or people have, but which many people have difficulty finding the money for. For each of the following things please tell me the number from the showcard that best explains whether you (and your family or partner) have it or not.

MATDEPA

Do you (and your family or partner) have a holiday away from home for at least one week a year, whilst not staying with relatives at their home?

SHOWCARD

- 1 I/we have this
- 2 I/we would like to have this but cannot afford this at the moment
- 3 I/we do not want/need this at the moment
- 4 Does not apply

MATDEPD

Do you (and your family or partner) have enough money to keep your house in a decent state of repair?

SHOWCARD

- 1 I/we have this
- 2 I/we would like to have this but cannot afford this at the moment
- 3 I/we do not want/need this at the moment
- 4 Does not apply

MATDEPF

Do you (and your family or partner) have... money to make regular savings of £10 a month or more for rainy days or retirement?

SHOWCARD

- 1 I/we have this
- 2 I/we would like to have this but cannot afford this at the moment
- 3 I/we do not want/need this at the moment

4 Does not apply

MATDEPH

Do you (and your family or partner) have money to replace or repair major electrical goods such as a refrigerator or a washing machine, when broken?

1 I/we have this

2 I/we would like to have this but cannot afford this at the moment

3 I/we do not want/need this at the moment

4 Does not apply

Comments and suggested amendments

- WG are currently reviewing these items to establish whether a shorter suite of questions could be used. The aim is to reduce the household material deprivation questions to five items and to reduce the child material deprivation questions to five items. We estimate that with these cuts the material deprivation questions would contribute approximately 1.4 minutes to mean interview length.

2.3 Missing variables of interest?

In addition to reviewing the suggested CORE items this review aims to look at whether additional classification questions should be included to allow for other types of sub-group analysis.

2.3.1 Coverage of the 2010 Equalities Act

Part of the scoping review is to ascertain whether the CORE items assess all protected features of the 2010 Equalities Act. Under the act there are nine protected characteristics.⁴ There is an argument that the integrated survey should attempt to assess each of characteristics, in order to assess whether experiences and outcomes vary between different groups. Table 2.3.1 summarises the protected characteristics and whether or not they are currently being assessed in the CORE questions.

Table 2.3.1 Protected characteristics and whether they are currently assessed

Protected characteristic	Currently assessed?	Details
1. Age	Yes	Collected via DtBirth at the start of the respondent interview
2. Disability	Yes	Collected via LongIII and LongIIIlimit in the health demographics section
3. Gender reassignment	No	Not currently collected.
4. Marriage and civil partnership	Yes	Collected via MarStat2 at the start of the respondent interview
5. Pregnancy and maternity	Yes	Pregnancy is collected in the WHS section. 'Living with children' is collected at RelResp2 at the start of the respondent interview.
6. Race	Yes	Collected via Ethnicity during the demographic section of the CORE module.
7. Religion or belief	Yes	Collected via Religion during the demographic section of the CORE module.
8. Sex	Yes	Collected via Gender at the start of the respondent interview
9. Sexual orientation	Yes	Collected via SexInt during the CORE module.

Therefore, eight out of nine protected characteristics are already been assessed by the proposed questions. The only characteristic not being captured is 'gender reassignment.' Currently ONS has not developed a harmonised question on gender-

⁴ See <http://www.legislation.gov.uk/ukpga/2010/15/contents>

reassignment.⁵ However, other surveys have collected this information and the Equality and Human Rights Commission has suggestions for items.⁶ Therefore it would be possible to include a short question (or questions) on gender-reassignment in the self-completion element of the interview. This should have minimal impact on overall survey length (but see concerns mentioned below under “Comments”).

2.3.2 Other classification items

It was noted by the reviewers that the current CORE questionnaire contains no measures on the following areas:

- Household income
- Socio-economic classifications (industry, occupation, and other information used to generate standardized NS-SEC codes).

We appreciate that in the new integrated questionnaire space is at a premium. However, these variables are likely to be relevant when analysing other areas of interest. For example, income is related to well-being, material deprivation, health etc. A simple banded income question should not much add additional length to the mode (approximately 0.25 minutes).

Socio-economic classifications would potentially add more time as they would involve collecting open descriptions of occupation and industry (SIC/SOC codes) as well as follow-up items on management responsibilities. In total these items would contribute approximately 1.5 minutes to the average interview length.

Comments and suggested amendments

- After discussion WG have decided not to add an additional items to the CORE questions due to concerns about the length of the module:
 - No question will be added on gender reassignment. The rationale for this is there is no harmonised question to measure this and, even if a question could be developed, it is likely to be viewed as intrusive for some respondents. It was also felt the results wouldn't be useful for cross-tabulation purposes due to the small numbers falling within this group.
 - The NS-SEC questions will not be added as they are currently not frequently used for analysis.
 - There were concerns about the accuracy of a short-form income question and the reliability of the data collected.

⁵ Further information can be found in the ONS 'Trans Data Position Paper': <http://www.ons.gov.uk/ons/guide-method/measuring-equality/equality/equality-data-review/index.html>

⁶ See <http://www.equalityhumanrights.wales.nhs.uk/sitesplus/documents/1120/collecting-equality-data-%20%282%29.pdf>

3 Review of Welsh Health Survey (WHS) questions

This chapter describes the review of WHS questions in the integrated questionnaire. The aim of the review was:

- To provide length estimations on administration time (and how the length of the questions could be reduced).
- To explore possible alternatives to the Health and Wellbeing module (the SF36).
- To identify other questions that could do with additional refinement. In particular we focussed on whether questions should be refined given the move from paper administration to interviewer (or CASI) administration.

The main findings from each of these areas are discussed in the following sections. Some additional findings on the current use of WHS data can be found in Appendix A.

3.1 Questionnaire content and duration

WHS questions cover health and illness, lifestyle and use of services. Information is currently collected using a self-completion questionnaire. As WHS is a self-completion the information collected is relatively 'high level', providing key information on each topic rather than large amounts of detailed information.

The original WHS modules flagged for inclusion in the integrated questionnaire were as follows:

1. **Health service use** (use of GPs, nurses, hospitals, dentists etc).
2. **Medicines**
3. **Illnesses and other health problems**
4. **Untreated illness**
5. **Health and wellbeing** (the SF36 and details of a limiting long-standing illness)
6. **Smoking.**
7. **Alcohol consumption**
8. **Fruit and vegetable consumption**
9. **Exercise**
10. **Carers**
11. **Additional questions for cross analysis** (height, weight and pregnancy).

In the first interim report we predicted that all the above modules would take approximately 24 minutes to administer on average. The target length was 14 minutes. Since producing the first interim report the following sections have been cut from the WHS questionnaire in consultation with WG:

- Health service use (this cut reduced the predicted mean length by approximately four minutes)
- Medicine (this cut reduced the predicted mean length by approximately 0.6 minutes)
- Untreated problems or symptoms (this cut reduced the predicted mean length by approximately 0.7 minutes)

The Health and Wellbeing module has been replaced with the health demographic questions from the National Survey (see CORE module content in section 2.1). The SF36 has been replaced with WEMWBS (see section 3.2). These substitutions reduced the predicted mean length of the WHS modules by approximately four minutes).

The module on Illnesses and other sections has been split into two parts, a section called 'Vision, Hearing and Teeth' in the CAPI and a section called 'Illnesses and other health problems' in the CASI. Question on accidents, injuries and poisonings (acc-accplace) and stomach upsets (stomnum-dtomdr) have also been dropped. These cuts reduced the predicted mean length by approximately 0.6 minutes)

The WHS alcohol module has been replaced with the HSE alcohol module. Having looked at the timestamp data from HSE we do not think this change will result in any time savings. However, the structure of the HSE questions is more suited to computerised administration.

The revised timing estimates for the WHS section are shown in the table overleaf:

Table 3.1: Predicted module length of WHS sections

Module section	Contribution to predicted average length (mins)
1. Carers	0.4
2. Vision, hearing and teeth	0.9
3. Smoking	1.3
4. Fruit and vegetable consumption	3.6
5. Illnesses and other health problems	1.4
6. Additional questions for cross analysis (BMI pregnancy)	0.6
7. Alcohol consumption (HSE items)	2.1 ⁷
8. Exercise	1.9
9. Wellbeing (WEMWBS)	2.58
TOTAL	14.6

Currently the aim is for the WHS questions to last approximately 14 minutes. Therefore, we recommend that all the questions that are currently retained can continue to be included in the field test.

3.1.1 Are there shorter existing health questions?

We have reviewed questions from other national health surveys, including the Health Survey for England (HSE)⁸, Scottish Health Survey (SHeS)⁹ and Understanding Society (USoc)¹⁰. The questionnaire modules from dedicated health surveys tend to be

⁷ Based on CAPI timestamp data- includes additional questions on whether ever drinks, reasons for not drinking and frequency of drinking in addition alcohol consumed over the last seven days.

⁸ Health Survey for England reports and resources on Health and Social Care Information Centre website. HSE 2013: <http://www.hscic.gov.uk/catalogue/PUB16076>, HSE 2012: <http://www.hscic.gov.uk/catalogue/PUB13218>

⁹ Scottish Health Survey, on Scottish Government website. SHeS 2013: <http://www.scotland.gov.uk/Publications/2014/12/6634>

¹⁰ Understanding Society – Guide to content: <http://www.understandingsociety.ac.uk/documentation/mainstage/long-term-content-plan>.

longer because the CAPI interviews are somewhat longer – HSE is around 60 minutes- and focus only on health questions. Understanding Society is a more general survey with a health component, covering health status, disability, lifestyle and wellbeing. The USoc questions tend to be shorter and often have a slightly different focus to WHS.

Rather than asking shorter set of questions, WG could consider making some of the current WHS modules 'rotations' rather than including them every year. This approach is used to other health surveys. For example HSE has a core set of questions that are asked every year which it supplements with rotating and ad hoc modules.¹¹ Fruit and vegetable intake and exercise, although collected some years, are not part of the core HSE questions. Adopting this approach would reduce the average length of the WHS questionnaire by 2-3 minutes.

Questionnaires, include health status, lifestyle and wellbeing (SF12 and GHQ in CASI), social support: <https://www.understandingsociety.ac.uk/documentation/mainstage/questionnaires>
Overview of health assessment – medicines and bio-measures as part of nurse visit: <https://www.understandingsociety.ac.uk/about/health/health-assessments>; CAPI and CATI questionnaires – nurse health assessments: <https://www.understandingsociety.ac.uk/documentation/health-assessment/questionnaires>

¹¹ The HSE core module comprises of general health, wellbeing smoking, alcohol, health services use, medicines, and measurements of height, weight and blood pressure.

3.2 Alternatives to the Health and Wellbeing module (the SF36)

The biggest contributor to average interview length in the original WHS questionnaire was the health and wellbeing module. This is mostly comprised of the SF36 items. We estimated this module would add over six minutes to the average questionnaire length. A range of alternative wellbeing measures are available. This section discusses the relative merits of different approaches.

3.2.1 Measuring wellbeing

Measures of wellbeing are increasingly included on surveys. However, the word 'wellbeing' can be used to refer to a range of different things. At its most broad, it can be used to cover all the different domains that affect someone's quality of life. For example, the ONS Measuring National Wellbeing programme collects and reports data on ten areas: relationships, personal finance, health, governance, national economy, natural environment, work/education, neighbourhood, and personal wellbeing. The last of these, personal wellbeing, is also widely referred to as subjective wellbeing. This relates to how people experience and evaluate their lives. This is also closely linked with mental wellbeing, which has been described as 'feeling good and functioning well', or as positive mental health.

Here, we assume that it is subjective / personal / mental wellbeing that is of particular interest, rather than an assessment of all aspects of life. There is a wide range of measures available to get at subjective wellbeing. These tend to divide into three main methodological approaches:

- **Experience-based** (which asks how someone was feeling at a specific time or doing a specific activity);
- **Evaluative** (where someone is asked to rate their level of satisfaction, overall or with an aspect of their life); and
- **Eudaimonic** (which focuses more on people's perceptions of value, meaning and purpose in life, and on social relationships).

The focus on measuring wellbeing has also sought to address the 'deficit model' in social research that has tended to capture negative experiences and adversity, rather than positive capital. Some measures, for example the widely used Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS), deliberately include only positive traits.

There is also debate about the extent to which positive mental wellbeing is simply the opposite end of a spectrum from mental illness. The two have been found to be very strongly inversely related, but may also be measuring different constructs rather than just being the opposite of each other. Some reference is made here to measures of mental illness, but generally these are considered to be measuring a somewhat different construct from subjective wellbeing.

Finally, it is important to note that the 'hedonic treadmill' (the fact that people adjust to changes in circumstances and tend to revert to a stable wellbeing 'set point') means

that measuring wellbeing has greatest utility when it comes to picking up on inequality between groups rather than for looking at trends over time.

3.2.2 Single item measures

A number of single item measures have been used on surveys to provide a quick summary of people's wellbeing. These include:

- Life satisfaction.
- Happiness.
- Optimism (e.g.: 'Over the next ten years, on balance do you expect to have more negative experiences or more positive experiences in your life?')

Benefits:

- Short.
- Life satisfaction, the most widely used, is closely associated with most other measures of subjective wellbeing, and so is a fair proxy for overall subjective wellbeing.

Drawbacks:

- Only covers one aspect of wellbeing, for example some people have high life satisfaction (high evaluative wellbeing) combined with feeling low (low hedonic wellbeing).

3.2.3 ONS4

As part of the Measuring National Wellbeing programme ONS identified four single-item measures which together span the 3 key approaches to measuring 'personal wellbeing'. However, they draw on existing questions that have a long history of use on British surveys. The four ONS questions are:

- Overall, how satisfied are you with your life nowadays? (evaluative approach)
- Overall, to what extent do you feel the things you do in your life are worthwhile? (eudemonic approach)
- Overall, how happy did you feel yesterday? (experience approach)
- Overall, how anxious did you feel yesterday? (experience approach)

All are answered using a 0 to 10 scale where 0 is 'not at all' and 10 is 'completely'. These questions are currently already included in the proposed integrated questionnaire as part of the CORE module.

