



Understanding and addressing underrepresentation in a postal survey of social care users

Nick Smith and Juliette Malley

DP2843

Acknowledgements

This is an independent report commissioned and funded by the Policy Research Programme at the Department of Health under the Quality and Outcomes of Person-Centered Care Research Unit's) Engagement stream. The views expressed in this report are those of the authors and are not necessarily those of the Department of Health.

The engagement stream also funded a more general review of engaging seldom heard populations in research (Beadle-Brown et al., 2012). This review draws upon some of the findings of that review.

Contents

Acknowledgements.....	1
Executive summary.....	5
Introduction	10
Potential causes of underrepresentation	12
Excluded from the sampling frame	13
Failure to receive survey request.....	14
Inability to respond to survey request.....	14
Refusal to participate	15
Summary	16
Methods for minimising nonresponse.....	17
Methods for encouraging participation.....	17
Payment and incentives	17
Confidentiality, anonymity and trust.....	18
Following up nonrespondents.....	19
Summary.....	21
Methods for Enabling Participation	21
Consent and recruitment	21
Adapting the questionnaire.....	22
Alternative modes of data collection	23
Summary.....	26
Alternative approaches when participation in not possible.....	27
Proxy instruments for gathering quality of life or outcomes data	28
Alzheimer’s Disease-Related Quality of Life (ADRQL)	32
Camberwell Assessment of Need for the Elderly (CANE).....	34
EQ-5D.....	36
DEMQOL-Proxy	38
Quality of Life in Alzheimer’s disease (QoL-AD)	40
Quality of Life Assessment Scale (QOLAS).....	42
Evaluation of Quality of Life Instrument (EQLI).....	43
Quality of Life Interview Schedule (QUOLIS).....	45

The Comprehensive Quality of Life Scale – Intellectual Disabilities (COM QOL-ID)	46
Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39)	48
Stroke-Specific Quality of Life (SS-QOL)	50
Summary.....	50
Observational Tools.....	56
Dementia Care Mapping (DCM)	59
Short Observation Framework for Inspectors (SOFI)	64
The Therapeutic Environment Screening Survey for Nursing Homes (TESS-NH).....	64
Observable Indicators of Nursing Home Quality Instrument (OIQ)	66
The Adult Social Care Outcomes Toolkit – Care Homes version (ASCOT-CH3)	69
Section summary	72
Discussion and recommendations for the Adult Social Care Survey.....	75
Appendix 1: Methods for the systematic review of proxy tools	83
Appendix 2: Methods for the systematic review of observational tools	84
References	85
Table 1 Key characteristics of the proxy tools reviewed	30
Table 2: Details of information found regarding each of the proxy tools.....	31
Table 3: Summary of the question types and proxy requirements of the reviewed tools	51
Table 4: Summary of the question types and proxy requirements of the reviewed tools	53
Table 5: A summary of the characteristics of the reviewed observational tools.....	58
Table 6: Summary of the types of documentation available for each instrument	59
Table 7: Summary of the characteristics of the observational tools, in terms of requirements around usage and testing of validity, reliability and sensitivity	73
Box 1: Key features of the Adult Social Care Survey (ASCS)	11
Box 2: Types of proxy questions	29
Box 3: Sample questions from ADRQL.....	32
Box 4: Sample questions from CANE	34
Box 5: Sample questions from EQ-5D.....	37
Box 6: Sample questions from DEMQOL and DEMQOL-Proxy	39
Box 7: Sample questions from QOL-AD	41
Box 8: EQLI Domains	44
Box 9: Sample items from EQLI.....	44

Box 10: Sample questions from COM QOL-ID	46
Box 11: Sample questions from SAQOL	48
Box 12: DCM8 behaviour categories and codes	60
Box 13: DCM8 personal detraction and enhancer types.....	61
Box 14: Sample items from TESS-NH	65
Box 15: TESS-NH Domains.....	66
Box 16: Sample items from OIQ.....	67
Box 17: Sample ASCOT ratings guidance	70
Figure 1 Factors affecting nonresponse to postal surveys	82

Executive summary

1. Evidence from postal surveys, such as the Adult Social Care Survey (ASCS), is increasingly being used in England to monitor and improve the performance of social care services.
2. Past experience of these surveys shows that certain groups of social care users (e.g. younger, non-white and mental health users) are under-represented due to their reduced propensity to respond to the survey request. In addition, people with cognitive and memory impairments, such as intellectual disabilities (PWID) and dementia, are also likely to be under-represented due to difficulties with consenting to take part and completing a postal survey.
3. Underrepresentation can undermine the usefulness of the data for performance improvement and monitoring. For example, it can create perverse incentives, such as staff not focusing on those persons excluded from the survey, and distort management priorities because managers lacking information about the excluded groups may focus only on improvements for the groups that are included in survey data.
4. This report is organised into three sections. In the first section, we begin with a discussion of the potential causes of underrepresentation in postal surveys. In the second part of the report, we look at methods of improving representation in surveys, such as the ASCS. Finally, we draw together and discuss the findings of the review and present recommendations for the ASCS.
5. The literature on underrepresentation and nonresponse to postal surveys discusses a number of reasons certain groups tend to be underrepresented which can be categorised as being excluded from the sampling frame, failure to receive the survey, being unable to respond and refusing to respond.
6. Of key importance for the ASCS, given the nature of the survey population, certain disabilities appear to make a person more likely to be a nonrespondent. In particular, adults who have cognitive or memory impairments, such as dementia or learning disabilities, are highlighted as being particularly susceptible to being excluded from the sample or, when included, finding that they are unable to respond.

7. An interesting omission from the literature reviewed was the recognition that it was not just learning or cognitive disabilities that can impair a person's ability to respond, but that physical disabilities too can create real challenges for responding to postal surveys.
8. A clear problem for the ASCS is identifying the extent of underrepresentation, particularly of ethnic groups and people with certain types of disabilities.
9. This review has highlighted that it is not only the characteristics of users but the behaviour of the organisation delivering the survey that can affect response rates. Indeed, most theories of survey participation posit that different sets of influences act on the sample members to determine their likelihood of participation; non-participation is not a constant attribute of a person. Response propensity is rather the result of the interaction between the attributes of the survey, the attributes of the sample member and their environment.
10. Research on encouraging people to participate in postal surveys suggests a number of methods to counter nonresponse. Clear messages emerge around the effectiveness of following up nonrespondents with additional mailings to improve response rates, although it is not clear whether this improves response rates from hard-to-reach groups.
11. Lack of trust was an important reason for survey nonresponse but evidence around ways of building trust is more limited. Whilst there is good evidence that university sponsorship and personalising materials can improve response rates, these methods are not particularly helpful for the ASCS.
12. There is good evidence that incentives can help response rates especially amongst certain seldom-heard groups, such as younger people and some black and minority ethnic (BME) groups. However, the review found that unconditional prepaid incentives are far more effective than conditional incentives. There are concerns, though, that incentives are a subtle form of coercion.
13. The review found four key methods that can be used to enable participation in postal surveys. Firstly, by recognising that the study information section of a postal survey can be a barrier to participation if it does not appear to be relevant to the respondent; secondly, by adapting the questionnaire to reflect the 'individualised' needs of the participant, such as versions in other languages and Easy Read versions

for adults with learning disabilities; thirdly, by employing alternative methods of data collection, such as face-to-face or telephone interview, alongside the postal survey; and fourthly, by encouraging or even providing support to complete the survey.

14. The most solid evidence is around the provision of alternative methods. Both face-to-face and telephone interviews have higher response rates than postal surveys and, in the case of face-to-face interviews, there is evidence that they are more effective for engaging people with physical impairments.
15. There is also evidence around the value of adapting questionnaires to make them more understandable to certain groups. Adaptations tend to be focused on either versions being made available in other languages or in Easy Read versions for adults with learning disabilities.
16. The review also highlighted a lack of research into the effectiveness and effects of having someone to help answer the questionnaire.
17. Two potential alternative approaches were identified where participation was not possible: asking either a carer or family member to act as proxy for the service user and rate their quality of life; and using observers to rate service users' quality of life. Both approaches are examined in this review.
18. The review identified 11 quality of life instruments that had either been developed as proxy instruments or had been used with proxies. The instruments developed for eliciting outcomes or quality of life information from proxies tended to be condition-specific, for example focusing on stroke or Alzheimer's disease. The bulk of the instruments found were specific to older people with dementia but there were also studies where proxies were used to elicit outcomes for people with intellectual disabilities and adults who had suffered a stroke
19. The majority of the tools use evaluation-based questions. The researchers appear not have considered the theoretical problems associated with this choice and largely formulated questions that asked proxies to answer as if they were the recipient or to answer how they think the recipient would have answered had he/she been able. This approach is likely to be extremely challenging and prone to error as the proxy is unlikely to utilise the same judgement criteria as the recipient would have used. The lack of consideration given to the format of the questions is reflected in the development of the instruments. Very few of the instruments reviewed underwent any significant development work with proxy respondents. The ADRQoL and

DEMQOL are exceptions since both were designed with proxy input and were extensively tested with proxies.