Benefits:

- Relatively short.
- Used widely, so lots of comparable data.

Drawbacks:

- The data cannot be summarised in a single combined score.

3.2.4 Satisfaction measures

On surveys, satisfaction measures are commonly used to get at a subjective evaluation of life overall, or with an aspect of life (e.g. neighbourhood, relationships, job).

Benefits:

- Used widely, so lots of comparable data.
- Can cover a variety of domains.

Drawbacks:

- Reflects expectations, which in turn are driven by characteristics and context. For example, older people are more likely to be satisfied with the same health care than younger people; people are more likely to be satisfied with their job in time of high unemployment than when unemployment is low. This makes interpreting trends difficult.

3.2.5 WEMWBS and SWEMWBS

Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) is a widely used, validated 14 item survey measure of subjective wellbeing. There is also a 7-item version (the Short-WEMWBS (SWEMWBS))¹². It is now included alongside the ONS4 on the Measuring National Wellbeing programme for the personal wellbeing domain. Since 2007 it has been the key indicator used to monitor trends in subjective wellbeing in Scotland. All the items are positively worded, and it covers both hedonic (feeling good) and eudaimonic (getting on with others and functioning well) perspectives of wellbeing.

- I've been feeling optimistic about the future
- I've been feeling useful
- I've been feeling relaxed
- I've been feeling interested in other people
- I've had energy to spare
- I've been dealing with problems well
- I've been thinking clearly
- I've been feeling good about myself
- I've been feeling close to other people
- I've been feeling confident
- I've been able to make up my own mind about things
- I've been feeling loved
- I've been interested in new things
- I've been feeling cheerful

¹² WEMWBS is free to use but is copyrighted. A user guide for WEMWBS is available here: <http://www.healthscotland.com/uploads/documents/7551-WEMWBS%20User%20Guide%20Version%201%20June%202008.pdf> (accessed February 2015). A user guide for SWEMWBS is available here: <http://www.healthscotland.com/uploads/documents/14092-SWEMWBSSept2007.pdf> accessed February 2015)

Benefits:

- Respected and validated.
- Used widely, so lots of comparable data.
- Produces an overall score, easier for reporting.

Drawbacks:

- Longer than the ONS4.

3.2.6 Measuring ‘low wellbeing’

Measures of mental ill health, quality of life, and functioning such as the Short-Form (SF), the General Health Questionnaire (GHQ), and the revised Clinical Interview Schedule (CIS-R) are not featured here because they fit more with a medical model of deficit and limitation. However, while the scoring - and thresholds applied to the scores - for these measures tend to focus on identifying people with mental distress or a poor quality of life, they also include individual items which capture positive aspects of wellbeing. For example, the GHQ12 has items on feeling capable of making decisions and enjoying playing a part in day-to-day activities. We summarise some of these below, as they are commonly used as measures of wellbeing:

GHQ12 – the General Health Survey 12 (which also comes in longer versions, such as the GHQ30), is widely used (for example, on the Health Survey for England). Benefits are that it includes positive items such as being able to concentrate and feel reasonably happy; drawbacks include that its primary output is a score that identifies people experiencing mental distress.

EQ5d – the EQ5d Health Questionnaire is increasingly being used, it is short – just 5 items – but covers a deficit model (asking about things like depression /anxiety, and problems with mobility and self-care). Furthermore, there is a fee for usage.

SF12 – the Short Form 12 (and longer versions like the SF36). Includes positive items like vitality and energy and provides comparability with the Welsh Health Survey; drawbacks are the high license fee for use.

3.2.7 Summary and recommendations

In summary, we recommend (where possible) avoiding reliance on:

- A single item (such as happiness or life satisfaction) as this does not capture the different components of wellbeing.
- Measures of mental illness or distress (such as the GHQ12) as these tend to be based on measuring deficit, rather than positive emotions and functionality.

Often choice will be driven by space available, with longer measures bringing gains in cover and validity. Given this, we recommend the following:

Space available	Recommended measure
4 items	ONS4 (life satisfaction, happiness, anxiety, worthwhile)
7 items	Short-Warwick Edinburgh Mental Wellbeing Scale (SWEMWEBS)
14 items	WEMWBS

After discussion with WG it was agreed for the WEMWBS questions to be trialled as an alternative to the SF36.

It should be noted that multiple approaches for measuring wellbeing could be adopted. For example, the ONS4 could be included every year as part of the CORE module and WEMWBS could be asked as part of a rotating module.

3.3 Adjustments for change in mode

Questions have been taken from a paper self-completion format in the current WHS, for inclusion in the CAPI questionnaire of the new survey. A few general amendments would be beneficial to make the questions work in CAPI. For example, showcards are desirable to display answer options for some questions. Showcards will be beneficial as when answer categories are presented visually they are easier for participants to retain. In the self-completion WHS answer categories are also presented visually so introducing showcards should also help ameliorate the risk of mode effects.

Examples of where showcards should be used include:

- Questions on NHS versus private health service use (denser, chirser, physser, osteoser)
- The extent to which activities are limited (Ilti2).

There are a few general question framing issues to consider when developing the CAPI instrument. For example:

- Some follow-up questions need to be amended so that they make sense in an interviewer administered context. For example when asking about longstanding illness (illcde1-4) instead of saying “Please specify the main illness below” the question could be replaced by “What is your main illness?”
- Making sure adequate interviewer instructions are included as to whether definitions should be read out. Some definitions that were in the WHS paper questionnaire are no longer present. For example, the definition of pharmacy should be included for pharser, as it is on the self-completion.
- It might be helpful to repeat the question stem for longer yes/no lists e.g. those on health services. For example the phrase ‘During the past 12 months have you used...’ should be repeated to make sure respondents are clear about what they are answering and what time frame they should consider.

3.3.1 Summary and recommendations

After discussion with WG it was agreed that a number of the questions mentioned above (denser, chirser, physser, osteoser, pharser, Ilti2) have since been removed from the final version of the questionnaire. This was because of overlap with existing National Survey items and in order to make length savings. Should the questions be reinstated the above recommendations will still apply.

3.4 Comments on specific modules

We suggest that most modules can be retained in their current form with minor changes to certain questions. The WHS is designed for self-completion by respondents and questions are therefore clearly worded and relatively easy to understand. The questions are predominantly factual, asking respondents about the services they have used, diagnosed health problems and their health related behavior so most will transfer relatively well to CAPI.

Recommendations on specific modules are summarised below.

3.4.1 Illnesses and other health problems

This module collects information about whether people are currently receiving treatment for a range of illnesses and health problems. The questions ask about a range of illness like cardiac and respiratory illnesses, mental health, diabetes and other longstanding illness. The list of conditions included is currently under review.

Questions about health conditions (especially mental health conditions) may be perceived as sensitive by respondents. This may lead to underreporting by some respondents. WG could consider putting the entire module questions in CASI Alternatively they could trail the questions in CAPI *and* CASI just at the testing stage to compare responses. However, it may not be possible to detect a significant difference if the sample size used at the test stage is small.

One area queried by WG was whether ‘treated conditions’ or ‘diagnosed conditions’ should be collected in the questionnaire, and what approach was adopted in other National Health Surveys.

- WHS asks directly about ‘treated’ illness, for a range of illnesses including cardiac and respiratory diseases, diabetes.
- SHeS 2013 covers a similar range of illnesses using a series of questions for each illness of interest. The questions ask broadly: whether someone has the illness, whether they have been told this by a doctor (doctor diagnosed) and whether they are using medicines, tablets or pills ...” (treated)¹³.
- HSE 2013 asks about a relatively narrow range of illness – diabetes and highBP. Like SHeS, a series of questions is used for each illness of interest, to establish whether the respondent has been told by a doctor they have an illness, and whether they are currently receiving treatment. Note that the focus of HSE varies from year to year, and may cover a greater range of illness in other years¹⁴.
- In addition all three surveys ask about longstanding illness, and limiting longstanding illness. HSE and SHeS record up to 6 longstanding illnesses, while WHS records the main illness.

¹³ See technical report: <http://www.scotland.gov.uk/Publications/2014/12/6634/downloads> (pp 128-132).

¹⁴ See technical report: <http://www.hscic.gov.uk/catalogue/PUB16076>.

Therefore all three surveys provide a measure of treated illnesses. HSE and SHoS are longer health interviews so can accommodate more questions (in this case to differentiate between treatment and diagnosis). WHS on the other hand is self-completion so the questionnaire has to be short and simple. As the integrated survey also needs to be kept concise retaining the current WHS questions (about treated illness) might be the most suitable option, unless there is a strong argument from data-users for inclusion of the additional items on diagnosis.

In addition to asking about health conditions this module of questions also asks about eyesight, hearing and dental health. WG have proposed asking questions about eyesight and hearing every 2 years. Including these in alternate years would be possible and is unlikely to have large context effects on neighbouring questions. However, the impact on average survey length would be relatively small. We estimate that this sub-module of 4 questions contributes less than one minute to average interview length. We note that WHS data has been used to develop the primary eye care policy and wonder whether collecting data in alternate years would be sufficient to support this (see Appendix A on how WHS data is currently used). Therefore, our current timing estimates are based on these questions being included every year.

The questionnaire contains an item called 'Teeth'. The current question asks whether respondent has 21 or more of their own teeth, or less than 21. A finer breakdown would be useful to explore the link between dental health and health in general. We suggest that this question could be replaced with the Adult Dental Health Survey (ADHS) questions 'Natural' and 'Denture' (see the box below). If space is an issue, WG should consider using just 'Natural'.

Adult Dental Health Survey (ADHS) questions on dental health

{Ask all}

NATURAL [Source= ADHS 2011]

How many natural teeth have you got? Is it...

Running prompt

INTERVIEWER - Include wisdom teeth - adults usually have up to 32 teeth, including the 4 wisdom teeth.

INTERVIEWER - Exclude implants to replace missing teeth

- (1) none at all,
- (2) at least 1 but less than 10,
- (3) between 10 and 19,
- (4) or do you have 20 or more natural teeth?
- (5) Have some natural teeth but don't know how many (SPONTANEOUS ONLY)

{Ask all}

DENTURE [Source= ADHS 2011]

Do you have (require) a denture, even if you don't wear it?

- (1) Yes
- (2) No

Comments and suggested amendments

- Consider the placement of the questions on current health conditions. Some of these items may be better placed in in the CASI.
- 'Teeth' has been replaced with the more detailed ADHS items.

3.4.2 Health and wellbeing (SF-36)

The current WHS questions include the SF36 (36 question short-form). This is a standardised measure of functional physical and mental health included in the survey since 2003/4. The questions cover general health, longstanding and limiting longstanding illness and the impact of physical and mental health on day to day activities.

Alternative options for Health and Wellbeing measures are described in section 3.2. Whatever form the final wellbeing questions take, a CASI mode should be considered. Questions about mental health and wellbeing are likely to be perceived as sensitive by some respondents and a self-completion format may elicit more honest responses.

Comments and suggested amendments

- It is recommended that WEMWBS is always administered in a self-completion format. It has been tested in both paper and CASI self-completion modes.¹⁵
- As the survey includes a CASI module it would be more cost-efficient to include the WEMWBS items in the CASI rather than have it as a separate paper self-completion questionnaire.

3.4.3 Smoking

The WHS questions cover smoking, giving up smoking, where people smoke, whether respondents use e-cigarettes, and their exposure to passive smoke. Other health surveys include questions about smoking, but the information collected varies from survey to survey. On HSE, for example, the smoking questions cover similar topics, but also include more detailed questions to establish the number of cigarettes smoked. The HSE module is considerably longer than the WHS module and may be too long for the integrated questionnaire.

We suggest that the WHS questions are retained in their current format as the smoking module is relatively short and provides the information needed by WG to monitor smoking behaviour.

The WHS has introduced an item on e-cigarettes (this will be included in 2015). This question is asked of all sample members who are currently smokers or who have smoked in the past. This is because although e-cigarettes are seen as a tool for giving up smoking there are reports from HSE of non-smokers using e-cigarettes. We recommend these questions are retained.

¹⁵ See user guide: <http://www.healthscotland.com/uploads/documents/7551-WEMWBS%20User%20Guide%20Version%201%20June%202008.pdf>

Comments and suggested amendments

- We suggest that the WHS questions are retained in their current format. The placement of these questions (CAPI or CASI) can be reviewed further as part of the field test.

3.4.4 Alcohol

The WHS covers alcohol consumption (frequency), and consumption on heaviest drinking day in the last week. We feel that the measures on the heaviest drinking day do not transfer well to CAPI format.

We suggest that the alcohol questions are replaced with the HSE questions which collect the same information, but are designed for CAPI. The HSE module has more items but is heavily routed so respondents are not asked all questions. The WHS questions `alcodrink` to `oth2scan`, would be replaced with the equivalent HSE questions `drinkL7` to `othL7qc` (see Appendix B). The initial questions (`freqalc`, `nodrink`) can be retained in their current form.

Although alcohol consumption questions are asked in other face-to-face surveys there is a danger that moving to a CAPI format could discourage reports of heavy drinking. WG could consider trialling the questions in both CAPI *and* CASI just at the testing stage to compare responses. However, it may not be possible to detect a significant difference if the sample size used at the test stage is small.

Comments and suggested amendments

- We suggest that the WHS questions on alcohol are replaced with HSE equivalent questions (that have been specifically designed for computer based administration as they are heavily routed). These questions can be found in Appendix B.
- WG to consider whether these questions should be included in the CASI section of the interview.
- In addition to asking questions on consumption over the last seven days there is an interest from data-users in measuring 'usual alcohol consumption.' Since 2011 HSE has collected information on 'average weekly consumption in the last 12 months.' The questions used to measure this are included in Appendix B. Based on timestamp data including these questions would add approximately 2.6 minutes to mean interview length. These questions are currently being considered as a potential rotating module.

3.4.5 Fruit and vegetable consumption

The WHS questions on fruit and vegetable consumption are designed to measure the number of portions of fruit and vegetables consumed in the last 24 hours. This is an estimate that is relatively stable from year to year so, if annual data is not required, these questions could be asked in alternate years. They could also be asked of a subsample of respondents.

Although there are other sets of questions that measure fruit and vegetable intake (such as those used on HSE) the other modules are considerably longer. For example, HSE asks two questions for every type of fruit and vegetable: whether the respondent consumed that food and, if so, how much they consumed. This approach would take longer to administer than the current WHS module. Therefore, we think the current approach should be retained.

There may be a social desirability bias which could be increased if questions are asked in CAPI instead of a self-completion. That is, respondents may report higher fruit and vegetable consumption when asked these questions by an interviewer than when answering a self-completion questionnaire. There are two ways to reduce this social desirability effect with the current set of questions:

1) To ask the questions in CASI, although this would be a long module and may not be the best use of a limited CASI resource;

2) To reassure respondents that it's okay to say no – “If you haven't eaten any of the following just say ‘none’”.

Small tweaks in question wording will be required if the questions are asked in CAPI. For example, the current introduction says ‘using the measures below, how much of the following did you eat yesterday?’ If the questions are orally administered no information on measures will be shown ‘below’. Instead information on measures will have to be read out for each question. Additional information could also be provided (based on HSE) so that interviewers could explain the measures if needed.

Comments and suggested amendments

- Consider asking these questions in CASI. If this is not considered a good use of resources interviewers should be prompted to reassure respondents that ‘none’ is a suitable response.
- If the questions are administered in CAPI minor tweaks to the question stem will be required.
- Consider asking these questions as part of a rotating module rather than including them every year if time is limited.

3.4.6 Exercise

This exercise module will be introduced in WHS 2015, to provide physical activity estimates that are more in line with the updated physical activity guidelines. Currently we estimate the exercise question will contribute 1.9 minutes to the average interview length.

This module is based on the short IPAQ (a self-completion questionnaire) and similar to the questions introduced to HSE this year (also in self-completion format). USoc contains a slightly shorter set of exercise questions. However, these questions focus on walking in the last 4 weeks (rather than the last week). These may be of less utility if analysts are interested in assessing whether people are meeting recommended weekly targets on exercise.

The new questions will be less comparable with other surveys, like HSE, if asked in CAPI rather than self-completion. Therefore CASI would be our preferred mode. If a CAPI approach is adopted, a showcard reminding people of the definitions of moderate and vigorous exercise should be introduced.

Some revisions may make the current questions easier to answer, for example allowing respondents to report their time exercising in their preferred units (minutes or hours).

Comments and suggested amendments

- Retain the current questions. If space is at a premium consider asking the questions in alternate years.
- Consider asking these questions in CASI.
- Respondents should be allowed to enter time spent doing an activity in their preferred reference period (minutes or hours).

3.4.7 Other questions (care, BMI and pregnancy variables)

A single question is asked about caring or unpaid care. This can be retained in its current format, with a showcard used to present answer options. The positioning of this question could be reviewed (as it would also sensibly fit in the core module alongside questions on working status).

No changes are required for the additional analysis variables.

4 Welsh Outdoor Recreation Survey (WORS)

The next module reviewed was the items from the Welsh Outdoor Recreation Survey. The aim of the review was:

- To provide length estimations on administration time.
- To identify questions that may be at risk of mode effects.
- To identify other questions that could do with additional refinement. As part of this we explored other sources of suitable questions on outdoor recreation.

4.1 Module content and duration

The original WORS module suggested for the integrated questionnaire contained the following subjects:

1. Outdoor activities in the last 12 months
2. Outdoor activities in the last 4 weeks
3. Details of the most recent visit (activity, location, duration, mode of transport, cost, purpose)
4. 'General questions' (barriers to outdoor activities and latent demand)
5. Biodiversity.
6. Demographics (Car and bike access)

Our old timing assumptions for the WORS modules (as described in the first interim report) were that all the items suggested above add approximately *12 minutes* to mean interview length. Since producing the first interim report the following sections have been cut from the WORS questionnaire:

- Biodiversity
- Extra demographics (car and bike usage).

It has also been clarified that the intention is for the WORS questions to only be asked of a sub-sample of 6,000 participants.

Based on these changes our new timing estimates for the WORS sections are as follows:

Table 4.1: Predicted module length of WORS sections

Module section	Contribution to predicted mean length (mins)
1. Outdoor activities in the last 12 months	0.5
2. Outdoor activities in the last 4 weeks	0.2
3. Details of the most recent visit (activity, location, duration, mode of transport, cost, purpose etc)	2.9
4. 'General questions' (barriers to outdoor activities and latent demand)	1.4
TOTAL	5.1

It is anticipated the full WORS module will be asked every four years. In addition a subsample of questions will be asked every two years. We predict this sub-sample of questions will take approximately 1.8 minutes to administer.

The current aim is for the full WORS module to take no longer than 5.75 minutes to administer and the shortened module to take no longer than 1.75 minutes to administer. Therefore we recommend that the current questions can be retained for the field-test without the need for additional cuts.

4.2 Minimising mode-effects

In the past WORS has been administered via the telephone. One potential issue of changing modes is the risk of introducing 'mode effects' that will disrupt the time-series. Mode effects are changes in the distribution of data that are not attributable real changes in the domain of interest. Instead, they are attributable to differences in the format in which questions are administered. Certain types of question are more at risk of mode effect than others, depending on their content and format.^{16,17} NatCen were asked to review WORS items and to comment on whether any of the proposed WORS items were at risk and how these risks could be ameliorated.

¹⁶ Campanelli, P., Nicolaas, G., Jäckle, A., Lynn, P., Hope, S., Blake, M., & Gray, M. (2011). A Classification of Question Characteristics Relevant to Measurement Error and Consequently Important for Mixed Mode Questionnaire Design. Royal Statistical Society Conference Proceeding. London.

¹⁷ Dex, S., & Gummy, J. (2011). On the experience and evidence about mixing modes of data collection in large-scale surveys where the web is used as one of the modes in data collection. National Centre for Research Methods (NCRM) Review Paper.

The main risk factor when converting CATI questions to CAPI questions is creating mode effects via **changing the perceptual pathway** (e.g. moving from aural presentation to visual presentation). This type of could occur if answer options are changed from being read out over the phone to being displayed on a showcard. Participants will select more options from an aurally presented list (where they have to say Yes / No to each item) compared to a visually presented list where they are asked to 'check all that apply' (CATA). This is thought to be due to the fact participants reading showcards will employ a form of satisficing (short-cutting) by not fully reading the list of answer options (see Dex & Gummy, 2011). The risk of this satisficing (and therefore of mode effects) increases with the number of answer options displayed on the showcard.

We identified a number of items in the WORS questions which we believe are at risk of mode effects occurring. These are former Yes / No readouts that have been converted to CATA questions with showcards. An example of this type of question is Q1a (shown in Table 4.2 overleaf).

We have made recommendations on where we feel showcards should be used and where they should not. Our recommendations are based on:

- Whether or not the options were formally read out (or whether they were interviewer coded from open responses). Options that were read out have a higher risk of mode effects occurring.
- The number of answer options listed (longer lists are higher risk).
- Whether one or multiple answer options that can be coded (multi-coded questions are higher risk).

The second cause of mode effects when converting CATI questions to CAPI is in relation to scalar formats. There is some evidence to suggest that survey respondents give more positive ratings to scale questions by phone compared to other modes¹⁸.. It is unclear what triggers this effect.

Having reviewed the WORS items there is only one scalar question (Q23 on how respondents rate their concerns over biodiversity). This question has now been cut from the questionnaire. Therefore we do not think this risk is present in the questions that are retained.

¹⁸ Campanelli, P., Nicolaas, G., Jäckle, A., Lynn, P., Hope, S., Blake, M., & Gray, M. (2011). A Classification of Question Characteristics Relevant to Measurement Error and Consequently Important for Mixed Mode Questionnaire Design. Royal Statistical Society Conference Proceeding. London

Table 4.2: Example of a question at risk of mode effects:

Question	Risk	Solution
<p>Q1a Which of the following activities have you undertaken at least once during visits to the outdoors in the last 12 months?</p> <p>SHOWCARD</p> <p>Walking Road cycling Off-road cycling or mountain biking Horse riding Fishing Rock climbing or caving Motorised watersports Other watersports Swimming outdoors Snowsports Field sports (shooting/hunting) Air sports (hang gliding/ballooning) Wildlife watching Running Sightseeing or visiting an attraction Off-road driving or motorcycling Picnicking (as a significant part of the activity) Visiting children's playgrounds/taking children to play areas outdoors Informal games (e.g.Frisbee or golf) Other outdoor activities SPECIFY - None of these</p>	<p>Move from Yes/No readout to CATA showcard.</p> <p>The number of activities reported in the last 12 months could decrease if there is move from using a Yes/No readout to a CATA plus showcard approach.</p> <p>As monitoring outdoor activities is the primary purpose of WORS this would be of concern.</p>	<p>Retaining a Yes / No readout approach would ameliorate this risk.</p> <p>This approach would create a slight increase in administration time length.</p>

4.3 Comments on specific sections

In addition to looking at length and potential mode effects we reviewed the questions for general areas of improvement. As part of this we reviewed other questions on outdoor recreation. The main sources of questions we reviewed were:

- The Monitor of Engagement with the Natural Environment (MENE, 2013). This is the English equivalent of WORS, sponsored by Natural England, Defra and the forestry commission.
- The Scottish Recreation Survey (SRS, 2012). This is the Scottish equivalent of WORS, sponsored by Scottish Heritage.

Both the above surveys are conducted face-to-face (as part of a CAPI omnibus). The topic coverage in MENE and SRS was broadly similar to that used in WORS. However, there were individual variations in question wordings and the answer options available.

For example they included variations in terms of:

- The wording of the introduction
- Types of activities engaged in
- Places in which activities were undertaken
- Reasons for participating in activities.
- Reasons for not participating in activities

Table 4.3 overleaf summarises our views on potential revisions for each of the questionnaire sections.

Table 4.3: Additional revisions for WORS module

Section	Potential issue and recommendations
Introduction	The current introduction is vague in relations to what types of outdoor space should be included and which should be excluded. SRS and MENE use slightly different introductions. Revisions and further testing may be beneficial.
Activity lists (Q1a, Q2, Q4, Q4c, Q18)	The answer option 'informal pursuits' could do with further clarification. SRS and MENE use slightly different lists of activities.
Place lists (Q7a, Q7b, Q20, Q21)	There is a conceptual overlap between some answer options listed (e.g. beach, sea, other coastline). SRS and MENE use slightly different lists that do not contain this overlap. Currently the answer options 'Historic sites' and 'Gardens' are listed for Q21 but not the other items. If these are in-scope they should be included in all place lists.
Barriers (Q16, Q16a, Q19)	<p>The original questions on barriers to visiting the outdoors currently are interviewer coded based on respondent's verbatim responses. The list of available codes is currently long and unwieldy. It uses 27 codes many of which are quite similar.</p> <p>We recommend the approach for collecting information on barriers is reviewed. Firstly we recommend that the code list is refined so overlapping codes are removed. In addition we recommend showing the revised code list on a showcard and getting respondents to select all the barriers that apply to them. This approach would encourage respondents to acknowledge and report more barriers than the current question format.</p>

4.4 Summary and recommendations

- We estimate the WORS material flagged for inclusion every two years will add approximately *1.75 minutes* to mean interview length. We estimate the material flagged for inclusion ever four years will add approximately *5.1 minutes* to mean interview length.
- We recommend that showcards are not used for questions that previously relied on a Yes / No readout as this could introduce mode-effects with no gain in quality.
- The list of answer codes could be refined and reduced in places. For example, similar items (sea, beach, coastline) could be combined into a single code. The extent to which this is feasible is dependent on what codes are considered important for analysis. Alterations to the answer options used have been made in consultation with WORS data users. We recommend that showcards are used for questions on barriers that were previously interviewer coded. This could introduce mode effects but would reduce the likelihood of inconsistent

coding practices. The code lists should be refined to minimise the conceptual overlap between answers.

5 Length estimates of further questionnaire sections

This chapter provides length estimations of all sections of the draft questionnaire other than WHS and AAS sections (which are discussed in the preceding chapters) The sections discussed here have been taken from:

1. The National Survey for Wales (NS):
2. The Arts in Wales Survey (AIWS); and
3. The Active Adult Survey (AAS).

Length estimations of each of these will be discussed separately in the following sections.

5.1 National Survey (NS)

The National Survey modules being retained contain a wide variety of topics. The current intention is that some of topics will be asked of everyone in the sample whereas others will only be asked of a sub-sample.

In the second interim report our predicted mean administration time for the additional National Survey questions was approximately 13 minutes.

Since producing the second interim report two items on access to a car and transport behaviour have been added back into the active travel section (WORS D6 and AtUsMode). These items have been added as:

- Car usage was thought to be a useful demographic variable for cross-analysis, but also useful to support future small estimation modelling work.
- To reduce the likelihood of context effects occurring for Active Travel sections.

Other than this the timing estimates have undergone no further alterations (see the table overleaf):

Table 5.1 Estimated length of National Survey modules

Module section	Contribution to predicted length (mins) ¹⁹	Comments on sub-sampling
1. Accommodation	0.5	-
2. Community cohesion	0.3	Asked of a sub-sample of 2,000 (assumes 16.5% of survey sample)
3. Discrimination	1.3	-
4. Museums	0.9	-
5. Active travel (children)	0.5	-
6. Education	0.6	-
7. Play	0.2	-
8. Personal care plan	0.1	-
9. Local health service	0.5	Asked of sub-sample of 4,500 (assumes 37.5% of survey sample)
10. GP services	0.9	Section asked of sub-sample of 6,000 (assumes 50% of survey sample)
11. Hospitals	0.6	Asked of sub-sample of 10,000 (assumes 83.3% of survey sample)
12. Ambulance services	0.6	-
13. Social care services	1.2	-
14. Use of health/ social care in Welsh	0.1	Asked of sub-sample of 2,000 (assumes 16.5% of survey sample)
15. Satisfaction with Local Authority services	1.0	-
16. Democracy and understanding of the local council	0.4	-
17. Future generations	0.3	Asked of a sub-sample of 2,000 (assumes 16.5% of survey sample)
18. Welsh Government	0.2	Asked of a sub-sample of 3,000 only (assumes 25% of the sample)
19. Satisfaction with education	0.1	Asked of a sub-sample of 3,000 only (assumes 25% of the sample)
20. Satisfaction with health services	0.1	Asked of sub-sample of 4,500 (assume 37.5% of survey sample)
21. Internet	1.4	-
22. E-safety	0.2	-
23. Transport and active travel (adults)	1.6	-
TOTAL	13.4	-

¹⁹ Please note length estimates take routing into account where e.g. questions asked of parents only)

Our current prediction is that the NS questions will contribute 13.4 minutes to mean interview length. We think this assumption should be tested during the field-test.