20. A clear picture emerges from the literature; that there are differences between proxy and self-reports, and that proxy reports should not be seen as substitutes for self-report. Some tools acknowledge this and make it clear that, when using the tool, both proxy and self-report instruments should be administered, rather than the proxy being purely a substitute for the service user.
21. The extent of proxy-recipient agreement depends on the characteristics of both the proxy and the user, and can vary according to the subject of the question. Thus, family members are usually found to report lower quality of life than is self-reported whereas professionals often report higher levels of quality of life than is self-reported.
22. Proxy tools do offer insight into the quality of life experienced by service users whose impairments would make participation in a postal survey impossible. It is clearly important to recognise that proxy reports can never be a substitute for self-reports and that this has implications for the analysis and interpretation of the data gathered. However, this should not be seen as a barrier to adopting proxy tools as a way of gathering the experiences of people who are unable to provide their own views in a structured format.
23. The review identified five quality of life observation based instruments. Not surprisingly, all of the tools we found were designed for use within residential settings such as care homes, nursing homes and sheltered housing.
24. Most of the observational tools reviewed in this paper were designed for people with dementia, although there was a split between tools that attempted to measure an individual resident's quality of life and those which look at rating a residential setting. Those that focused on rating the setting tended to be more highly structured than those that look at the individual service users. Moreover, those that focused upon the service users tended to require training prior to use.
25. Since observation is a much more expensive activity than sending out questionnaires, it is clear that any tool used for the ASCS would need to be short and require preferably none, or at least minimal, training. The shortest and least demanding tools are those that focus on the care environment. However, a tool that focused on the care environment would only be applicable for generating ratings in

care homes or other institutional settings. It would not be suitable for people living in their own homes.

26. The review suggests nine recommendations in order to counter nonresponse and improve the representativeness of the ASCS:

- a) Improve the data held by CASSRs on disability of clients.
- b) The focus of future ASCS development should be on methods for gathering data from people with cognitive impairments.
- c) Revise the Easy Read version of the ASCS.
- d) Develop a strategy to provide support to service users to complete the ASCS.
- e) Improve the quality of the translated versions of the ASCS.
- f) The introduction of stricter guidelines for Local Authorities around following up nonrespondents.
- g) Consider greater use of alternative modes of data collection, with the caveat that extensive use be introduced only following a randomised trial to investigate the effect of mode on responses.
- h) CASSRs should use their position to improve response rates.
- i) Consider the greater use of incentives, with the caveat that extensive use be introduced only following a randomised trial to investigate the effect on responses and response rates.

Introduction

Evidence from postal surveys, such as the Adult Social Care Survey (ASCS), is increasingly being used in England to monitor and improve the performance of social care services. Past experience of these surveys shows that certain groups of social care users (e.g. younger, non-white and mental health users) are under-represented due to their reduced propensity to respond to the survey request (Department of Health, 2001; Department of Health, 2002; Department of Health, 2003; Department of Health, 2004; The NHS Information Centre, 2006; The NHS Information Centre, 2008; The NHS Information Centre, 2009; The NHS Information Centre, 2010). Moreover, in the 2011 ASCS, people who lacked the capacity to consent¹ to take part in the survey were excluded to meet the requirements of the ethics committee, meaning that people with dementia and adults with learning disabilities were underrepresented. Underrepresentation can undermine the usefulness of the data for performance improvement and monitoring of services. For example, it can create perverse incentives, such as staff not focusing on those persons excluded from the survey, and distort management priorities, because lacking information about the excluded groups, managers may focus only on improvements for the groups that are included in survey data.

A further concern is that response rates to the user surveys are not particularly high, in the region of 40 to 60 per cent depending on the focus of the survey, in terms of client group and service category (Department of Health, 2001; Department of Health, 2002; Department of Health, 2003; Department of Health, 2004; The NHS Information Centre, 2006; The NHS Information Centre, 2008; The NHS Information Centre, 2009; The NHS Information Centre, 2010). Although these response rates are good for postal surveys, they do appear to be declining. For example, response rates for the older people's home care survey have declined from 61 per cent in 2003 to 58 per cent in 2006 and to 53 per cent in 2009 (Department of Health, 2003; The NHS Information Centre, 2006; The NHS Information Centre, 2009). Falling response rates may further undermine the usefulness of the data collected since they give the impression that the data is of poor quality and is likely to be biased due to the large number of nonrespondents. As various studies have shown, this is not necessarily the case since bias only occurs where the factors that influence a sampled

¹ In order to assess capacity to consent the following criteria were used to ensure that service users understood what they are agreeing to do in relation to the survey:

- Does the service user understand that the survey seeks their views about the quality of their life and their services?
- Does the service user understand that lots of people will be asked these questions and they are not being singled out?
- Does the service user understand that these questions are being asked to understand how happy people are with their care and support services and assess their experiences of local care services?

person's propensity to respond also influence the survey estimate of interest. Importantly, therefore, the degree of bias is dependent on both the pattern of missingness in the data as well as the extent of missing data (Groves, 2006; Rubin, 1976) and means that bias may be present even where response rates are high if the survey has an unfortunate pattern of missing data (Bootsma-van der Wiel et al., 2002). Nevertheless, this means that any unfortunate pattern of missingness to the ASCS that causes bias is likely to be compounded by the high levels of nonresponse.

The aim of this report is twofold: to review the methods that can be used to address under-representation in postal surveys by aiding, encouraging or enabling the participation of those who have difficulty responding or choose not to respond to postal surveys; and to review methods that can be used to gather data about people who are currently excluded from surveys, like the ASCS, because they lack the capacity to consent to participate. Specifically, we examine the use of proxy respondents and observational methods. Since the review was undertaken to support improvements in the representativeness of the ASCS, we focus on evidence and tools about quality of life from the health and social care literature.

Box 1: Key features of the Adult Social Care Survey (ASCS)

- The ASCS is conducted by 151 of the 153 local councils in England with adult social services responsibility (CASSRs), following national guidance on sampling, data management and collection. The guidance is written by the NHS Information Centre (NHSIC) and reviewed by the Social Services User Survey Group (SSUSG) which has representation from CASSRs, the Department of Health, academics and other interest groups.
- Samples are drawn by CASSRs from their client record database. The sample is drawn from the population of publicly-funded social care service users, and includes people with a variety of disabilities and problems and receiving a range of services. The population comprises of people with sensory and physical disabilities, learning disabilities (or intellectual impairments), mental health problems and / or substance misuse problems, and also a small group of people designated as 'vulnerable', which includes asylum seekers and benefit claimants. Services received include residential services (e.g. personal care only homes and nursing care homes) as well as community-based services, such as home care and day centres, and other forms of low level or one-off support, such as equipment, transport and meals.
- All CASSRs are required to conduct the survey as a postal survey. The NHSIC provides alternative versions, including an Easy Read version for people with learning disabilities and a version for residents of care homes, the latter of which has minor modifications to the content (replacing 'home' with 'care home') and information sheet. The NHSIC also

provides a version in some languages other than English, although these versions have not undergone forward- and back-translation to ensure the sense is maintained.

- CASSRs are allowed some room to diverge from the requirements to conduct a postal survey to enable participation of hard-to-reach groups and to follow up nonrespondents. The NHSIC provides an interview script for face-to-face and telephone interviews for this purpose and provides guidance around gaining informed consent.
- CASSRs are required to send out only one reminder to each nonrespondent.
- The process is slightly different for residents of care homes. CASSRs also send a letter to the care home manager to gain their support before sending the questionnaire to the service users.
- CASSRs do not receive additional funds to conduct the ASCS, despite its mandatory nature. CASSRs are expected to find resources from within their total social care budget.

This report is organised into three sections. The first section begins with a discussion of the potential causes of underrepresentation in postal surveys to provide some background to the review of methods to encourage or enable participation. This discussion draws on a systematic review carried out by the policy research unit in Quality and Outcomes of person centred care (QORU) looking at engaging seldom-heard populations in research (Beadle-Brown et al., 2012) and a non-systematic review conducted for the analysis of nonresponse within the ASCS and UES (Malley and Fernandez, 2012). In the second part of the report we look at methods of improving representation in surveys, such as the ASCS. This is split into two subsections: The first subsection looks at minimising nonresponse by using methods to either encourage or enable participation. This discussion draws on an Office of National Statistics report (Williams and Betts, 2010), the systematic review carried out by Beadle-Brown et al. (2012) and an additional very rapid review of the literature to fill in the gaps. The second subsection looks at alternative data collection approaches for instances where users are unable to participate due to cognitive, intellectual or communication difficulties. We have focused specifically on the use of proxy respondents and observational methods, and have conducted two systematic reviews of the literature to identify those tools that could be applied to assess the quality of life of adult social care users. The methods used for these reviews can be found in Appendices 1 and 2. Finally, we draw together and discuss the findings of the review and present recommendations for the ASCS.

Potential causes of underrepresentation

Broadly, we can think of underrepresentation in postal surveys as having two causes: a failure to include population sub-groups within the sampling frame and a failure on behalf of the sampled person to respond. In postal surveys the causes of failing to respond can be divided into three categories: failure to receive the survey request (for example, because of non-delivery or interception by another and failure to forward onto the intended recipient), refusal to participate, and inability to respond (for example, due to illiteracy in English,

physical or cognitive impairments). Due to nature of postal surveys it is rarely possible to determine within which of these three categories non respondents fall. Nevertheless, the categorisation provides a conceptual framework within which to discuss the literature.

Excluded from the sampling frame

The *researcher's assumptions* about the competencies of certain groups can have a significant influence on research design and can lead to their exclusion from the sampling frame (Nind, 2009; Proctor, 2001). Researchers may exclude people because they assume they are not coherent or lucid enough to express a view (Coucill et al., 2001; Nind, 2009; Proctor, 2001). For example, Aldridge (2007) suggests that people with learning disabilities may not be included as they do not 'fit' with researchers' ideas of what a participant should be and how they should respond. Similarly, a study comparing US and British researchers' attitudes to the inclusion of minority populations in research found that many researchers working in the field of asthma did not feel the inclusion of ethnic minorities in their studies was worth the extra effort they felt this entailed (Sheikh et al., 2009).