5.2 Arts in Wales Survey (AIWS)

The suggested AIWS module contains sections on the following subjects:

1. **Attendance:** This includes frequency of visiting different arts events (cinema, musicals, operas, theatre, concerts, galleries etc.) and reasons for attendance.
2. **Participation:** This includes participation of art forms (playing an instrument, acting, dance, photography, etc.).
3. **Barriers** to attending art events and activities.
4. **Attitudes to art and culture.**

Since producing the second interim report one item has been removed from the AIWS module (Q9g on attitudes to what constitutes art). It has also been decided that the AIWS sections on participation and attendance will only be asked of a sub-sample of 6,000 participants, and the remaining sections will only be asked of a sub-sample of 3,000 participants.

Based on these changes are new timing estimates for the AIWS sections are as follows.

Table 5.2 Estimates of how each section contributes to predicted average length

Module section	Contribution to predicted average length (mins)	Comments on estimates
1. Attendance	2.0	Section asked of sub-sample of 6,000 (assumes 50% of survey sample)
2. Participation	1.5	Section asked of sub-sample of 6,000 (assumes 50% of survey sample)
3. Barriers	0.2	Asked of a sub-sample of 3,000 only (assumes 25% of the sample)
4. Attitudes to art and culture	0.5	Asked of a sub-sample of 3,000 only (assumes 25% of the sample)
TOTAL	4.2	

5.3 Active Adult Survey (AAS)

The suggested AAS module contains questions on the following subjects:

1. **Participation:** This includes participation in *outdoor* sports and activities, *indoor* sports and activities and *outdoor pursuits*.
2. **Sports club membership:** This includes questions on what types of clubs the participant is a member of, frequency of attendance and receipt of coaching.
3. **Latent demand:** This includes questions on what types of sport or activities participants would like to do more of in the future and what could encourage future participation.
4. **School experience:** This includes two items on whether the participant enjoyed school sport.
5. **Volunteering:** This includes items on all volunteering, sports volunteering (including frequency, type of activity and age-group worked with) and coaching.
6. **Social:** This includes items on perceived behavioural norms about sport and items on how the participants rate their physical abilities in relation to their peers.

Some minor amendments were made to the AAS section prior to the length estimates being calculated. The aim of the alternations was to reduce the length of some the showcards used. The changes made consisted of:

- Reducing the number of activities on the 'indoor activities' showcard in the participation section. Individual types of dance, martial arts and fitness or gym classes are now asked as separate follow-up questions.
- Having three questions to capture latent demand rather than one. One question captures latent demand for outdoor activities, one question captures latent demand for outdoor pursuits and one question captures latent demand indoor activities. To facilitate administration the revised questions on latent demand have been moved to the participation section. Questions on 'encouraging future participation' now form their own sub-section.

Currently, it is thought that the AAS module will be rotated and asked every five years. In total we estimate the entire AAS module will add approximately *10 minutes* to average interview length on the years they are included. The table overleaf provides estimates of how much each section contributes to predicted length.

Table 5.3 Estimates of how each section contributes to predicted average length

Module section	Contribution to predicted average length (mins)
1. Participation and latent demand	3.3
2. Sports club membership	1.2
3. Encouraging future participation	0.8
4. School experience	0.6
5. Volunteering	0.9
6. Social	3.1
TOTAL	9.9

There is an interest in asking the 'participation' questions every year instead of every five years. We estimate these participation questions will add approximately *2 minutes* to mean interview length.

5.4 Overall questionnaire length

The aim is for the integrated questionnaire to have an average administration time of 45 minutes. There is some scope for variation in length by year, provided the average length across a five year period is around 45 minutes.

Based on our length estimations we feel the current questionnaire is roughly the correct length. However, the relative contribution of each module to total length is not yet within target parameters.

5.4.1 Contribution of annual components

The three questionnaire sections that will be included every year are the CORE questions, the National Survey questions and the WHS questions. We estimate that these sections, in their current format, will have a mean administration time of under 40 minutes (see table 5.4.1 below).

Table 5.4.1: Estimated mean length of annual components

Survey component	Subsampling	Rotation	Current estimated contribution to mean length (mins)	Target contribution to mean length (mins)
CORE	None	Every year	11.2	8.5
NS	Some-varies per module	Every year	13.4	16.25
WHS	None	Every year	14.6	14
Total contribution of annual questions			39.2	38.75

Therefore, although the estimated mean length of the annual questions is close to the length target the CORE questions may take longer to administer than was originally intended. Whether or not this occurs in practice should be verified during the field-test. If the CORE questions do contribute more than 8.5 minutes mean interview length than two options should be considered:

- Reducing the target length of other annual components; or
- Reducing the set of annual CORE questions further.

It is our opinion that material deprivations may be a candidate for rotating rather than including annually in the CORE module. Our rationale for this is that these items may be less useful for group comparisons/ cross-tabulations compared to the other CORE items (such as the standardised socio-demographic variables). However, we recommend that a decision regarding this is made in consultation with data-users after the field-test has been carried out and full information on actual administration time is available.

5.4.2 Contribution of rotating modules

The plan is for the WORS, AIWS and AAS questionnaire modules to be rotated, that is not asked every year. Table 5.4.2 below indicates our predictions on mean interview length in years where the full WORS, AIWS and AAS modules are included.

Table 5.4.2a: Estimated mean length of rotating components

Survey component	Subsampling	Rotation	Estimated contribution to mean length (mins)	Estimated survey length once annual components included
WORS	Asked of 6,000	Full module every 4 years	5.1	44.3
AIWS	Some modules asked of 6,000, others asked of 3,000	Full module every 5 years	4.2	43.4
AAS	None	Full module every 5 years	9.9	49.1

As shown in the above table it could be that average interview length is slightly higher than desirable in years when the full AAS module is included. However, it should be stressed that the length estimates we have provided are prone to a margin of error. Furthermore, there is scope for variation in interview length by year, provided the average length across a five year period is around 45 minutes. The table below shows the predicted contributions of each rotating modules to interview length averaged out across a five year period.

Table 5.4.2b: Estimated mean length per year of rotating components over five years

Survey component	Subsampling	Rotation	Estimated contribution to mean length per year over five years (mins)	Target contribution to mean length per year over five years (mins)
WORS	Asked of 6,000	Full module every 4 years	1.3	1.75
AIWS	Some modules asked of 6,000, others asked of 3,000	Full module every 5 years	0.8	1.00
AAS	None	Full module every 5 years	2.0	3.5
Total contribution per year over five years			4.1	6.25

Therefore, before making a decision as to whether further items should be cut from the AAS module we recommend checking the length estimates provided as part of the field-test.

Assuming the length estimates we have provided are correct two years out of five there will capacity to include around 10 minutes worth of additional questions. These could include the shortened version of the AAS (lasting approximately 2 minutes), the shortened version of the WORS module (lasting approximately 1.8 minutes) and then any other ad hoc modules of questions that have a mean administration time of around six minutes.

6 Questionnaire order

After receiving all the proposed questionnaire modules in terms of estimated length NatCen reviewed the overall topic coverage in order to recommend a questionnaire order.

The rationale when suggesting an order was to:

1. Provide a logical narrative or flow (for instance by clustering sections on similar topics together).
2. To start the questionnaire with non-sensitive, universally relevant questions to ease participants into the questionnaire. More sensitive questions should come later in the instrument after the interviewer has had time to develop rapport.
3. The most sensitive items should be included in a separate self-completion (CASI) module. The aim of this is to reduce the risk of inaccurate reporting and question refusals. However, we attempted to keep the number of CASI items to a minimum (to keep the interview streamlined for participants less confident using computers).
4. To have 'rotating modules' or sub-sample modules towards the end of the questionnaire. We did this to reduce the risk of question order effects.

For some questions two or more of the above considerations were mutually contradictory. For example some questions would logically sit within one section because of their subject matter but were also candidates for CASI because of their sensitivity and/or were also flagged as potential candidates for rotating modules. In these cases we have suggested an item positioning on a case-by-case basis.

6.1 Areas of overlap

In addition to looking at the overall ordering of items we have reviewed potential sources of overlap and duplication between the different sections. The duplications identified are described below:

6.1.1 NS and WHS sections on 'use of health services'

There is some conceptual overlap between the NS section on GP services (asked of a sub-sample only) and the WHS sections on health-service use (which are currently flagged as an area for cutting or rotating). The material covered in both sections is summarised below:

Area	Items from NS	Items from WHS (currently cut)
GP service used	<ul style="list-style-type: none"> • Whether used GP services in last 12 months. • Satisfaction with GP services (various domains). 	<ul style="list-style-type: none"> • Whether talked to GP services in last 2 weeks (and how many times). • Whether given a prescription in the last 2 weeks. • Whether seen practice nurse in last 2 weeks (and how many times).

There is also some overlap between the NS section on hospitals (asked of a sub-sample only) and the WHS sections on health-service use (which is currently flagged as an area for cutting or rotating). The material covered in both sections is summarised below:

Area	Items from NS	Items from WHS (currently cut)
Hospital usage	<ul style="list-style-type: none"> • Whether any NHS hospital appointments in the last 12 months. • Whether last appointment was as a day-patient, inpatient or outpatient. • Satisfaction with care received. 	<ul style="list-style-type: none"> • Whether any A&E appointments in last 12 months. • Whether any day-patient appointments in last 12 months (and whether this was NHS or private). • Whether any inpatient appointments in last 12 months (and whether this was NHS or private). • Whether any outpatient appointments in last 12 months (and whether this was NHS or private).

Finally the general scope of the NS sections on other health services is more limited than the scope previously included in WHS (see below):

Area	Items from NS	Items from WHS (currently cut)
Other health services	<ul style="list-style-type: none"> • Whether used <i>ambulance services</i> in the last 12 months. 	<p>Collects information on use of the following services in the last 12 months:</p> <ul style="list-style-type: none"> • Dentists (and whether private or NHS) • Chiropodists (and whether private or NHS) • Physiotherapists (and whether private or NHS) • Osteopaths (and whether private or NHS) • Opticians • Health visitors • GP OOH services • NHS direct • Pharmacists • Reasons for going to the dentist • Flu jabs

We recommended that WG and the consultation group discuss the data needs for these sections to see if there is any way the questions can be combined to meet the data needs of both NS data-users and WHS data users. As a result of these discussions the decision was made to retain the NS questions, as they help in the production of relatively high-profile satisfaction figures, and to drop the WHS questions, as there is not a strong stakeholder demand for these results.

6.1.2 NS and WHS questions on walking

Both NS and WHS contain questions on walking. The NS questions are included in the 'active travel' module and the WHS questions are included in the module 'Exercise'. The scope of the two sections are illustrated below.

Area	Items from NS	Items from WHS
Walking	<ul style="list-style-type: none"> • Whether has walked as a form of transport in the last 3 months. • Frequency walking (daily, weekly etc). 	<ul style="list-style-type: none"> • Amount of time spent walking in last 7 days. • Walking pace. • Whether walking acts as a 'physical activity'.

The purpose of these two sections is quite different (one is asked from a transport perspective and the other is asked from a health perspective). Given these two different remits we do not think the questions need to be combined in terms of improving user-experience.

6.1.3 NS and AIWS questions on 'Arts/Museums/Heritage'

The integrated questionnaire contains a small number of questions from the National Survey on museum visits taken from the section on 'Art/Museums/Heritage'. The rest of the questions on 'Art/Museums/Heritage' from the NS have been dropped for the purposes of the large-scale test, because there are already questions on arts and heritage in the AIWS module.

It is possible that the museum questions could be reframed and combined with the questions used in AIWS on frequency of attending different types of arts performances (cinema, theatre, art gallery). Whether or not this is appropriate is dependent on the needs of data-users. One main difference between the NS questions and the AIWS questions is that for the NS participants are instructed to only consider visits to museums in Wales whereas in the AIWS no such geographical restriction is mentioned. Similarly, in the NS questions participants are prompted to include visits to museum cafes or shops even if they did not visit an actual museum exhibit (MusVisit). If these details are important to NS data-users we recommend that the original NS museum questions are retained in their current format.

6.1.4 AIWS, AAS, WORS

There is a small amount of overlap between the coverage of AIWS, WORS and AAS sections. For example:

- AIWS contains items on frequency of participation in sport and dance in the last 12 months (these areas are covered in more depth in AAS).
- AIWS contains an item on the frequency of visiting the seaside in the last 12 months (this area is also covered in WORS).

However, as WORS, AAS and AIWS items will be rotated (e.g. most participants will not be asked to answer more than one of these modules during the interview) and this overlap is small we do not think the overlap should not have any detrimental effect in terms of user-experience. Therefore the current level of overlap is acceptable in terms of user-experience. The level of overlap between rotating modules should continue to be reviewed each survey wave. If there is any overlap between questions that are asked of the same participants the extent of this overlap should be examined. Data-users should be consulted about any overlapping content and whether overlapping questions could be consolidated whilst still meeting the informational requirements of all parties.

6.2 Questionnaire outline

After discussions with WG the following questionnaire order was agreed prior to the context-review process:

Core – Address check
Core – Household selection
Core – Respondent selection
Core – Household information
Core – Start of respondent interview
Core – Welsh language
Core – Demographics, qualifications, employment
WHS – Carers
Core – Household material deprivation / financial inclusion
Core – Child material deprivation
Core – Tenure
NS – Accommodation
NS – Community cohesion
NS-Discrimination
NS – Active travel – children
NS – Education
NS – Play
Core – Health demographic questions
NS – Personal care plan
WHS – Vision, hearing and teeth
WHS – Smoking
WHS – Fruit and vegetables
NS – Local health services
NS – GP services
NS – Hospitals
NS – Ambulance services
NS – Social care services
NS – Use of health / social care services in Welsh
NS – Satisfaction with local authority services
NS – Democracy and understanding of local councils

NS – Future generations

NS – Welsh Government

NS – Satisfaction with education system

NS – Satisfaction with health services

NS – Internet

NS – e-Safety

NS – Active travel

NS – Arts / museums / heritage

AIWS – Attendance

AIWS – Participation

AIWS – Barriers

AIWS – Attitudes to arts and culture

WORS – Introduction

WORS – Visits in the last 12 months

WORS – Visits in the last 4 weeks

WORS – Most recent visit

WORS – General questions

AAS – Current participation and latent demand

AAS – Sports club membership

AAS – Encouraging future participation

AAS – School experiences

AAS – Volunteering

AAS – Social

CASI WHS – Illnesses and other health problems

CASI WHS – Pregnancy and BMI

HSE – Alcohol consumption [Revised CASI version]

CASI WHS – Exercise

CASI Core – Wellbeing/ ONS4

CASI Warwick-Edinburgh Mental Well-being Scale [WEMWBS]

CASI Core - Sexual orientation and religion

Core – Recontact and linkage

Core – Interviewer check information

The following questionnaire sections were dropped (pending agreement from the consultation group):

WHS – Health service use

WHS – Untreated problems or symptoms

WHS – Wellbeing items (SF36). These have been replaced with the Warwick-Edinburgh Mental Well-being Scale (WEMWBS)

WHS – Medicines

WHS – Alcohol. This has been replaced with the HSE module that captures equivalent information.

WORS – Biodiversity

WORS- Extra demographics

6.2.1 Further research on question order during the field-test

We recommend that some of the assumptions we have made about question order are investigated further as part of the field test later this year. In particular we recommend reviewing what sections should be prioritised for inclusion in the CASI rather than in the CAPI. As described in Chapter 3 there are both advantages and disadvantages of specific sections being included in CASI. This applies to the following sections:

- WHS - Health related behaviours (smoking, alcohol consumption, fruit and vegetable consumption and exercise).
- WHS - Pregnancy and BMI
- Core - Sexual orientation, religion, income.

The advantage of CASI administration is that participants may be less inclined to give socially desirable answers, and may feel more comfortable divulging personal information. The advantage of CAPI is that interviewers can provide assistance in relation to important definitions or provide specific reassurances regarding why information is being collected. CAPI administration requires less effort from participants as they are not responsible for data entry tasks. Therefore, there is a need to minimise the length of the overall CASI module section.

The field-test could investigate how well these questions perform in both CAPI and CASI using a series of split ballot experiments where participants get asked different questions in different modes. Which mode to opt for would be based on:

- Length of module administration in both modes.
- Participants' ratings of 'ease of completion' and 'sensitivity of questions' in both modes. This could be assessed using participant debriefing questions at the end of the interview.
- Whether or not mode has had any impact data distributions across the two conditions.

In general we recommend that the field-test checks the timing assumptions we have made in both the first interim report and in Chapter 1 and that detailed time-stamp data is collected for each questionnaire sub-section.

7 Reviewing the questionnaire for context effects

The next phase of the review involved assessing the new questionnaire to assess which questions, if any, are at risk of context effects caused by the change in their positioning or the change in preceding material. The risks identified in the context effect review are described in the following sections.

7.1 CORE questions

The review of the CORE questions revealed a number of questions at risk of context effects. These risks (and potential solutions) are summarised in Table 7.1 below.

Table 7.1: Risk of context effects within the CORE modules

Questionnaire section	At risk variables	Description of risk and recommendations
Core-Demographics	National identity questions NatEdWel	<ul style="list-style-type: none"> In the current version of the new questionnaire items on 'national identity' are now directly after questions on 'Welsh language ability'. There is a risk that fewer people will describe their identity as 'Welsh' in the new questionnaire if they just answered questions on Welsh language ability (and say they cannot speak Welsh). We recommend moving the questions on 'Welsh language ability' to later in the questionnaire (so they come after 'national identity' items).
	SexOrient Religion	<ul style="list-style-type: none"> These questions are now asked in the self-completion element rather than in CAPI. This means there is less scope for participants to query why this information is being collected. We recommend that the CASI should provide some additional preamble explaining why this information is being collected (because in CAPI this information could be provided by the interviewer). The positioning of this question (CAPI or CASI) should be reviewed as part of the field-testing.

CASI Core-Wellbeing	Questions on wellbeing (ONS4) WbSatLife, WbLifeWrth, WbHapYest, WbAnxYest.	<ul style="list-style-type: none"> • In the 2014-15 NS the ‘wellbeing’ questions were directly preceded by one of the following: Satisfaction with health provision / Satisfaction with education provision / Welsh Government / Democracy. • In the new questionnaire the ‘wellbeing’ questions are directly preceded by the health behaviour items (e.g. smoking, alcohol and exercise). • There is a risk that in the original NS version ‘wellbeing’ will be judged in relation to ‘satisfaction with local’ services and in the new version wellbeing will be judged in relation to ‘satisfaction with health behaviours.’ • We recommend the wellbeing questions are moved to the start of the CASI questionnaire. A short introduction should be added to the start of the CASI to minimise interference from any preceding CAPI questions.
----------------------------	--	--

In addition to the comments related to potential context effects the following comments were made in relation to the CORE items:

- It was suggested that questions on access to the internet (IntHhHave) and access to a car (CarUse) should be considered as CORE questions even if their positioning does not change. This is because these questions will be useful for sub-group analysis across multiple questionnaire domains (satisfaction with local services, participation in outdoor activities, participation in outdoor sports or pursuits, and so on).
- It was suggested that the sections on household material deprivation and child material deprivation could be moved to later on in the questionnaire due to their potential sensitivity. It was felt participants may feel more comfortable answering these items later on in the interview once more rapport with the interviewer has been developed.

7.2 National Survey for Wales

The review of the National Survey questions revealed a number of questions at risk of context effects. These risks (and potential solutions) are summarised in Table 7.2 below.

Table 7.2: Risk of context effects within the NS modules

Questionnaire section	At risk variables	Description of risk and recommendations
Discrimination	DcrimExp (1-12) DcrimExpOth	<ul style="list-style-type: none"> Some items have been added to the questionnaire on 'discrimination'. These follow on directly from the questions on 'community cohesion' included in the National Survey in 2014-15. There is the risk that participants may incorrectly assume that the discrimination questions are about discrimination they have faced within their local community (when in fact are about any form of discrimination encountered in the last 12 months). We recommend a clarification is made to the question stem (to explain we are interested in any discrimination regardless of where it happened). Some introductory text has been added to the questionnaire.
NS- Democracy and understanding of local councils	LocCouKnow LocCouSat	<ul style="list-style-type: none"> In the last version of NS questions on 'understanding of the role of local councillors' were directly preceded by questions on 'recycling.' In the new questionnaire the same questions are directly preceded by items on 'Satisfaction with local authority services.' There is a small risk that in the original version of the questionnaire satisfaction with the local councillor may have been related to satisfaction with recycling facilities, whereas in the new order satisfaction might be related to 'satisfaction with local authority services.'

		<p>Nonetheless we recommend the current questionnaire order is retained, as it seems the most logical order based on the items retained.</p>
NS- Active travel (Adults)	<p>AtFrqBikeOnly</p> <p>AtFrqBikePrt</p> <p>AtFrqWlkOnly</p> <p>AtFrqWlkPrt</p>	<ul style="list-style-type: none"> • A number of the original questions from the 2014-15 NS Active Travel section have been cut from the new questionnaire. This could alter both the comprehension and retrieval of the remaining questions. • In the original NS questionnaire the item 'AtLast7Days' contains an additional readout on what types of journey to exclude: <p>READ OUT: THE QUESTION DOES NOT COVER WALKING AND CYCLING THAT YOU DO PURELY FOR PLEASURE, FOR HEALTH REASONS, FOR TRAINING OR COMPETITION, OR FOR JUST WALKING THE DOG.</p> <ul style="list-style-type: none"> • In the new questionnaire this is included an interviewer instruction but is not explicitly marked as a readout. • We recommend that exclusion categories are marked explicitly as readouts in the new question. • In the original questionnaire the item 'AtUsMode' prompts people to consider multiple modes of transport (car, bus, train cycling, foot etc) and multiple locations (shops, work, GPs etc). This question has now been cut. Therefore recall on types of trips may be better in the original version of the questionnaire. • We recommend the decision to cut 'AtUsMode' is revoked on the basis it's removal could introduce context effects. We have added this item back into the questionnaire.

In addition to the issues related to potential context effects the following comments were made in relation to the National Survey items:

Questionnaire section	Variables	Comments
Multiple sections on satisfaction	All variables using the showcard of answer options 'Very satisfied' through to 'Very dissatisfied.'	<ul style="list-style-type: none"> • It was noted that currently all satisfaction scales are currently displayed as 'Very satisfied' through to 'Very dissatisfied.' • We recommend that the direction of the scale should be randomly rotated between participants in order to minimise the risk of positivity bias via primacy effects.
NS- Accommodation	HSat	<ul style="list-style-type: none"> • HSat asks 'How satisfied are you' whereas Lldat asks 'How satisfied or dissatisfied are you...' In one question satisfaction is inferred but in the other satisfaction is not inferred. • We recommend that all questions are neutrally phrased as far as possible. For satisfaction questions the stem should always be 'How satisfied or dissatisfied are you...'
NS- Play	Introduction and PlayoppU11	<ul style="list-style-type: none"> • The question PlayOppU11 repeats material that has just been read out in the section introduction. The introduction and the question could be contracted.
NS- Ambulance service	AmbSatSkill	<ul style="list-style-type: none"> • It was queried whether the item AmbSatSkill was about trust and confidence regarding ambulance services <i>in general</i> or satisfaction regarding the professionalism the crew who <i>attended</i>. It was felt that the concept being measured was slightly ambiguous.

7.3 Welsh Health Survey (WHS)

The review of the Welsh Health Survey questions revealed a number of sections at risk of context effects. These risks, and potential solutions, are summarised in Table 7.3 below.

Table 7.3: Risk of context effects within the WHS modules

Questionnaire section	At risk variables	Description of risk and recommendations
WHS – Carers	carerhrs	<ul style="list-style-type: none"> In the new questionnaire the question on care provision (carerhrs) is located in an early demographic section. In WHS this question comes after a suite of questions on health. There is a small chance that participants may think more broadly about types of care provided when answering the WHS question compared to answering the item in the new questionnaire. However, we do not think this risk in itself is enough to justify a change in item location. It should be noted that this question was used in the 2011 Census without the prior material contained in WHS.
WHS – Smoking	Likegup Guphlthp Guphlthg Gupreill	<ul style="list-style-type: none"> The WHS smoking module contains four questions on attempts to give up smoking. In the original WHS questionnaire the smoking questions came after a module of questions on illnesses and health problems. Therefore in the original WHS questionnaire participants may be more primed to think of the negative health consequences of smoking. Participants may be less likely to report failed quit-attempts in a face-to-face interview. Therefore, we recommend that the WHS-smoking section is moved to the CASI section. It should be placed after the questions on illnesses and health problems (assuming these are retained).
HSE – Alcohol consumption (CASI)	DrinkL7	<ul style="list-style-type: none"> In the new questionnaire the questions on alcohol consumption now come soon after the question on pregnancy (pregw). It was felt that pregnant women (who have drunk in the last week) may be less likely to answer the alcohol questions honestly in the new order of questions. We recommend the question on pregnancy is moved to a later part of the CASI questionnaire.

In addition to the comments on context effects the following observations were made in relation to the revised WHS sections.

Questionnaire section	Variables	Comments
WHS – Vision, hearing and teeth	All	<ul style="list-style-type: none"> • Currently there is no introductory material to this section. The move from the previous section (Personal Care plans) to this section is quite abrupt. • We recommend that a short introduction to this section is added.
WHS – Fruit and vegetable consumption	All	<ul style="list-style-type: none"> • Then fruit and vegetable questions have been adapted slightly to make them more suitable for computerised administration. We recommend that these revisions are tested as part of the field-test. • Currently it is undecided whether these items would be better asked in CAPI or CASI.
HSE – Alcohol	All	<ul style="list-style-type: none"> • The alcohol questions used have been taken from HSE (as these are more tailored to computerised administration). • The HSE questions have been adapted slightly to allow for self-completion via CASI. For example, in HSE drink sizes are coded by interviewers (with assistance provided at the briefing and via interviewer help-screens). In the revised CASI version there are some new participant instructions related to different wine glass sizes and beer sizes (pints, bottles or cans). • We recommend that these revisions are tested to ensure that participants understand what information should be recorded where. • Currently it is undecided whether these items would be better asked in CAPI or CASI. As discussed in 6.2 variations in CAPI and CASI placement could be investigated as part of the field-test.

7.4 Welsh Outdoor Recreation Survey (WORS)

No potential context effect risks were raised in relation to the positioning of the WORS questions within the new questionnaire.

It was noted that the scripting for Q15a (below) is currently unclear.

{ASK IF 1 – 20 AT Q2}

Q15a {SOURCE=WORS}

During this visit, how much money did you personally spend any money on any of the following items?

SHOWCARD

Food and drink
Petrol\diesel\LPG
Car parking
Bus\train\ferry fares
Hire of equipment
Purchase of equipment
Maps\guidebooks\leaflets
Gifts\souvenirs
Admission fees
Other items
Didn't spend any money
Don't know [SPONTANEOUS]

Is the intention of this question to find out whether *any money* was spent on each item, how much money was spent on *each item* or how much money was spent *in total*? We have assumed the aim is get an estimate of how much money was spent in total and adjusted the questionnaire wording accordingly.