Researchers may also be concerned about '*fragmented accounts*', where data is incomplete for some reason, and how they should interpret gaps and inconsistencies in the data (Lloyd et al., 2006; Proctor, 2001). Although the term fragmented account is used in the literature to discuss the issue in relation to more qualitative research, it is clear that it is equally relevant to quantitative research methods, such as postal surveys, where there may be concerns about high levels of missing data. Groups with cognitive impairments, such as adults with learning disabilities and people with dementia, may return questionnaires with large numbers of uncompleted items. Missing data can create difficulties at the analysis stage and researchers may decide that, given the costs of collecting data and the sample size needed for analysis, it is not worthwhile to include such groups.

A further issue may be determining the extent of exclusion where the *population is difficult to define*. For example, categorisations of ethnic minority groups tend to be quite broad – British Asian, Black British and so on – which can make it difficult to identify whether there is underrepresentation of ethnic minorities. When using categorisations, the implicit assumption is that everyone within a certain category has similar experiences, but this is not always the case. Minority populations, such as Black or Asian, can be diverse and include people of 'mixed race' or less visible ethnic groups, such as migrant workers (Garland et al., 2006). Similarly, categorisations such as 'people with intellectual disabilities' mask the diversity within such groups and make it difficult to identify underrepresented conditions. Such groups may include people with profound intellectual or communication impairments who may be less able to participate in research, and it would be unwise to assume that their views are the same as those who can speak for themselves.

Failure to receive survey request

A key reason that someone may fail to receive the survey request is *poor quality data* about the address or other details of the intended recipient. Inaccuracies in the data used for the sampling frame, perhaps caused by errors in inputting data or poor updating of records, may occur more often amongst certain groups of people, for example, those who move home more frequently or have less contact with the record keeping organisations. In addition to this, surveys may sometimes be mistaken as junk mail and thrown away (Buckley et al., 2007; Harkins et al., 2010).

Surveys may also be *intercepted* preventing them from reaching their intended recipient. Gatekeepers, including GPs, care managers, support workers, carers and family, may intercept the survey and not pass it on through their desire to protect the intended recipient of the survey (Atkinson and Flint, 2001; Cambridge and McCarthy, 2001; Howard et al., 2009; McKeown et al., 2010; McNally, 2003; Oliver-Africano et al., 2010; Rugkasa and Canvin, 2011; Tuffrey-Wijne et al., 2008; Zermansky et al., 2007). The co-operation of gatekeepers will depend on the type of relationship they have with the potential participants, how they perceive the research and their judgement about who should be involved (McKeown et al., 2010). Previous research has shown that the potential for participants to not receive the survey questionnaire may be more acute in certain settings. Developmental work on the ASCS (Malley et al., 2010), for example, found that English care homes operated many different methods for distributing mail to residents and that staff often censored mail, considering some inappropriate for the residents.

Inability to respond to survey request

Questionnaires, delivered as a self-completion version or by interview, can pose particular challenges for some groups of people, such as adults with learning disabilities and people with dementia, due to their limited cognitive capabilities and the demands placed on them by the questionnaire. Questions about time, quantitative judgements, direct comparison questions, abstract concepts and generalised judgements can all be challenging. However, people with intellectual disabilities are a heterogeneous population, in terms of personal history, linguistics and cognitive ability, which, some argue, makes it unlikely that a single questionnaire will be valid for the whole population (Finlay and Lyons, 2002). The same issue would also be true of other groups, such as people with dementia.

For older people with cognitive impairments such as dementia, work by Malley et al. (2010) has suggested that highly structured questions with fixed answer options are difficult to answer. This finding is supported by work by McKee et al. (2002) looking at the quality of life of residents of a nursing home, using the Schedule for the Evaluation of Individual

Quality of Life (SEIQoL-DW).² The study found that the vast majority of residents could not complete the interview due to poor physical health (45%) or confusion (28%), or refused to participate (10%). Indeed, only 10 out of the 60 participants actually completed the interview and of these, half were judged by the interviewer to have had a poor understanding of the interview process. Mckee et al. noted that the SEIQoL-DW had been previously validated for use with healthy older people and those with mild cognitive impairment and concluded that its use was problematic when extended to either confused or frail older people.

Other types of disability may also affect response propensity to a postal survey due to the nature of the task and its interaction with the nature of the impairment. A postal survey of parents of children with ophthalmic disorders found that households with no other *visually impaired* family members were more likely to respond, suggesting a potential link between the presence of visual impairments and response to a postal survey (Rahi et al., 2004).

There is also evidence that *disability severity*, as reflected in proximity to death, affects response to postal surveys. Kauppi et al (2005), for example, found a relationship between response propensity and proximity to death, with nonrespondents being more likely to die two years after the study was conducted.

Refusal to participate

Some people refuse to participate because they are *not interested in the research topic*. It may not be a priority in their lives or they may not feel it will benefit other people or themselves (Gilbert, 2004; Harkins et al., 2010; Rugkasa and Canvin, 2011; Williams et al., 2007; Woodall et al., 2011). Another reason for refusal may be that the potential *participant does not define him/herself as part of the research population* as it is defined on the introductory letter or information sheets. For example, Williams et al (2007) describe how potential participants in their study, which focused on the physical health of older people living in areas of socio-economic deprivation, regarded themselves as physically active and therefore not part of the intended research population.

Trust is an important factor affecting participation in research. Often people do not trust the person or institution asking them to take part in the study, which can affect their propensity to respond (Andrews, 2005; Brown and Scullion, 2010). For example, some studies found that past negative experiences with the health or social services or broader institutions create a culture of mistrust and suspicion about the purpose of the study (Harkins et al., 2010; Tuffrey-Wijne et al., 2008). Participants may have concerns about privacy or

² The SEIQoL-DW is an interview based tool which takes between 20 and 30 minutes to administer. The respondent has to choose five domains that are important to their quality of life and rate their current level of satisfaction in each. Finally, the relative importance of each domain is determined via weighting.

confidentiality, worrying about results being reported back to carers or care managers (Proctor, 2001; Ulivi et al., 2009; Young and Chesson, 2006). Trust also relates to research fatigue; people are weary of researchers targeting particular communities, such as traveller communities, and then disappearing (Brown and Scullion, 2010).

A related issue that may influence nonresponse is that some people may think they are not able to answer questions, or provide the 'right' answer (Proctor, 2001). There may also be other circumstances that make them feel uncomfortable and unable to respond, for example Young and Chesson (2006) surmise that people who need support from staff may be less willing to respond.

Summary

The literature on underrepresentation and nonresponse to postal surveys discusses a number of reasons certain groups tend to be underrepresented, which can be categorised as being excluded from the sampling frame, failure to receive the survey, being unable to respond and refusing to respond. Of importance for the ASCS, given the nature of the survey population, certain disabilities appear to make a person more likely to be a nonrespondent. In particular, adults who have cognitive or memory impairments, such as dementia or learning disabilities, are highlighted as being particularly susceptible to being both excluded from the sample and, when included, finding that they are unable to respond. Moreover, often those with cognitive or memory difficulties may have a gatekeeper who decides not to pass the survey request on to them, making it more likely that people in these groups fail to receive the survey request. There are, therefore, multiple reasons why people with learning disabilities and cognitive impairments are likely to be underrepresented in postal surveys such as the ASCS.

An interesting omission from the literature reviewed was the recognition that it is not just learning or cognitive disabilities that can impair a person's ability to respond but that physical disabilities too can create real challenges for responding to postal surveys. In a study of adults who require help to complete surveys, Malley et al (2010) found, that the kinds of help people required to participate in surveys encompassed purely physical aspects of the survey process, such as writing on the survey form or posting the response back. Given the possibility that people with physical impairments may overcome their disabilities easily by asking for help to complete the questionnaire it is not clear to what extent physical disabilities are really a barrier to response. The ability of people with physical impairments to respond may well depend on the availability of people to help them but we were unable to find any research to support this viewpoint.

A clear problem for the ASCS is identifying the extent of underrepresentation, particularly of ethnic groups and certain types of disabilities. Most CASSRs report ethnicity according to six categories (white, mixed, Asian, Black, Chinese and other), which are clearly too broadly specified to pick up underrepresentation of the growing and relatively new immigrant

communities, such as people from Eastern Europe. Ethnicity is also a poor indicator for literacy in English, so it will be impossible within the ASCS to assess the extent to which people who are illiterate in English are underrepresented. Additionally CASSRs use very broad categories to record disability (physical and sensory impairment, learning disability, mental health, substance misuse, and vulnerable people) and few CASSR systems allow for cases to have disabilities in more than one of these very broad categories. It is therefore very difficult to assess the extent of underrepresentation amongst people with cognitive impairments and for people with certain types of learning disability. This review has highlighted that it is not only the characteristics of users but the behaviour of the organisation delivering the survey that can affect response rates. Indeed most theories of survey participation posit that different sets of influences act on the sample members to determine their likelihood of participation and that non-participation is not a constant attribute of a person. Rather, response propensity is the result of the interaction between the attributes of the survey, the attributes of the sample member and their environment (Dillman et al., 2009; Groves and Heeringa, 2006; Groves et al., 2000). Since the ASCS survey is run independently by each of the 153 CASSRs in England following centrally-set guidance, the survey attributes vary by CASSR. Issues like trust and the quality of records, both identified as affecting response propensity, will depend on the CASSR. In the first instance on the relationship the CASSR has with its service users and in the second the CASSR's internal management systems and record keeping. Whilst it is clearly important to focus on how the survey can be improved to make it more accessible to people with certain long-term conditions, CASSRs also have a role to play in improving their management processes to ensure they attain good response rates.