7.5 Arts in Wales Survey (AIWS)

The review of the AIWS questions revealed a number of questions at risk of context effects. These risks are summarised in Table 7.5 below.

Table 7.5: Risk of context effects within the AIWS modules

Questionnaire section	At risk variables	Description of risk and recommendations
AIWS- Attendance	All	<ul style="list-style-type: none"> • The two questions immediately preceding the AIWS questions (MusVisit and MusSat taken from the NS section on Arts/Museums/Heritage) explicitly state to only include museum visits in Wales. Participants may incorrectly assume this exclusion also applies to the AIWS questions on participation. • As described previously WG should consider whether the NS-Museum questions could be merged with the AIWS questions and whether the exclusion criteria they have previously used are still required. • If the data-needs of the NS-Museum questions remain unchanged the AIWS section on participation would benefit from a new introduction. This introduction should clarify that visits both in Wales and elsewhere should be included.

It was also noted that the scripting for Q8a (below) is currently unclear.

{ASK ALL}

Q8 {AIWS 2010, Q11C}

Below is a list of things which people say prevent them from attending or participating in arts events or activities. Which, if any, prevent you from going to or taking part in the following at all or more often than you currently do...?

SHOWCARD. ROTATE ORDER. MULTICODE

1. It is difficult to find the time
2. It costs too much
3. Family commitments/childcare
4. My health is not good enough
5. Lack of access/facilities for disabled people

6. I had a bad experience the last time I went
7. Don't know
8. Other (specify)

Is the intention of this question to find out barriers to specific types of activities (as was the case in AIWS 2010 when a list of activities was also listed at this question)? Or is it intended to be a single question to assess barriers to *all* types of art events?

Based on feedback from AIWS data-users it was agreed that the aim of this question should be to collect information on barriers to all types of art event, and the questionnaire has now been updated accordingly. It is recommended that this question could be re-tested to ensure that participants continue to think about all types of art event when answering.

Finally it was noted that the attitude question Q9g may not be collecting accurate assessments of what constitutes as art (see below).

{ASK ALL}

Q9g {AIWS 2010}

When I think about the arts, I think about activities like going to the theatre, listening to classical music or visiting an art exhibition.

ROTATE ORDER. SINGLE CODE

1. Strongly agree
2. Agree
3. Neither agree nor disagree
4. Disagree
5. Strongly disagree
6. Don't know

Currently this question is asked after participants have already answered questions on a range of different art forms. Therefore, participants will already have been prompted to consider theatre, classical music and art exhibitions as art forms prior to the question being asked. Based on feedback from AIWS data-users this question has now been cut.

7.6 Active Adults Survey (AAS)

No potential context effect risks were raised in relation to the positioning of the AAS questions within the new questionnaire. However, the following comments were raised in relation to the AAS volunteering section.

Table 7.6: Further comments on Active Adult Survey Modules

Questionnaire section	Variables	Comments
AAS volunteering section	QPART46	<p>It was felt that the wording of the QPART46 (on any volunteering) was somewhat ambiguous and should be reviewed.</p> <p>We recommend replacing it with the following item which has been adapted from the 'Taking Part' survey.</p> <p>[ASK ALL] VOLWORK {Source= Adapted from the Taking Part survey} I would now like to ask you about any voluntary work you may have done. During the last 12 months, have you done any voluntary work?</p> <p>READ OUT: This could be organising or helping to run an event or club, coaching, mentoring, campaigning, conservation, raising money, providing transport or driving.</p> <p>INTERVIEWER: THIS COULD INCLUDE SPORT OR NON-SPORT RELATED VOLUNTEERING</p> <p>1. Yes 2. No</p>
	QPART47	<p>It was felt that the wording of the QPART47 (on any volunteering related to sport) should also be reviewed.</p> <p>In particular it was felt that the exclusion criteria (on not including transporting someone to an activity) may not always be read out by interviewers. This could result in levels of volunteering being over-estimated. We recommend that the exclusion criteria should be made a 'read out' rather than left to the interviewer's discretion.</p>

8 Findings from test interviews

In the final phase of the project a small number of test interviews were conducted in-office to assess the overall flow of the instrument.²⁰ Four interviews were conducted with interviewees being routed to different sub-sections in each interview. All test interviews took between 40 minutes and 50 minutes to complete.

In general positive comments were made on the overall order of each interview, and interviewees had no comments regarding improving the order of topics. Two main issues were raised in relation to the general flow of the questionnaire. These were:

1. Inconsistent use of answer categories; and
2. Inconsistent use of reference periods.

A number of small queries were also raised regarding specific questions. All of these areas are described further in the following sections.

8.1 Consistency of answer options

Currently the five-point 'agree-disagree' answer scale is phrased in four different ways throughout the questionnaire, depending on the original source material. For example, in some scales the term 'agree' is used for the second scale option, in others 'tend to agree', in others 'somewhat agree' and in others 'agree slightly.' It was felt that the flow of the interview would be improved if a consistent scale was used throughout.

It was also noted that some 'agree-disagree' modules started with the general instruction: 'To what extent do you agree or disagree with each of these statements' whereas other modules (for example the NS Community cohesion) repeated 'To what extent do you agree or disagree...' at the start of each question asked. The people interviewed felt the latter approach felt more repetitive and long-winded.

We recommend a consistent wording should be used for all agree-disagree scales. We have updated the questionnaire accordingly. We have opted for the version used in the National Survey for Wales (tend to agree) as this format was most prevalent in the draft questionnaire. Agree-disagree sections of the questionnaire should start with the introduction 'To what extent do you agree or disagree with each of the following' and then have a list of statements.

Another consistency issue noted was in relation to the labelling of answer options on showcards. Currently, some showcards label each answer option with a letter (A, B, C etc) whereas others label each answer option with a number (1, 2, 3 etc). Prior research conducted by NatCen indicates that participants have no preference between showcards labelled alphabetically or numerically although a consistent approach to showcard labelling within an interview is preferred. As a small number of the questions proposed have more than 26 answer options (i.e. those used in the Active Adult Survey

²⁰ Please note: Interviews started at the 'start of the respondent interview' section of the questionnaire. Interviewer administrated items and householder selection items were not included in the test interviews.

that list types of activity) we recommend that answer options on showcards are always labelled with numbers rather than letters.

8.2 Consistency of reference periods

The other flow issue described during the test interviews was that the reference periods used in the questions change at several points. For example, some of the questions ask about the last 12 months, others ask about the last three months and others ask about the last four weeks. Interviewees described having to concentrate in order to follow which time-period was being discussed.

Despite this we recommend that the reference periods used within each of the questions need to be retained in order to maintain the time-series data. We recommend that the variation in reference periods should be discussed during interviewer briefings, and interviewers should be trained to emphasis the reference period at each question.

8.3 Additional comments

A small number of additional issues were raised about specific questions during the test interviews. These issues are summarised in the table below.

Table 8.3: Additional comments raised during the test interviews

Questionnaire section	Question reference	Issues raised	Recommendation
WHS- Vision hearing and teeth	hearaidt	The question hearaidt starts with the clarification 'If you usually wear a hearing aid.' However, now only people who usually wear a hearing aid are routed to this question. Therefore this introduction is superfluous and does not flow well.	Remove the text 'If you usually wear a hearing aid' from the question.
NS- GP Services	Section introduction	The section is introduced with the statement 'As I mentioned earlier please think about the NHS when answering these questions.' However, only a sub-sample of people will have asked the previous questions.	Remove the text 'As I mentioned earlier' as not all respondents will have been asked the earlier questions.
NS- Democracy and understanding of local councils	LocCouKnow and LocCouSat	LocCouKnow and LocCouSat are two attitudinal items on whether people feel they understand what their local councillor does and whether people feel their local councillor works closely with the local community. Currently these questions are only	Make these two items 'Ask all' rather than 'Ask if LocCouCont= 1.Yes.'

		asked if the respondent has had any contact with their local councillor in the last 12 months (if LocCouCont= 1.Yes). However, interviewees felt people may have an opinion on their local councillor even if they had not had any contact with them in the last 12 months. They felt these questions should be asked of everyone.	
AIWS- Attendanace	Q2	Interviewees were unclear what performance was being referred to at Q2 (the most recent performance).	Reverse the order of Q2 and Q3.
WORS	Introduction	The introduction says 'in this survey.' This sounds odd as the module is towards the end of the survey.	Replace 'in this survey' with 'The next questions are about...'

The latest version of the questionnaire has been updated with the above recommendations.

Appendix A. How WHS data is currently used

WHS estimates are widely used by WG, local government, NHS and third sector organisations. WG, for example, uses WHS data to develop policies, to set targets and monitor improvements in health, to inform its public health outcomes framework, and to monitor health and health inequalities. Local government use WHS to inform local health needs assessments that are used to prioritise, plan and commission services and programmes. Public Health Wales Observatory also produce a range of outputs, such as sub local authority results, reports on specific topics like alcohol and health, tobacco and health, and health inequalities.

Examples of how WHS data are used are shown in Table A1 overleaf.

Table A1: Examples of how the WHS data are used

	WHS content	Uses of WHS data
	All content	<p>Target setting and monitoring: Our Healthy Future; Fairer Health Outcomes for All; Tackling Poverty Action Plan; likely to be used in the combined outcomes framework for Communities First, Families First and Flying Start.</p> <p>Policy Development: e.g. Carer strategy, primary eye care policy</p> <p>NHS performance monitoring: e.g. delivery plans for major conditions</p> <p>Needs assessment in Townsend NHS resource allocation formula; local health needs assessments which inform planning, priorities and allocation of services</p> <p>Reports and information: Chief Medical Officer's annual report, with focus on ill health and health inequalities; information and general context for speeches, briefings, media etc.</p> <p>As basis for further outputs: Public Health Wales Observatory (PHWO) have generated a range of outputs including: sub LA analysis of results; profiles of lifestyle and health in Wales; measuring inequalities and reports on tobacco, alcohol, obesity and cardiovascular disease.</p>
1	Health services	Used in general reports, resource allocation and performance monitoring described above
2	Medicines	Used in general reports, resource allocation and performance monitoring described above
3	Illnesses and other health problems	Used in general reports, resource allocation and performance monitoring described above PHWO cardiovascular disease website
4	Untreated problems or symptoms	
5	Health and wellbeing	Used in general reports, e.g. CMO report, PHWO profile of lifestyle and health
6	Smoking	Programme for Government; Tobacco Control Action Plan; smoking cessation targets; identifying local areas with highest smoking prevalence for targeted campaigns; informing local policy development on tobacco control; Public Health Wales Observatory and Welsh Government report on Tobacco and health in Wales ²¹
7	Alcohol	Programme for Government; Public Health Wales Observatory report on Alcohol and Health in Wales ²²
8	Fruit and vegetable consumption	
9	Physical activity	Programme for Government; Creating an Active Wales
10	BMI	Programme for Government; Public Health Wales Observatory Obesity web resource

²¹ See <http://www.wales.nhs.uk/sitesplus/922/page/59800>

²² See <http://www.wales.nhs.uk/sitesplus/922/page/75229>

Appendix B. HSE – Alcohol consumption

We have suggested that WHS questions about alcohol consumption in the last week (**alcodrink** to **oth2scan**) are replaced with the equivalent HSE questions. The wording and routing of the HSE items are shown below.

Alcohol consumption in the last week

Drink

I am now going to ask you a few questions about what you drink - that is if you drink. Do you ever drink alcohol nowadays, including drinks you brew or make at home?

- 1 Yes
- 2 No

{IF Drink = No}

DrinkAny

Could I just check, does that mean you never have an alcoholic drink nowadays, or do you have an alcoholic drink very occasionally, perhaps for medicinal purposes or on special occasions like Christmas and New Year?

- 1 Very occasionally
- 2 Never

{IF DrinkAny = Never}

AlwaysTT

Have you always been a non-drinker or did you stop drinking for some reason?

- 1 Always a non-drinker
- 2 Used to drink but stopped

{IF AlwaysTT = Used to drink but stopped}

WhyTT

Did you stop drinking because of a particular health condition that you had at the time?

INTERVIEWER: IF RESPONDENT SAYS PREGNANCY, CODE YES

- 1 Yes
- 2 No

{IF (Drink = Yes) OR (DrinkAny = very occasionally)}

DrinkOfft

SHOW CARD J1

Thinking now about all kinds of drinks, how often have you had an alcoholic drink of any kind during the last 12 months?

- 1 Almost every day
- 2 Five or six days a week
- 3 Three or four days a week
- 4 Once or twice a week
- 5 Once or twice a month
- 6 Once every couple of months
- 7 Once or twice a year
- 8 Not at all in the last 12 months

{IF DrinkOfft <> Not at all in the last 12 months}

DrinkL7

Did you have an alcoholic drink in the seven days ending yesterday?

- 1 Yes
- 2 No

IF DrinkL7 =Yes THEN

DrnkDay

On how many days out of the last seven did you have an alcoholic drink?

Range: 1..7

IF DrnkDay = 2 to 7 days THEN

DrnkSame

Did you drink more on *one of the days/some days than others*, or did you drink about the same on *both/each of those days*?

1 Drank more on one/some day(s) than other(s)

2 Same each day

WhichDay

Which day *last week* did you *last have an alcoholic drink/have the most to drink*?

1 Sunday

2 Monday

3 Tuesday

4 Wednesday

5 Thursday

6 Friday

7 Saturday

DrnkType

SHOW CARD I2

Thinking about last (*answer to WhichDay*), what types of drink did you have that day?

CODE ALL THAT APPLY

1 Normal strength beer/lager/cider/shandy

2 Strong beer/lager/cider

3 Spirits or liqueurs

4 Sherry or martini

5 Wine

6 Alcopops/pre-mixed alcoholic drinks

7 Other alcoholic drinks

8 Low alcohol drinks only

IF DrnkType = Normal strength beer/lager/cider/shandy THEN

NBrL7

Still thinking about last (*answer to WhichDay*), how much normal strength beer, lager, stout, cider or shandy (excluding cans and bottles of shandy)

did you drink that day? INTERVIEWER: CODE MEASURES THAT YOU ARE GOING TO USE.

1 Half pints

2 Small cans

3 Large cans

4 Bottles

IF NBrL7=Half pints THEN

NBrL7Q(1)

ASK OR CODE: How many half pints of normal strength beer, lager, stout, cider or shandy (*excluding cans and bottles of shandy*) did you drink that day?

Range: 1..97

IF NBrL7Q = Small cans THEN

NBrL7Q(2)

ASK OR CODE: How many small cans of **normal strength beer, lager, cider or shandy** did you drink that day?

Range: 1..97

IF NBrL7=Large cans THEN

NBrL7Q(3)

ASK OR CODE: How many large cans of **normal strength beer, lager, cider or shandy** did you drink that day?

Range: 1..97

IF NBrL7=Bottles THEN**NBrL7Q(4)**

ASK OR CODE: How many bottles of **normal strength beer, lager, cider or shandy** did you drink that day?

Range: 1..97

NBotL7

ASK OR CODE: What make of **normal strength beer, lager, stout, cider or shandy** did you drink from bottles on that day? INTERVIEWER: IF RESPONDENT DRANK DIFFERENT MAKES CODE WHICH THEY DRANK MOST.

Text: Maximum 21 characters

IF DrnkType = Strong beer/lager/cider THEN**SBrL7**

Still thinking about last (*answer to WhichDay*), how much **strong beer, lager, stout or cider** did you drink that day? INTERVIEWER: CODE MEASURES THAT YOU ARE GOING TO USE.

1 Half pints

2 Small cans

3 Large cans

4 Bottles

IF SBRL7=Half pints THEN**SBrL7Q(1)**

ASK OR CODE: How many half pints of **strong beer, lager, stout or cider** did you drink on that day?

Range: 1..97

IF SBrL7=Small cans THEN**SBrL7Q(2)**

ASK OR CODE: How many small cans of **strong beer, lager, stout or cider** did you drink on that day?

Range: 1..97

IF SBrL7=Large cans THEN**SBrL7Q(3)**

ASK OR CODE: How many large cans of **strong beer, lager, stout or cider** did you drink on that day?

Range: 1..97

IF SBrL7=Bottles THEN**SBrL7Q(4)**

ASK OR CODE: How many bottles of **strong beer, lager, stout or cider** did you drink on that day?

Range: 1..97

SBotL7

ASK OR CODE: What make of **strong beer, lager, stout or cider** did you drink from bottles on that day? INTERVIEWER: IF RESPONDENT DRANK DIFFERENT MAKES CODE WHICH THEY DRANK MOST

Text: Maximum 21 characters

IF DrnkType = Spirits THEN**SpirL7**

Still thinking about last (*answer to WhichDay*), how much spirits or liqueurs (such as gin, whisky, brandy, rum, vodka, advocaat or cocktails) did you drink on that day?

CODE THE NUMBER OF SINGLES - COUNT DOUBLES AS TWO SINGLES.

Range: 1..97

IF DrnkType = Sherry THEN

ShryL7

Still thinking about last (*answer to WhichDay*), how much sherry or martini, including port, vermouth, Cinzano and Dubonnet did you drink on that day?

CODE THE NUMBER OF GLASSES.

Range: 1..97

IF DrnkType = Wine THEN

WineL7

Still thinking about last (*answer to WhichDay*), how much wine, including Babycham and champagne, did you drink on that day?

INTERVIEWER: please note that respondent may give answer in bottles and glasses. Please code the relevant option.

INTERVIEWER: CODE THE MEASURE THE RESPONDENT USED

1 Bottle or parts of bottle

2 Glasses

3 Both bottles or parts of bottle, and glasses

IF WineL7= 1 (Bottles or part of bottle)

WL7Bt

INTERVIEWER: code the number of 125ml glasses drunk **from the bottle** by the respondent. E.g. If they drank half a bottle, code 3 glasses. Press <F9> for information

CODE THE NUMBER OF GLASSES.

1 BOTTLE =6 GLASSES

½ BOTTLE=3 GLASSES

1/3 BOTTLE=2 GLASSES

¼ BOTTLE=1.5 GLASSES

1 LITRE=8 GLASSES

½ LITRE=4 GLASSES

1/3 LITRE=2.5 GLASSES

¼ LITRE=2 GLASSES

Range: 1..97 (ALLOW FRACTIONS)

F9 for WL7Bt

If respondent has answered in bottles or litres convert to glasses using the information provided on the screen. For example if a respondents said they shared a bottle with one other person and they shared it equally code 3 glasses.

IF WineL7= 2 (Glasses)

WL7Gt

CODE THE NUMBER OF GLASSES (**drunk as glasses**).

Range: 1..97 (ALLOW FRACTIONS)

WL7Gtz

Were you drinking from a large, standard or small glass?

INTERVIEWER: If respondent drank from two or three different size glasses, please code all that apply.

INTERVIEWER: please note that if respondent was drinking in a pub or wine bar and had a small glass, this would usually be 175ml.

1 Large glass (250mL)

2 Standard glass (175 mL)

3 Small glass (125 mL)

IF DrnkType = Alcopops/pre-mixed alcoholic drink

PopsL7

Still thinking about last (*answer to WhichDay*), how much **alcoholic soft drink** ('alcopop') did you drink on that day? INTERVIEWER: CODE

MEASURES THAT YOU ARE GOING TO USE

- 1 Small cans
- 2 Stanadard Bottles (275ml)
- 3 Large Bottles (700ML)

IF PopsL7 = Small cans THEN

PopsL7Q(1)

ASK OR CODE: How many small cans of **alcoholic soft drink** ('alcopop') did you drink on that day?

Range: 1..97

IF PopsL7= standard sized Bottles THEN

PopsL7Q(2)

ASK OR CODE: How many standard bottles of **alcoholic soft drink** ('alcopop') did you drink on that day?:

Range: 1..97

IF PopsL7= LargeBottles THEN

PopsL7Q(3)

ASK OR CODE: How many large bottles of **alcoholic soft drink** ('alcopop') did you drink on that day?:

Range: 1..97

IF DrnkType=Other THEN

OthL7TA

Still thinking about last (*answer to WhichDay*), what other type of alcoholic drink did you drink on that day? CODE FIRST MENTIONED ONLY.

Text: Maximum 30 characters

OthL7QA

How much (*name of 'other' alcoholic drink*) did you drink on that day? WRITE IN HOW MUCH. REMEMBER TO SPECIFY HALF PINTS/ SINGLES/GLASSES/ BOTTLES.

Text: Maximum 30 characters

OthL7B

Did you drink any other type of alcoholic drink on that day?

- 1 Yes
- 2 No

IF OthL7B=Yes THEN

OthL7TB

Still thinking about last (*answer to WhichDay*), what other type of alcoholic drink did you drink on that day? CODE FIRST MENTIONED ONLY.

Text: Maximum 30 characters

OthL7QB

How much (*name of 'other' alcoholic drink*) did you drink on that day? WRITE IN HOW MUCH. REMEMBER TO SPECIFY HALF PINTS/SINGLES /GLASSES/ BOTTLES.

Text: Maximum 30 characters

OthL7C

Did you drink any other type of alcoholic drink on that day?

- 1 Yes

2 No

IF OthL7C=Yes THEN

OthL7TC

Still thinking about last (*answer to WhichDay*), what other type of alcoholic drink did you drink on that day?

CODE FIRST MENTIONED ONLY

OthL7QC

How much (*name of 'other' alcoholic drink*) did you drink on that day?

WRITE IN HOW MUCH. REMEMBER TO SPECIFY HALF PINTS/ SINGLES/ GLASSES/ BOTTLES.

Text: Maximum 30 characters

There is also an interest in collecting information on 'usual consumption' as part of a rotating module. The questions used by HSE to measure this are shown in the box below.

Usual weekly alcohol consumption in the last 12 months

{IF DrinkNow = 1 or DrinkAny = 1}

Intro I'd like to ask you whether you have drunk different types of alcoholic drink in the last 12 months. I'd like to hear about ALL types of alcoholic drinks you have had.

If you are not sure whether a drink you have had goes into a category, please let me know. I do not need to know about non-alcoholic or low alcohol drinks.

(PRESS <F9> AT FOLLOWING QUESTIONS FOR MORE INFORMATION ABOUT WHAT SHOULD BE INCLUDED AT THE DIFFERENT DRINKS CATEGORIES).

NBeer

SHOWCARD J1

I'd like to ask you first about normal strength beer, lager, stout, cider or shandy which has less than 6% alcohol. How often have you had a drink of normal strength beer, lager, stout, cider or shandy (excluding cans and bottles of shandy) during the last 12 months?

(NORMAL = LESS THAN 6% ALCOHOL BY VOLUME.)

<F9> FOR INFO ON DRINKS TO BE INCLUDED HERE.

- 1 Almost every day
- 2 5 or 6 days a week
- 3 3 or 4 days a week
- 4 once or twice a week
- 5 once or twice a month
- 6 once every couple of months
- 7 once or twice a year
- 8 not at all in last 12 months

{IF NBeer = 1 – 7}

NBeerM

How much NORMAL STRENGTH BEER, LAGER, STOUT, CIDER or SHANDY (excluding cans and bottles of shandy) have you usually drunk on any one day during the last 12 months?

FIRST CODE TYPE OF MEASURE AND THEN CODE NUMBER OF EACH MEASURE. CODE ALL THAT APPLY.

1. Half pints
2. Small cans
3. Large cans
4. Bottles

NBeerQ

How many (half pints/ small cans/ large cans/ bottles) of NORMAL STRENGTH BEER, LAGER, STOUT, CIDER or SHANDY (excluding cans and bottles of shandy) have you usually drunk on any one day during the last 12 months?

Range 1..97

{IF Drinknow = 1 or DrinkAny = 1}

SBeer

Now I'd like to ask you about STRONG BEER OR CIDER which has 6% or more alcohol (eg Tennant's Extra, Special Brew, Diamond White). How often have you had a drink of strong BEER, LAGER, STOUT or CIDER during the last 12 months? STRONG=6% AND OVER ALCOHOL BY VOLUME. USE HELP SCREEN FOR OTHER DRINKS TO BE INCLUDED HERE.

- 1 Almost every day
- 2 5 or 6 days a week
- 3 3 or 4 days a week
- 4 once or twice a week
- 5 once or twice a month
- 6 once every couple of months
- 7 once or twice a year
- 8 not at all in last 12 months

{IF SBeer = 1 – 7}

SBeerM

How much STRONG BEER, LAGER, STOUT or CIDER have you usually drunk on any one day during the last 12 months?

FIRST CODE TYPE OF MEASURE AND THEN CODE NUMBER OF EACH MEASURE.

1. Half pints
2. Small cans
3. Large cans
4. Bottles

SBeerQ

For each type of measure of strong beer ASK OR RECORD, How many (half pints/ small cans/ large cans/ bottles) of STRONG BEER, LAGER, STOUT or CIDER have you usually drunk on any one day during the last 12 months?

Range: 1..97

{IF Drinkknow = 1 OR DrinkAny = 1}

Spirits

SHOWCARD J1

How often have you had a drink of SPIRITS OR LIQUEURS, such as gin, whisky, brandy, rum, vodka, advocaat or cocktails during the last 12 months?

<F9> FOR OTHER DRINKS TO BE INCLUDED HERE.

- 1 Almost every day
- 2 Five or six days a week
- 3 Three or four days a week
- 4 Once or twice a week
- 5 Once or twice a month
- 6 Once every couple of months
- 7 Once or twice a year
- 8 Not at all in last 12 months

{IF Spirits = 1 – 7}

SpiritsQ

How much SPIRITS OR LIQUEURS, such as gin, whisky, brandy, rum, vodka, advocaat or cocktails have you usually drunk on any one day during the last 12 months? CODE THE NUMBER OF SINGLES -

Range: 1..97

{IF Drinkknow = 1 or DrinkAny = 1}

Sherry

SHOWCARD J1

How often have you had a drink of SHERRY OR MARTINI including port, vermouth, Cinzano and Dubonnet, during the last 12 months?

<F9> FOR OTHER DRINKS TO BE INCLUDED HERE.

- 1 Almost every day
- 2 Five or six days a week
- 3 Three or four days a week
- 4 Once or twice a week
- 5 Once or twice a month
- 6 Once every couple of months
- 7 Once or twice a year
- 8 Not at all in last 12 months

{IF Sherry = 1 – 7}

SherryQ

How much SHERRY OR MARTINI, including port, vermouth, Cinzano and Dubonnet have you usually drunk on any one day during the last 12 months? CODE THE NUMBER OF GLASSES

Range: 1..97

{IF Drinknow = 1 or DrinkAny = 1}

Wine

SHOWCARD J1

How often have you had a drink of WINE, including Babycham and champagne, during the last 12 months?

<F9> FOR OTHER DRINKS TO BE INCLUDED HERE.

- 1 Almost every day
- 2 Five or six days a week
- 3 Three or four days a week
- 4 Once or twice a week
- 5 Once or twice a month
- 6 Once every couple of months
- 7 Once or twice a year
- 8 Not at all in last 12 months

{IF Wine = 1 – 7}

WineQ

How much WINE, including Babycham and champagne, have you usually drunk on any one day during the last 12 months? CODE THE NUMBER OF GLASSES.

INTERVIEWER: code the number of 125ml glasses drunk from the bottle by the respondent. E.g. If they drank half a bottle, code 3 glasses. Press <F9> for information

CODE THE NUMBER OF GLASSES.

1 BOTTLE =6 GLASSES

½ BOTTLE=3 GLASSES

1/3 BOTTLE=2 GLASSES

¼ BOTTLE=1.5 GLASSES

1 LITRE=8 GLASSES

½ LITRE=4 GLASSES

1/3 LITRE=2.5 GLASSES

¼ LITRE=2 GLASSES

Range: 1..97

BWineQ2

SHOW CARD J3

Were those mainly ...READ OUT...

IF RESPONDENT USUALLY DRINKS IN A PUB OR WINE BAR AND HAD A SMALL GLASS, THIS WOULD USUALLY BE 175ML.