Methods for minimising nonresponse

In this section we look at approaches that can be used to ameliorate nonresponse to surveys. The literature is organised into methods for encouraging participation, which address nonresponse due to refusal to respond to the survey, and methods for enabling participation, which address either failure to receive a survey request or inability to respond to the survey request. It should be noted that much of the literature is based on reflections of researchers and philosophically-inspired ideas about best practice rather than research evidence. Therefore for each method identified in the literature we summarise the key debates, guidance and evidence, before presenting the evidence in more detail where this is available.

Methods for encouraging participation

Payment and incentives

There appears to be mixed evidence about whether incentives and payment to participate in a survey actually help nonresponse rates. Whilst Simmons and Wilmot (2010), in their

review of the literature on incentives in surveys, find that incentives do increase response rates significantly, other research and research specific to the ASCS has found that giving incentives makes no discernible difference to response rates (Williams and Betts, 2010). It should be noted though that the analysis of the ASCS did not differentiate types of incentives and utilised variation in a cross-sectional dataset rather than trial-based research.

There is some debate around the ethics and benefits of providing payment to participants which, although not specific to seldom-heard populations, is of relevance. The core concern is that the promise of financial reward in any project information sheet may exert a sense of coercion or obligation, bringing into question the 'voluntary' nature of participation (Rugkasa and Canvin, 2011). A further issue raised is that payments may be blunt instruments that do not specifically address underrepresentation from seldom-heard groups. Rather, payment will simply reward those participants wishing to be involved, but who needed that extra 'nudge'. On this point, however, the literature suggests that some of the key groups who do not respond to surveys (e.g. younger people, people from BME groups, those with lower incomes) do tend to be most likely to be encouraged to respond by an incentive.

Simmons and Wilmot (2010) provide an overview of evidence on incentives in social surveys. The key conclusion is, as noted above, that incentives, however small, do increase response rates in postal surveys as well as in telephone and face-to-face interviews. They also note that the majority of evidence and opinion is that unconditional prepaid incentives are more effective than conditional incentives, the latter of which are often burdensome for participants. In a meta-analysis conducted by these authors it was shown that unconditional incentives increase the response rate by 19% compared to 4.5% when conditional incentives were used. Their review also noted studies which found that non-monetary incentives, such as pens and stamps, could be equally effective in increasing response rates.

The paper also deals with some of the issues around the use of incentives, such as concerns around incentives causing both decreases in data quality, as people give substandard response to the survey, and increases in response bias, as incentives do little to help those who cannot reply. Both of these critiques of incentives were found to have little support in the evidence they reviewed. Targeting incentives (e.g. to those groups who traditionally do not respond to surveys) was shown to be perceived as 'unfair' by those who did fill in surveys but, as one study based in the United States showed, did not stop those who did not get payment from responding. Simmons and Wilmot did not look at the ethical issues around incentives.

Confidentiality, anonymity and trust

The earlier section on refusal to respond noted that trust, or more accurately the lack of it, was a key reason why some people may choose not to respond to a survey request.

Research suggests that surveys that try to create a personalised experience and build a sense of relationship between the organisation carrying out the research and the potential respondents have increased levels of participation (Edwards et al., 2009; Scott and Edwards, 2006). Likewise, surveys that have clear endorsement from a university also tend to get higher response rates (Edwards et al., 2009). Dillman et al. (2009) list a number of ways in which trust in the organisation carrying out the survey can be enhanced: In addition to sponsorship by a legitimate authority, such as a university, and ensuring confidentiality, they suggest stressing the importance of the task and providing a token of appreciation in advance (see section on incentives).

Literature on participation does not really consider the issue of building trust between local authorities and recipients of services. The ONS report on ASCS methodology (Williams and Betts, 2010) , however, does make clear recommendations about confidentiality and anonymity and thus does begin to deal with issues of trust, albeit trust in what happen to your data. The report suggests that, in addition to the already existing protocol that provides details about how long data will be kept, information given to respondents should also include details of how data will be destroyed. Moreover, any agreement to take part in further research should, according to the report, be done via a separate card rather than in a section of the survey itself so not to compromise anonymity.

Edwards et al (2009) conducted a systematic review of studies reporting the randomised controlled trials of methods designed to increase response to postal and electronic surveys. Their search, which looked in 14 databases, found 481 eligible studies. Their review found 14 trials, involving 21,628 participants, which evaluated the effect of University endorsement or sponsorship on response rates. Response rates were found to be increased by more than a quarter when the survey came from a university compared to other types of organisations such as government departments or commercial organisations.

Scott and Edwards (2006) reported on a meta-analysis of the impact of hand addressed / personally signed letters on the response rate of postal surveys based on 14 studies, mostly from the United States, of randomised controlled trials and involving just over 12,000 participations. Their analysis found that the numbers of questionnaires returned could be improved through the use of direct personalised mailing (i.e. addressing the individual by name, such as Mrs Smith) rather than by any generic appellation (e.g. Dear Participant). Adding hand-written signatures of the researchers will further increase responses. The size of the impact of this strategy on the proportion of questionnaires returned was predicted to be between four and ten per cent depending on the baseline response proportion when using neither intervention.

Following up nonrespondents

The literature on following up nonrespondents in postal surveys presents a very clear message that using such strategies helps to improve response rates (Dillman et al., 2009;

Fowler Jr., 2009; Wensing et al., 1999; Wensing and Schattenberg, 2005). Moreover, state survey organisations, such as ONS (Williams and Betts, 2010) and Public Works and Government Services Canada (2007), both suggest that following up nonrespondents is an approach for not only increasing overall response levels but also counteracting response bias as follow-up contacts may draw in responses from groups that are less likely to respond.

ONS (Williams and Betts, 2010) also suggest the follow-up contacts should draw upon the multiple contact approach where different stimuli are used at each stage of contact (e.g. the first reminder is a thank you card that doubles as a reminder and the second reminder is a letter with replacement questionnaire) as it is more effective. Both the ONS and the Canadian government office also encourage the collection of data around who responds and when, in order to inform decisions about who to target with response reminders.

The use of a different method for follow-up contacts is endorsed by Dillman et al.'s (2009) survey creation guide which draws on a systematic review by Auster and Janda (2009) who found that those who received a final contact by telephone (following an invitational letter) were 1.5 times more likely to take part in a study compared to those who did not.

Edwards et al (2009) conducted a systematic review of methods to increase response in postal and electronic surveys. Their search, which looked in 14 databases for studies which were unconfounded, randomised controlled trials of methods designed to increase response rates, found 481 eligible studies. The review found 19 trials which, involving 32,778 participants, evaluated the impact of following up nonrespondents with either a repeat mailing or a telephone call and found that such strategies increased the odds of response by more than 25%. 11 trials in their review, involving 8,619 participants, evaluated the impact of sending another questionnaire to respondents and found that the odds of response were increased by a half. 5 other trials, involving 2,254 participants, included in this review, however, found that a follow up carried out via the telephone had no effect on response. This goes against guidance for survey creation, which suggests using a follow up to boost response, and is supported by an analysis of data from 3 more studies, involving 13,922 participants. Their review also found no difference in response rates between studies which carried out a reminder within 31 days of the initial contact and those who waited until at least 31 days had passed.

In their study of combining postal survey data with face-to-face interview data, **Allison et al. (2003)** found that following up nonrespondents with a visit and offering an interview raised the response rate from 30% to 75%. They did, however, note that this type of approach would require full ethical approval before the start of the study.

Summary

This section reviews methods that may be employed to encourage people to respond to survey requests. Clear messages emerge around the effectiveness of following up nonrespondents with additional mailings to improve response rates, although it is not clear whether this improves response rates from hard-to-reach groups. There is more debate around the effectiveness of changing the mode of follow-ups (e.g. by post or by telephone) and much of the outcome may depend on the combination of modes chosen and the survey population. However, this is usually seen as the best way of ensuring follow-ups engage hard-to-reach groups. We address this issue further in the next section where we consider alternative modes of data collection as a way of engaging groups who are less likely to respond.

Lack of trust has been shown to be an important reason for survey nonresponse but evidence around ways of building trust is more limited. Whilst there is good evidence that university sponsorship and personalising material can improve response rates, these methods are not particularly helpful for the ASCS. Other guidance around anonymity and confidentiality, and policies for data storage and destruction are considered best practice but there is no evidence about their effectiveness in improving response rates or engaging hard-to-reach groups.

There is good evidence that incentives can help response rates, including those amongst certain seldom-heard groups, such as younger people and some BME groups. However, the systematic review found that unconditional prepaid incentives are far more effective than conditional incentives. The use of all types of incentives raise ethical questions about coercion and creating a false sense of obligation but this would seem to be particularly the case for unconditional incentives. Indeed their effectiveness may well be due to exactly this: that people feel a sense of obligation to respond having been given a pen or money by the sender of the questionnaire. Some CASSRs do include incentives, but these are conditional, usually in the form of prize draws, and analysis of variations in response rates across CASSRs suggests that they are not effective ways of reducing nonresponse.

Methods for Enabling Participation

Consent and recruitment

Many of the studies reviewed by Beadle-Brown et al. (2012) discuss the process of consent and made suggestions about appropriate or helpful approaches that could ensure the inclusion of seldom-heard groups. We do not summarise this literature here as it is more applicable to interview-based surveys than postal surveys. However, for the few people who are given interview version of the ASCS or should the ASCS move beyond a primarily postal approach, the Beadle-Brown et al. (2012) review provides a useful summary of the issues and guidance.

There is some debate in the literature about whether recruitment/survey materials should avoid the use of diagnostic terms, such as dementia or schizophrenia, to encourage the recruitment of people who, along with their relatives, may be unhappy with particular labels or diagnoses (McKeown et al., 2010). For example, Hellstrom et al. (2007) used the term 'memory problem' to avoid using 'dementia' with prospective participants and their family, unless they brought up the term dementia themselves. Woodall et al (2011) suggest that, while researchers should use language carefully, they should also seek to engage with potential participants' understandings of their illness or condition. Just as there are problems with overly-technical language, euphemisms can also present problems since they may lead people to think they are not the target of the research.

Adapting the questionnaire

Although structured questionnaires have been used successfully with populations who, it was thought, may struggle to respond to such formats (Mindham and Espie, 2003a) questionnaires often need to be adapted or, as the next section will explore, be rejected in favour of a different mode of survey administration in order to enable these groups to more fully participate in postal surveys. This section looks at some of the ways questionnaires can be adapted to increase response rates for people who have cognitive, sensory, language or memory impairments or groups for whom English is not their first language.

Translations

It is not surprising that the literature on research participation points out that the involvement of people from ethnic minority communities in both quantitative and qualitative research is facilitated by providing the research materials in different languages. For example, in some studies, project information sheets were provided in audio format and a range of languages, whilst face-to-face interviews were offered and conducted in a range of languages (Allison et al., 2003; Lloyd et al., 2008; Rooney et al., 2011; Rugkasa and Canvin, 2011). For translated information to be effective, however, Williams and Betts (2010) also point out that the option to use survey materials in a different language needs to be made clear to any potential participants at an early stage. Another consideration is that any translation of supporting information or data tools will require considerable pre-testing, on-going development and back-translation before it is meaningful for different communities (Allison et al., 2003).

Easy Read

Most of the work on developing Easy Read versions of questionnaires is aimed at adults with learning disabilities. Some of the key messages are outlined below but, as guidance on survey creation suggests (Dillman et al., 2009), all surveys, regardless of the target population, should aim to be easy to understand and complete.

Literature on research participation suggests that questions need to be simplified (e.g. Likert-type scales reduced) and illustrated (e.g., smiley faces for response options and pictures for

questions) in Easy Read or learning disability versions of questionnaires. The use of different coloured stickers will help participants identify priorities (Finlay, 2001; Gordon et al., 2007; Nind, 2009; Schmidt et al., 2010; Young and Chesson, 2006). Show-cards, it is argued, are necessary to illustrate topics covered through open questions (Mindham and Espie, 2003b). Throughout the development of any standardised tool the validity of responses should be tested through the repetition of difficult questions and acquiescence through the use of reverse wording and nonsense questions, with any conflicting information being noted for later discussion (Finlay, 2001).

Malley et al. (2010) reports on a study which developed an Easy Read version of the ASCS for people with learning disabilities.³ Drawing on interviews and focus groups with adults with learning disabilities, the authors devised an Easy Read version of the ASCS and also noted key points for making sure the adapted questionnaire would be as widely understood as possible. They concluded that not only should the questionnaire use simple language, but that questions should be accompanied by illustrations that both illuminate the questions topic but also the answer options. Participants in the study found smiling/unhappy faces aided understanding of answer options. Moreover, the study found that too many answer options often caused problems for people with learning disabilities and that seven answer options, as found in some of the questions in the ASCS,- and, in some cases, even five - was too many for some people in the study. The study also noted that an Easy Read version in itself would not be enough to enable everyone to participate and that other formats and help answering questions would often be required. Although the study developed an Easy Read version, neither its reliability nor validity were evaluated due to time constraints.

Alternative modes of data collection

Literature on survey methods suggests that one of the best ways to reduce nonresponse is to use more than one mode to collect data (Dillman et al., 2009; Fowler Jr., 2009; Williams and Betts, 2010). Studies by Scott et al. (2006) and Manfreda (2008) provide evidence that mixed mode surveys have better response rates than those that just using a single method⁴. However, it is clear that different modes of survey administration may impact upon the data collected or, in other words, that there may be a mode effect on responses (Williams and Betts, 2010). Jackle et al. (2006) report on a study looking at mode effects in telephone and face-to-face interviews and noted that, whilst there was no evidence of satisficing⁵, there

³ This was part of larger study that looked at various aspects of administering the ASCS to groups that may struggle to participate in the standard postal survey.

⁴ Most the recent studies looking at response rates across different survey mode and mixed modes concentrate upon web or email based survey formats.

⁵ Where a respondent does not perform optimally during all stages of the cognitive process.

were clear indications that respondents were more likely to give socially desirable response in the telephone interview format.

Despite the issues around mode effects, the literature is clear that using multiple modes enables greater participation from groups that might otherwise struggle with a single format (Dillman et al., 2009; Fowler Jr., 2009). Literature on research participation also supports multi-mode data collection as an approach to increase response rates from groups that may have difficulty responding to postal surveys (Rugkasa and Canvin, 2011; Tuffrey-Wijne, 2007; Wilson et al., 2010). Indeed, the ONS review (Williams and Betts, 2010) of modes of survey data collection suggested that offering a range of ways in which data could be collected (e.g. postal surveys, face to face interviews and telephone interviews) was an effective method of increasing participation, particularly with groups who have cognitive difficulties. As such, study designs should allow and plan for the use of a variety of methods with different groups of people within the same study (Gilbert, 2004).

Face-to-face interviews

The literature suggests that using a face-to-face interview to collect data enables more people to participate in a study than using a postal questionnaire. For example, Seymour et al. (2001) found that the only way to fully enable older people with physical disabilities to take part was to change the mode of administration of a particular tool (in this case, the SF-36) from postal self-completion to a face-to-face interview.

However, simply changing the mode of administration is often not enough to really enable the participation of particular groups, such as people with dementia or adults with learning disabilities., and questionnaires, questions, response options and supporting information may need to be adapted in a variety of ways. These issues are all summarised in Beadle-Brown et al. (2012).

Telephone interviews

Telephone interviews have been found to be effective in increasing response rates in surveys. For example, in one study they were used to contact and interview nonrespondents using a postal survey on lower back pain (Lall et al., 2012). In this study, using telephone interviews raised the response rate of the survey from 71% to 85%. Despite this, there is not a great deal of evidence regarding how effective telephone interviews are at increasing response rates amongst groups that are less likely to participate. Hoffman et al. (2010) argue that while there is some evidence that people who have had a stroke and their carers can participate in a survey via telephone interviews, the value of such survey modes is questionable for some groups with more severe impairments. Visual methods that facilitate the participation of adults with learning disabilities, for example, cannot be incorporated into telephone interviews.

Web-based data collection

A key development in survey data collection in recent years has been the increase in the use of web-based surveys, either as a sole method of data collection or a part of a mixed mode survey. However, as the response rates are generally lower for online surveys, it is important to consider the characteristics of potential respondents since many, particularly older people, still lack access to the internet or are infrequent users and would struggle to use this particular mode to respond. The ONS review (Williams and Betts, 2010) also notes that having an online option could, actually lower response rates as having to decide which mode to use to complete a survey acts as a break in the response process.

Support

Another method of enabling participation in a postal survey is to allow participants to have support from another person to complete the questionnaire. This strategy has not been widely explored in the literature on participation but Fraser and Fraser (2001) and Pawson et al. (2005) suggest that ensuring effective participation of people with learning disabilities requires an 'interpreter' (i.e. a professional or family member) to ensure that the views of individuals are appropriately conveyed. However, other commentators have expressed concerns with this approach arguing that supporters can act as 'gatekeepers', negatively influencing people's involvement and silencing views that may be contrary to existing policies (Brewster, 2004; Kaehne and O'Connell, 2010; Llewellyn, 2009; Nind, 2009).

The role of support in helping people complete surveys was explored in developmental work for the ASCS (Malley et al., 2010). In this work, support from family, care home workers and advocates was considered. What was clear was that there was significant variety in the types of help that were required by potential respondents to the survey, from assistance in filling in a questionnaire for those with physical disabilities through to help explaining questions and even guiding their answers for respondents with cognitive and intellectual disabilities. In keeping with the limited literature in this area, the authors of the study did note the potential for helpers to take over and answer for respondents, even in situations where the participants needed only a little help to answer for themselves. This was true for both family helpers and staff in care homes. However, it was clear that without such help many people would not be able to take part in the survey. Therefore, the recommendations from the study were that in cases where people could not respond to the survey without help then the use of helpers should be encouraged and be recorded in the survey response. Changes to the survey guidance were also suggested to clarify how both family and care home staff should give appropriate help and not answer for the respondent.

Malley, Caiels et al (2010) also explored the use of advocates as a method of enabling participation but were unable to identify other studies that had looked at this issue. They concluded that, whilst advocates were not currently a viable strategy for enabling people with cognitive impairments to complete the ASCS, they may be appropriate in other

circumstances. There were a number of reasons for this conclusion. Firstly, it was clear from their field work with both advocacy agencies and care home managers that there was not the required level of advocates or resources to assist the large number of people who needed help. Secondly, there are safety concerns about advocates going into the homes of service users, and both practical and resource issues about service users visiting the agency offices of advocate. Finally, it was felt that in order to help people fill in a questionnaire, the advocates would need to build a relationship with the service user and that this would take several visits. As a result of these considerations, the authors felt that only where there was an existing relationship between a survey participant and advocate would using an advocate to help complete a postal survey be a sensible course of action.

The study also noted that one possible mode for supporting some, albeit probably less severely impaired, potential respondents to surveys, was a telephone help line. In previous surveys examined by the authors some local authorities had set up a centralised helpline that were staffed with trained call centre operatives and had successfully minimised the need for costly face-to-face interviews.