Small Glasses (approx. 125ml)

Standard (approx. 175ml)

Or Large Glasses (approx. 250ml)

Bottles (Spontaneous Only)

{IF Drinknow = 1 or DrinkAny = 1}**Pops**

SHOWCARD J1

How often have you had a drink of ALCOPOPS (i.e. alcoholic lemonade, alcoholic colas or other alcoholic fruit-or-herb-flavoured drinks for e.g. Smirnoff Ice, Bacardi Breezer, WKD, Metz etc), during the last 12 months?

- 1 Almost every day
- 2 Five or six days a week
- 3 Three or four days a week
- 4 Once or twice a week
- 5 Once or twice a month
- 6 Once every couple of months
- 7 Once or twice a year
- 8 Not at all in last 12 months

{IF Pops = 1 – 7 }**PopsLY11**

How much ALCOPOPS or pre-mixed alcoholic drinks (i.e. alcoholic lemonade, alcoholic colas or other alcoholic fruit-or-herb-flavoured drinks) have you usually drunk on any one day during the last 12 months?

INTERVIEWER: CODE THE MEASURE(S) YOU ARE GOING TO USE.

Small cans

Standard Bottles (275ml)

Large Bottles (700ml)

{IF PopsLY11 = Small cans}**PopsQ11[1]**

ASK OR CODE: How many small cans of alcopops or pre-mixed drink have you usually drunk on any one day?

Range: 1..97

{IF PopsLY11=standard Bottles}

PopsQ11[2]

ASK OR CODE: How many standard sized bottles of alcopops or pre-mixed drink have you usually drunk on any one day?

Range: 1..97

{IF PopsLY11= large Bottles}

PopsQ11[3]

ASK OR CODE: How many large bottles of alcopops or pre-mixed drink have you usually drunk on any one day?

Range: 1..97

Appendix C. Copy of coder briefing instructions

P11748: Coding the questionnaire for context effects risks:

Coder briefing instructions

- The Welsh Government (WG) and its partner organisations such as Arts Council of Wales, Sport Wales, and Natural Resources Wales currently commission a number of large-scale social surveys in Wales.
- The Welsh Government wish to combine five of these existing surveys into a single survey. These five surveys that are being brought together are:
 1. The National Survey for Wales (NS)
 2. Welsh Health Survey (WHS)
 3. Active Adults Survey (AAS)
 4. Arts in Wales Survey (AIWS)
 5. Wales Outdoor Recreation Survey (WORS)
- NatCen has been commissioned to review the first draft of the new questionnaire prior to piloting taking place.
- One element of this review is to assess the new questionnaire for risks of *context effects*.

What is a context effect?

- The new questionnaire is mostly made up of questions that have been asked before (i.e. in one of the five Welsh surveys that are now being combined).
- However, just because a question has been asked before (and the wording has not changed) is not a guarantee that it will be understood in the same way between surveys.
- When completing a questionnaire survey participants do not view each individual question in isolation. Instead, the order in which questions are asked, and the prior material covered, can influence how questions are answered. The same questions can elicit different response distributions if the preceding material is different. This is an example of a *context effect*.
- The purpose of this review is to look at the new questionnaire to assess which questions, if any, are at risk of context effects caused by the change in their positioning or the change in preceding material.

Overview of the new questionnaire²³

The new questionnaire contains the following sections. The sections labelled 'core' are classification sections. The core sections contain material that is common to a number of different surveys (e.g. harmonized demographic questions).

The rest of the sections contain substantive questions material taken from one of the five existing surveys. Each section is labelled with which survey the questions have been taken from (NS, WHS, AIWS, WORS or AAS).

Core – Address check
Core – Household selection
Core – Respondent selection
Core – Household information
Core – Start of respondent interview
Core – Welsh language
Core – Demographics, qualifications, employment
Core – Household material deprivation / financial inclusion
Core – Child material deprivation
Core – Tenure
NS – Accommodation
NS – Community cohesion
NS- Discrimination
NS – Active travel – children
NS – Education
NS – Play
Core – Health demographic questions
NS – Personal care plan
WHS – Health service use
WHS – Medicines
WHS – Untreated problems or symptoms
WHS – Health and wellbeing
WHS – Smoking

²³ Note: The coder instructions shown here are a copy of the original documents used. A small number of items were dropped after the coder briefing and information on this was disseminated to coders by email. For a list of the questionnaire order reviewed please see section 6.2 of this report.

WHS – Alcohol
WHS – Fruit and vegetables
WHS – Exercise
WHS – Carers
NS – Local health services
NS – GP services
NS – Hospitals
NS – Ambulance services
NS – Social care services
NS – Use of health / social care services in Welsh
NS – Satisfaction with local authority services
NS – Democracy and understanding of local councils
NS – Future generations
NS – Welsh Government
NS – Satisfaction with education system
NS – Satisfaction with health services
NS – Internet
NS – e-Safety
NS – Active travel
NS – Arts / museums / heritage
AIWS – Attendance
AIWS – Participation
AIWS – Barriers
AIWS – Attitudes to arts and culture
WORS – Visits in the last 12 months
WORS – Visits in the last 4 weeks
WORS – Most recent visit
WORS – General questions
AAS – Participation
AAS – Sports club membership
AAS – Latent demand
AAS – School experiences
AAS – Volunteering

AAS – Social

CASI WHS – Illnesses and other health problems

CASI WHS – Additional questions for cross-analysis (Height, weight, pregnancy)

CASI Core – Wellbeing

CASI CORE - Income

CASI CORE - Sexual orientation and religion

Core – Recontact and linkage

Core – Interviewer check information

- Coder 1 and 2 will be reviewing all the sections labelled NS and AIWS for context effects.
- Coder 3 and 4 will be reviewing all sections labelled WHS, WORS and AAS for context effects.
- All coders will be looking at questionnaire sections *independently* of one another. Please do not confer with other coders who are looking at the same sections as you.

Task overview

1. Read all of the questions in your first assigned section of the new questionnaire.
2. Compare the section against the equivalent section in one of the old questionnaires (NS, WHS, AIWS, WORS and AAS). *Remember to also double-check the final items of the immediately preceding section.*
4. Decide whether you think *any* of the questions in the section are at risk of context effects using the code-frame provided overleaf.
5. Record the variable names of the 'at risk' questions in the recording sheet provided. For each variable list record 'yes' or 'no' to each type of risk factor. Write in an open description of the potential impact of the context effect.
6. Repeat the process with the next assigned section. All recording sheets should be completed and saved in the project folder by COP on Thursday 15 January 2015.

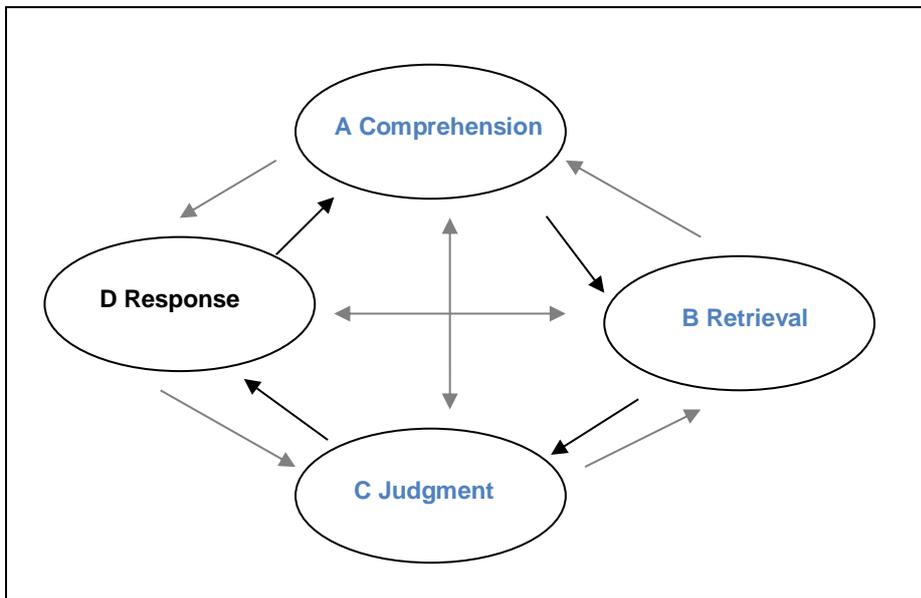
Code-frame: Risk of context effects

There are three risk factors we would like you to assess when reviewing the new questionnaire for possible context effects. These are:

- 1. Possible changes in comprehension.*
- 2. Possible changes in retrieval.*
- 3. Possible changes in judgment.*

These risks are based on three areas based on Tourangeau’s 1984 Question and Answer Model of question response (see Figure 1 below). Further details on how to assess risk on each of these areas is described in the following pages.

Figure 1: Elaborated Question and Answer Model



1. Change in comprehension

Could there be any change in question comprehension due to a change in preceding material?

- Could any of the questions be understood differently due to a change in how the section is introduced?
- Could any of the questions be understood differently due to a change in definitions provided by preceding questions?

Example

<u>Target question</u>	<u>Old question order</u>	<u>New question order</u>	<u>Impact</u>
QX. How many times did you visit your GP in the last 4 weeks?	<p>Q1. How many times did you visit your GP in the last 12 months? Include either seeing a doctor or a practice nurse.</p> <p>Q2. How many times did you see a GP in the last 4 weeks?</p>	<p>Q1. How many times did you visit a library in the last 4 weeks?</p> <p>Q2. How many times did you see a GP in the last 4 weeks?</p>	<p>Definition from a preceding question is no longer present.</p> <p>Comprehension of 'GP visits' could be affected.</p> <p>Fewer GP visits may be reported in the new question order (as it is no longer clarified that practice nurse appointments should be included).</p>

2. Change in retrieval

Could there be any change in information retrieval due to a change in preceding material?

- Are there any retrieval / memory prompts that were present in the old question order but not the new question order? This could include prompts from previous questions or introductory material.

Example

<u>Target question</u>	<u>Old question order</u>	<u>New question order</u>	<u>Impact</u>
QX. How many portions of fruit or vegetables did you eat yesterday?	<p>Q1. Which of the following did you eat yesterday? Tick all that apply.</p> <ul style="list-style-type: none"> • Fresh fruit • Tinned or cooked fruit • Dried fruit • Fruit juices or smoothies • Salad • Beans or pulses • Cooked vegetables • Other dishes containing mainly vegetables (such as soups, curries or stews) <p>Q2. How many portions of fruit or vegetables did you eat yesterday?</p>	<p>Q1. How many units of alcohol did you drink yesterday?</p> <p>Q2. How many portions of fruit or vegetables did you eat yesterday?</p>	<p>Retrieval cues/ memory prompts from a preceding question are no longer present.</p> <p>Recall of fruit and veg eaten could be affected.</p> <p>Fewer fruit and vegetable portions eaten may be reported in the new question order.</p>

3. Change in judgment: Contrast and consistency effects

Could there be any change in question adjustments due to a change in preceding material?

- These types of context effect could occur in relation to *attitude questions*. These include:
 - Agree-disagree questions,
 - Satisfaction questions,
 - Rating questions (good-poor. 0-10 scale, etc).
- People may rate things in different ways depending on what they are *comparing* them to. This can take the form of a contrast effect (where people try to make the two answers different to one another) or a consistency effect (where people try to make the two answers similar to one another).

Example 1: Contrast

<u>Target question</u>	<u>Old question order</u>	<u>New question order</u>	<u>Impact</u>
<p>QX. On a scale from 0-10 how enjoyable do you find your current job?</p>	<p>Q1. On a scale from 0-10 how enjoyable did you find your first job?</p> <p>Q2. On a scale from 0-10 how enjoyable do you find your current job?</p>	<p>Q1. On a scale from 0-10 how enjoyable did you find your last holiday?</p> <p>Q2. On a scale from 0-10 how enjoyable do you find your current job?</p>	<p>In old question order participants could contrast their current job to their first job.</p> <p>In the new question order participants could contrast their current job to their last holiday.</p> <p>Overall ratings of current job enjoyment could therefore p1 differ between the two questions (e.g. higher in the old order).</p>

Example 2: Consistency

<u>Target question</u>	<u>Old question order</u>	<u>New question order</u>	<u>Impact</u>
<p>QX. On a scale from 0-10 OVERALL how satisfied are you with the current state of the NHS in Wales?</p> <p><i>0 is not satisfied at all and 10 is highly satisfied.</i></p>	<p>Q1. On a scale from 0-10 how satisfied are you with the free NHS prescription available in Wales?</p> <p><i>0 is not satisfied at all and 10 is highly satisfied.</i></p> <p>Q2. On a scale from 0-10 OVERALL how satisfied are you with the current state of the NHS in Wales?</p> <p><i>0 is not satisfied at all and 10 is highly satisfied.</i></p>	<p>Q1. On a scale from 0-10 how satisfied are you waiting times for getting a GP appointment?</p> <p><i>0 is not satisfied at all and 10 is highly satisfied.</i></p> <p>Q2. On a scale from 0-10 OVERALL how satisfied are you with the current state of the NHS in Wales?</p> <p><i>0 is not satisfied at all and 10 is highly satisfied.</i></p>	<p>In old question order participants may be prompted to think about a positive feature of the Welsh NHS prior to answering the item on overall satisfaction.</p> <p>In the new question order participants may be prompted to think of a negative experience prior to answering the item on overall satisfaction.</p> <p>Overall ratings of satisfaction could differ between the two questions (e.g. higher in the old order).</p>

Important note:

- Some of the questions in the new questionnaire have not been asked previously in any of the five Welsh surveys named. These questions do not need to be reviewed as part of this exercise.
- Similarly, some of the questions have had their wording adapted since the last time the question was asked. The purpose of this exercise is to record risks brought about by changes in positioning, *not* by direct changes in question wording.

Appendix D. Coder recording sheet

Context effect: Coder recording sheet

Coder name: INSERT



Questionnaire section	Any risk of context effects detected in section? (Yes/No)	List all 'at risk' variable names	Risk of change in comprehension? (Yes/No)	Risk of change in retrieval? (Yes/No)	Risk of change in judgment? (Yes/No)	Description of impact	Any other coder comments