Summary

This section reviewed methods that can be used to enable participation in postal surveys and identified four key techniques to raise response rates: Firstly, ensuring that the information section of the postal survey is relevant to the respondent, as irrelevant information can serve as a barrier to participation; secondly, adapting the questionnaire to reflect the 'individualised' needs of the participant, such as versions in other languages and Easy Read versions for adults with learning disabilities; thirdly by employing alternative methods of data collection, such as face to face or telephone interview, to be used alongside the postal survey; and fourthly, by encouraging or even providing support to complete the survey. A scan of the literature reveals that evidence to support the effectiveness of the various approaches is uneven. The most solid evidence is around the provision of alternative methods. Both face-to-face and telephone interviews have higher response rates than postal surveys and, in the case of face-to-face interviews, there is evidence that they are more effective for engaging people with physical impairments. They are also likely to be easier for people with learning disabilities, who are likely to need help navigating a questionnaire, and for other groups, such as people with visual impairments. In this respect the ASCS may benefit from employing a mixed-mode approach which has been shown to be effective elsewhere and was recommended by the ONS in their review of the ASCS (Williams and Betts, 2010). However, one mode that is unlikely to be effective is web-based approaches which tend to have lower response rates than other modes and is dependent on access to and regular usage of the internet. The characteristics of respondents to the ASCS mean that a substantial proportion of potential respondents either do not have access to the internet or are infrequent users and would thus struggle using this particular mode to respond (Williams et al., 2007).

There is also evidence of the value of adapting questionnaires to make them more understandable to certain groups. Most adaptations focused on providing versions in other languages for respondents who do not speak English well or in an Easy Read format for adults with learning disabilities. However, experience suggests that adapted versions may not be enough to enable participation. First, and particularly in the case of the Easy Read version, recipients are likely to still need additional support to navigate and respond to the questionnaire. Second, although we were unable to find any studies specifically testing this issue, guidance suggests adapted versions should be sent to the recipient in place of the standard version (rather than be made available upon request) to be fully effective.

A further issue is the comparability of adapted versions with the standard version. There is a substantial literature on this for translating questionnaires into other languages that recommends several stages, including forward- and back-translation to ensure the sense of the questions and response options are maintained. However, in the case of Easy Read version, Beadle-Brown et al. (2012) were unable to find any discussion of these issues in literature from the United Kingdom. Easy Read versions also tend to have fewer response options which create additional problems at the analysis stage. Nevertheless, adapted versions of questionnaires are a good way of ensuring the engagement and collection of standardised and valid responses from people whose first language is not English or who have learning disabilities. None of the reviews we looked at identified or discussed using adapted versions for other groups who may benefit, such as people with dementia.

The review also highlights the lack of research into the effectiveness and effects of having someone to help answer the questionnaire. Research has demonstrated the need for people with learning disabilities to have support to complete questionnaires and the similar needs of people with dementia or other forms of cognitive impairment. Help to complete postal questionnaires can come in many forms (e.g. translation, interpretation, help reading, and help filling in the questionnaire) and may affect the quality of the data returned. For example, ad hoc translations will vary according to the translator and may alter the sense of the questions. In addition, helpers may find it hard not to influence the recipients responses again affecting the validity of the data returned. However, there is little research into the effect of having help and the impact of the different types of help on responses. Support is also clearly necessary for many groups of disabled people and is a relatively cheap way of ensuring their involvement. However, there is also the potential for continued exclusion if those supporting people to respond do not give appropriate levels of support (for example, answer on behalf of, instead of aiding the participant).

Alternative approaches when participation is not possible

Some people are unable to express their views even when the methods discussed above are used. Such people are likely to suffer from cognitive, intellectual and/or communication

impairments that make it difficult for them to form and communicate their opinions about specific areas of their quality of life in either written or oral formats. In the field of social care, the primary groups we would consider to fit into this category would be people with learning disabilities and severe dementia (Hoe et al., 2007), although there are also other groups of people, such as stroke survivors, whose condition may limit their cognitive capabilities.

In this section we consider how one might measure the quality of life or social care outcomes of people who are unable to express their views directly and focus on two broad methods that have been designed specifically for the purpose of eliciting such data in these cases: the use of proxy respondents and observational methods. Specifically we consider how the developers of the tools have addressed the challenge of measuring quality of life, given it is, arguably, a fundamentally subjective concept, and how they have assessed the validity and reliability of the tools.

Proxy instruments for gathering quality of life or outcomes data

A proxy respondent is someone who answers the survey questions on behalf of the intended survey recipient. Typically family carers, clinicians or staff providing social care, such as care workers or managers, are used as proxies. Our interest here is in understanding the processes researchers have gone through to develop proxy instruments or proxy versions of standardised instruments. In choosing the review questions we were guided by issues raised in the literature around the conceptualisation of quality of life (Rapkin and Schwartz, 2004; Schalock, 2004; Schwartz and Rapkin, 2004; Sprangers and Schwartz, 1999), philosophical issues around who has the right to speak for someone else, as well as more practical issues surrounding who is able to speak for the service user and the development of standardised instruments (DeVellis, 2003). The literature on this latter subject raised three questions that guided our review of the proxy tools:

1. How did the researchers decide upon the question format and what question format was chosen? (See **Box 2: Types of proxy questions** categorisation of question types.)
2. Did the researchers impose any restrictions around the choice of proxy or when a proxy should be used?
3. How have the researchers assessed the validity, reliability and sensitivity of the instrument, particularly in instances where the instrument is a proxy version of the standardised instrument?

Box 2: Types of proxy questions

Schwartz and Rapkin (2004) distinguish three types of quality of life questions: 'evaluation-based', 'perception-based' and 'performance-based' measures.

- Performance-based measures reflect the quantity and quality of effort, for example whether someone can walk to the shops.
- Perception-based measures gauge individual judgment concerning the occurrence of an observable phenomenon, for example how often someone walks to the shops.
- Evaluation-based measures are ratings of an experience, for example how difficult it was for the person to walk to the shops.

Schwartz and Rapkin (2004) note that whilst we may expect some disagreement between observers for all types of measures, particularly perception-based measures, which being based on observations are prone to bias, there is likely to be most disagreement for evaluation-based measures since judgements are made by people using their own (idiosyncratic) criteria. This characteristic makes evaluation-based measures, within which category the majority of quality of life measures fall, fundamentally subjective and therefore difficult (if not impossible) for proxies to answer on behalf of the intended survey recipient.

The methods used to identify instruments for this review are outlined in Appendix 1. As we describe in the appendix, we found it very difficult to identify tools using standard systematic review methods. We therefore drew on papers on proxy tools which were identified in the review carried out by Beadle-Brown et al (2012) and scanned them for references on both the testing and development of the tools as well as references about other proxy tools. We cannot therefore be sure that we have identified all the proxy tools available to measure quality of life. Nevertheless, we feel that our review gives a flavour of the approach taken by researchers to developing proxy tools and the usage of such tools.

Our review identified 11 quality of life instruments that had either been developed as proxy instruments or had been used with proxies. The instruments developed for eliciting outcomes or quality of life information from proxies tended to be condition-specific, for example focusing on strokes or Alzheimer's disease. The bulk of the instruments found were specific to older people with dementia but there were also studies where proxies were used to elicit outcomes for people with intellectual disabilities and adults who had suffered a stroke⁶. The instruments are summarised in Table 1.

⁶ There is also a large literature on parents as proxies for children. This literature is outside of the remit of this review.

Table 1 Key characteristics of the proxy tools reviewed

Instrument name	Client group	Main usage	Other uses	Developed as a proxy tool / version?
Alzheimer’s Disease-Related Quality of Life (ADRQL)	Dementia/Alzheimer’s disease	Research	None	Yes
Camberwell Assessment of Need for the Elderly (CANE)	Older people with any mental health problem, including cognitive performance	Needs assessment tool for clinical and research settings	Research	Partly as developed for assessment
Euroqol (EQ-5D)	Not specific Used with people with dementia and stroke survivors	Research	None	No, only used with proxies
DEMQOL-Proxy	Dementia/Alzheimer’s disease	Research	None	Yes
Quality of Life in Alzheimer’s disease (QoL-AD)	Dementia/Alzheimer’s disease	Research	None	Yes
Quality of Life Assessment Scale (QOLAS)	Dementia/Alzheimer’s disease	Research	None	Yes
Evaluation of Quality of Life Instrument (EQLI)	Intellectual disabilities	Research	None	Yes
Quality of Life Interview Schedule (QUOLIS)	Intellectual disabilities	Research	None	Yes
The Comprehensive Quality of Life Scale – Intellectual Disabilities (COM QOL-ID)	Intellectual disabilities	Research	None	Yes
Instrument name	Client group	Main usage	Other uses	Developed

				as a proxy tool / version?
Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39)	Stroke survivors, with post-stroke aphasia	Research	None	Yes
Stroke-Specific Quality of Life (SS-QOL)	Stroke survivors	Research	None	Yes

For each of the instruments we describe the features of the tool and its development, focusing on addressing the three issues described above: the format of the questions, the choice of and restrictions surrounding who is a proxy, and the validity and reliability of the measure. For some of the measures these details were sparse and we were not able to find information regarding the tool's development or obtain access to the tool, despite sending requests to the corresponding author. The availability of evidence for each of the tools is summarised in **Table 2**.

Table 2: Details of information found regarding each of the proxy tools

Instrument name	Copy of instrument obtained	Details of development and testing available
Alzheimer's Disease-Related Quality of Life (ADRQL)	Yes	Yes, quite detailed
Camberwell Assessment of Need for the Elderly (CANE)	Yes	Yes, quite detailed
Euroqol (EQ-5D)	Yes	n/a
DEMQOL-Proxy	Yes	Yes, very detailed
Quality of Life in Alzheimer's disease (QoL-AD)	Yes	Yes, quite detailed
Instrument name	Copy of instrument obtained	Details of development and testing available
Quality of Life Assessment Scale (QOLAS)	No	Yes, limited details
Instrument name	Copy of instrument	Details of development and

	obtained	testing available
Evaluation of Quality of Life Instrument (EQLI)	Yes	Yes, details on testing, early development published in Italian
Quality of Life Interview Schedule (QUOLIS)	No	Yes, details of testing
The Comprehensive Quality of Life Scale – Intellectual Disabilities (COM QOL-ID)	Yes	Yes, details of testing
Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39)	Yes	Yes, details of testing
Stroke-Specific Quality of Life (SS-QOL)	No	Yes, details of testing

Alzheimer’s Disease-Related Quality of Life (ADRQL)

The ADRQL, developed in the late 1990s, is a tool for assessing the health-related quality of life (HRQOL) of people with Alzheimer’s disease and states that the proxy must be either an informal caregiver (e.g. family member or friend) or a formal caregiver (nurse, therapist, paid carer etc.). To assess the individual’s HRQOL, caregivers engage in a standardised structured interview and are asked to identify observed behaviours that may reflect the person with Alzheimer’s HRQOL. The current version of the ADRQL has forty items covering five domains: social interaction, awareness of self, feelings and mood, enjoyment of activities and response to surroundings (Black, 2012) and, for each of these items, the caregiver must report whether or not the statement describes the person they care for over the two weeks prior to the interview. Thus, and as the sample questions below illustrate, the questions are perception-based.

Box 3: Sample questions from ADRQL

These statements are about relating to and being around other people. After each statement, please answer “Agree” if the statement describes Mr/Mrs/Ms..... in the last two weeks or answer “Disagree” if it does not.

- A1. He/She smiles or laughs when around other people
- A5. He/She talks with people
- A9. He/She smiles or laughs or is cheerful

These responses are used to create a score which gives an overall percentage from 0 to 100. Scores can also be calculated for the five domains. Scores for individual items are weighted to reflect how important carers feel they are to the quality of life of people with dementia.

Rabins et al. (1999) outline the development and testing of the ADRQL. The first stage was the creation of an item pool which drew upon both knowledge of other tools and the objectives of this tool. An expert panel of local health care professionals was then convened to create their own item pool, review the initial item pool and from these create domains. The panel's recommendations were nine domains, each with four or more items. Following further checking with nationally recognised experts the tool was then the subject of a focus group with family caregivers who were asked to identify both missing and inadequate items, and consider what behaviours they would use to evaluate the quality of life of somebody with Alzheimer's. The results of this focus group, alongside another consultation with local health care experts, fed into further revisions and the sorting the items into domains (see above for list of ADRQL domains). This final draft of the instrument was cognitively tested with three caregivers in order to assess whether they could understand the instrument and, as a result of these interviews, no changes were made. Preference weights were created during the final stage of the development and drew on work with 62 caregivers who were asked to rank the items within each domain from one to ten to reflect importance of the item to the health-related quality of life of people with Alzheimer's.

Kasper et al. (2009) outlines the testing of validity and reliability of the ADRQL in its current and original incarnation⁷. Using a sample of 310 people with Alzheimer's and their caregivers drawn from three settings, the community, nursing homes and assisted living facilities, the psychometric properties, validity and responsiveness of the instrument and also the validity of proxy responses were examined. The study found that the tool exhibited a good item internal consistency, a high correlation to scales hypothesized to be correlated with the ADRQL⁸, a good range of scores (albeit skewed towards the top end), and low levels of missing data. The construct validity and responsiveness of the ADRQL was also supported by its ability to discriminate among individuals based on cognitive and physical functioning, and behaviour. The study did not directly address proxy-subject agreement on quality of life; rather it assessed validity of caregiver responses by looking at how they were associated with caregiver characteristics. The study suggested that caregiver characteristics were "largely unrelated to scores" (275) and concluded that the revised version of the

⁷ The original ADRQL was revised slightly; dropping seven items (reducing it from 47 to 40 items) and re-arranging the domains items were attached to. Table 1 in Kasper et al. (2009) outlines the revisions.

⁸ The study hypothesised that that individuals with lower Mini-Mental State Examination Scores (MMSE) would also have low ADRQOL scores,

ADRQL exhibited improved measurement properties over the original and therefore should be used in preference over the original version.

Camberwell Assessment of Need for the Elderly (CANE)

Based upon the same model as the Camberwell Assessment of Need (CAN) (Phelan et al., 1995), the Camberwell Assessment of Need for the Elderly is a needs-assessment tool that is designed to be used in both clinical and research settings. Since it is a needs-assessment tool outcomes are operationalised around the concept of need and individuals are assessed on each item within the tools as experiencing no need, met need, unmet need or unknown, and followed up by questions about the level of help received by the service users in this area of their lives. The CANE consists of twenty-four domains covering, what may be termed as, generic quality of life domains, such as food, accommodation and activities, and also domains that are more specific to groups for which this tool was designed, such as psychological distress, memory and self-harm. As the sample questions below illustrate, the questions are primarily evaluation-based.

Box 4: Sample questions from CANE

Accommodation	
Does the person have an appropriate place to live?	
No needs	Has an adequate and appropriate home (even if in hospital). No need for assistance
Met need	Home undergoing adaptation/redecoration. Needs and is getting help with accommodation, e.g. in residential care, sheltered housing
Unmet needs	Homeless, inappropriately housed or home lacks basic facilities such as water, electricity, heating or essential alterations.
Not known	
Daytime activities	
Does the person have difficulty with regular, appropriate daytime activities?	
No needs	Adequate social, work, leisure and learning activities, can arrange own activities
Met need	Some limitation in occupying self, has appropriate activities organised by others
Unmet needs	No adequate social, work or leisure activities
Not known	

Unlike many other of the tools discussed in this section, the CANE is not designed solely for older people with either dementia or Alzheimer's disease. Rather the tool aims to identify unmet needs in older adults with any mental health condition, including depression and schizophrenia. Ratings for each of the domains can, in research settings, be gathered from a range of people including the service user, a member of staff or an informal carer. In clinical settings, the standard approach is staff assessment.

Two articles outline the development, psychometric properties and the testing of the agreement between needs ratings made by service users and proxies (carers and members of staff) of the CANE.

Reynolds et al. (2000) presents the process behind the development of the CANE which was based upon and extended by the creators of a CAN tool called the Camberwell Assessment of Need for Older Adults (CANOA). Specific aspects of the tool, including the layout and wording, were reviewed and further refined at a consensus conference involving 38 delegates, representing professional and voluntary organisations involved in the care of older people. The results were drawn together to create a CANE draft which was used in ten pilot interviews with older patients, their key member of staff and their carer. These interviews resulted in minor wording and item ordering revisions aimed at making the tool more user-friendly. This final draft was used in a range of tests with service users, staff and carers carried out in a number of locations, including several sites in England and Wales as well as tests in Sweden and the United States. Results of tests suggested that the tools had good face (as assessed via expert consensus and good Flesch reading scores) and content validity (all items were rated as at least moderately important by patients, carers and professionals) and showed reasonable construct validity (as assessed via an analysis of the convergence of ratings in the memory item and other items such as self-care, accidental self-harm and household skills). There was also found to be some correlation between CANE and other tools looking at health status: whilst a strong correlation was found for the Clifton Assessment Procedure for the Elderly-Behaviour Rating Scale⁹ (CAPE BRS) and the

⁹ The Clifton Assessment Procedure for the Elderly-Behaviour Rating Scale (CAPE BRS) is a questionnaire administered by nursing staff for use in clinical settings and was developed in the mid-1970s. The CAPE evaluates the presence and severity of impairment in mental and behavioral functioning and was intended for elderly long-term psychiatric patients. The questionnaire consists of two components: the Cognitive Assessment Scale (CAS) and the Behavior Rating Scale (BRS). The CAS includes a 12-item information and orientation subtest, a brief mental abilities test and a psychomotor performance test that involves tracing a line through a maze. The BRS contains 18 items and is completed by relatives or staff familiar with the patient's behavior. It covers physical disability including ADLs, apathy, communication difficulties and social disturbance.

Barthel index¹⁰, there was only a weak negative correlation with SF-36. Using Kappa scores, the study found excellent inter-rater reliability between patient, staff and carer scores, although the test-retest scores were only adequate.

Hancock et al. (2003) looked specifically at the agreement between the ratings of need made by service users and those made by staff and carers. The study also took place in a range of locations in England and Wales as well as some international settings (again Sweden and the United States), and involved a hundred and one older service users (and matched staff and carers) with a range of mental health issues (34% dementia, 10% schizophrenia, 43% depression/bi-polar disorder, 4% anxiety disorder, 8% other DSMIV disorders). Fourteen were not able to participate in an interview due to the severity of their condition. Agreement between the raters was compared in two ways. Firstly, paired sample t-tests were used to compare the frequency of met and unmet needs reported by service users, staff and carers, and showed that service users rated themselves as possessing significantly fewer met and unmet needs. However, when those with dementia were removed from the comparison there was no difference between service user, staff and carers suggesting that users with dementia were less likely to report that they had a need, be it met or unmet. The study also examined Kappa ratings of agreement and found a fair level of agreement between staff and service users (0.52) and between carers and service users (0.53). In both cases the proxies rated the service user's level of needs as higher than services users rated themselves. The study also noted that proxies and service users disagreed on certain aspects of the service user's life. For carers the largest disagreements with the ratings of service users were with regard to company, drugs and daytime activities. These areas, and the area of money, also had the largest disagreements between staff and service users. The study concluded that the views of staff, carers and service users were not inter-changeable.

EQ-5D

The EQ-5D is a well-established standardised instrument for measuring health outcomes and is designed as a tool for self-completion, primarily in postal interviews although it can be used in face to face interviews. It contains a descriptive section which asks a single question for each of five domains (mobility, self-care, usual activities, pain/discomfort, anxiety / depression) and a visual analogue scale for rating one's health between 0-100. Each question in the descriptive section has the following options: no problems, some problems, extreme problems, as the sample questions below show.

¹⁰ The Barthel Index is a ten item tool used to measure performance in activities of daily life.

Box 5: Sample questions from EQ-5D

Please indicate which statement best describes your own health today

Pain or discomfort

- I have no pain or discomfort
- I have moderate pain or discomfort
- I have extreme pain or discomfort

Anxiety or depression

- I am not anxious or depressed
- I am moderately anxious or depressed
- I am extremely anxious or depressed

Whilst there is no official proxy version of EQ-5D and therefore no official development work or assessment of the validity and reliability to outline, three studies look at the inter-rater reliability of the EQ-5D between people with dementia (2 studies) or stroke survivors (1 study) and their proxy respondents (staff and informal carers in the dementia studies and family caregivers in the stroke study). The first study with dementia sufferers used a slightly amended version of the EQ-5D tool to allow completion by proxies. However, the changes were minimal leaving the questions in the evaluative format. For example, the proxy version read “his/her own health” rather than “your own health” and “he/she has...” rather than “I have...” (Coucill et al., 2001). Therefore, the questions are best categorised as being evaluation-based. Unfortunately, the study of stroke survivors and their proxies does not report if any changes were made to either the wording or the administration of the EQ05D when used with proxies, although given the restrictions imposed by the owners of the tool around making changes to the instrument, it seems likely that any changes were minimal

Bryan et al. (2005) and Coucill et al. (2001) looked at the agreement between reports given by 64 people with dementia (identified by hospital records in a single NHS Trust in England and given a DSM-IV diagnosis of dementia of Alzheimer’s type and/or vascular dementia) and their proxies to the EQ-5D. Data was collected over two visits. On the first visit, a clinician interviewed the person with dementia to confirm a DSM-IV diagnosis and their carer completed a modified proxy version of EQ-5D to rate the person with dementia’s ability to carry out ADLs using the Bristol Activities Daily Living Scale (BADLS). A subset of carers (n=41) also completed the Neuropsychiatric Inventory¹¹. On the second visit, a researcher collected background information from the carer and administered the EQ-5D to

¹¹ The Neuropsychiatric Inventory assesses 12 behavioural disturbances: delusions, hallucinations, dysphoria, anxiety, agitation/ aggression, euphoria, disinhibition, irritability/ lability, apathy and aberrant motor behaviour, sleep and appetite and eating disorders). This tool is interview based.

the person with dementia. At this time the carer also completed the proxy version of the EQ-5D.

The studies found that informal carers reported worse outcomes across all of the domains contained within EQ-5D compared to the person with dementia, whilst the clinicians tended to report significantly better outcomes. As Bryan et al. (2005) put it “clinicians argue that the patients are more mobile, more able to wash and dress, more able to carry out their usual activities and less likely to suffer from anxiety or depression” (115). This pattern was particularly true for the domains of pain/discomfort and anxiety/depression. Indeed neither proxy (informal carer or clinician) had kappa (agreement) scores suggesting that agreement with the person with dementia was anything better than fair. However, the study suggested that people with dementia tended to over-estimate their quality of life. Responses from people with dementia accounted for over 90% of the ‘ceiling’ or ‘full health’ responses in the study. Moreover, it was data from clinicians that had the greatest correlation with the BADLS (carer-rated) data collected. The studies concluded that who rated health outcomes mattered.

Pickard et al (2004) report on the level of agreement on the EQ-5D between 124 stroke survivors and family caregivers who were acting as proxy respondents, both of whom completed the EQ-5D at baseline (2-3 weeks after a stroke) and were followed up six months later. Cross-sectional agreement between stroke survivors and their proxies was found to be generally acceptable ($ICC > 0.70$). The agreement was generally found to be higher at the six-month measurement than at the baseline. The proxies in this study also tended to report that the person they cared for had a worse quality of life than the individual’s self-report suggested, particularly in the self-care, pain/discomfort and anxiety/depression domains.

DEMQOL-Proxy

DEMQOL-Proxy, developed alongside DEMQOL, is a tool for eliciting the health-related quality of life (HRQOL) of people with dementia from their informal carer. DEMQOL-Proxy takes the format of a face-to-face interview and consists of 31 items and an overall question on quality of life. From these 31 items, two subscales – functioning and emotion – can be calculated. Importantly, DEMQOL-Proxy is different to DEMQOL in three main ways. First, DEMQOL has 28 items but only shares 14 of those with DEMQOL-Proxy. Second, DEMQOL produces different subscales (daily activities, memory, negative emotion and positive emotion). The reason, given by the developers, for these differences between the measures is that DEMQOL and DEMQOL-Proxy should be seen as complementary and not interchangeable (Smith et al., 2005a). They recommend using both tools in any study and specify that DEMQOL-proxy is not a tool to be used when an interview with the service users are not possible. However, there is preliminary evidence that it can be used without DEMQOL when the person is classified as having severe dementia (Smith et al., 2005b).

Finally, there is a difference in the way questions are phrased in DEMQOL compared to DEMQOL-Proxy, as the sample questions below illustrate. The questions in DEMQOL and DEMQOL-Proxy are best categorised as being evaluation-based measures.

Box 6: Sample questions from DEMQOL and DEMQOL-Proxy

DEMQOL (version 4)	DEMQOL-Proxy (Version 4)
I would like to ask you about your life. There are no right or wrong answers. Just give the answer that best describes how you have felt in the last week...	I would like to ask you about (your relative's) life, as you are the person who knows him/her best. There are no wrong or right answers. Just give the answer that best describes how... (your relative) has felt in the last week. If possible try and give the answer that you think (your relative) would give....
First I'm going to ask you about your feelings. In the last week have felt...	First I'm going to ask you about ... (your relative's) feelings. In the last week, would you say that... (your relative) has felt...
Cheerful A lot Quite a bit A little Not at all	Cheerful A lot Quite a bit A little Not at all
Now, I'm going to ask you about your everyday life. In the last week have you been worried about your everyday life. In the last week, how worried have you been about...	Now. I'm going to ask you about ... (your relatives) everyday life. In the last week, how worried would you say ... (your relative) has been about....
20. Not having enough company? A lot Quite a bit A little Not at all	28. Not having enough company? A lot Quite a bit A little Not at all

Smith et al. (2005b) and Smith et al. (2007) present the extensive development process behind both DEMQOL and DEMQOL-Proxy. First a conceptual framework was developed by drawing on relevant literature and qualitative interviews with people with dementia and their carers and expert opinion (Smith et al., 2005a). This initial work produced a pair of questions, one for the person with dementia and one for the carer, which were then tested and revised in the light of twelve pre-testing interviews with people diagnosed with dementia and their carers. Two matching 73-item versions were taken forward to the

preliminary field test which aimed to reduce the number of items by removing those with poor psychometric performance ratings, and begin an evaluation of subscales within the tool. The preliminary field test consisted of 130 people with dementia and 126 carers who were interviewed simultaneously but separately. Items were removed from the proxy and standard tools on the basis of missing data, endorsement frequencies and item redundancy. The result was two shorter tools with items both unique and in common. Whilst these resulting questionnaires did not match the conceptual framework closely, and may thus be challenged on content validity, the authors note the tools had much better psychometric properties than the initial version.

The aim of the final field test was to evaluate the acceptability, reliability and validity of the two questionnaires using standard psychometric methods, with one hundred and one people with dementia and 99 carers. In addition to using a number of other 'gold standard' measures of dementia and/or quality of life¹² in the interviews to validate DEMQOL/DEMQOL-Proxy, a test-retest subsample completed a baseline assessment of HRQL using DEMQOL/DEMQOL-Proxy and repeated the measurement at either two weeks or three months later. The broad finding of the final field test was that DEMQOL and DEMQOL-Proxy were comparable to the best available dementia-specific measures in mild and moderate dementia. More specifically, there was modest evidence of correlation between DEMQOL/DEMQOL-Proxy and the validating measures. Correlations between DEMQOL and DEMQOL-Proxy were found to be moderate for those with mild or moderate dementia and low for those who were classified as having severe dementia, highlighting the assertion that the two tools were complementary and not interchangeable. In other words, DEMQOL-Proxy assesses different aspects of HRQL than the DEMQOL.

Quality of Life in Alzheimer's disease (QoL-AD)

The QoL-AD is a tool for measuring quality of life of individuals with dementia, and consists of two separate tools: one designed to be given as an interview for the person with dementia and the other designed to be given as a self-completion questionnaire for the family caregiver. Both have the same thirteen items which, like many of the condition-specific tools, cover domains that are relevant to everyone (e.g. friends, ability to do things for fun and money) as well as those that are more relevant to people classified as having dementia (e.g. memory and mood). Each item is rated on a four-point scale running from poor to excellent. The sample questions below show the differences between the interview

¹² The mini-mental state examination (MMSE - a tool used to screen for cognitive impairment), The Dementia Quality of Life Instrument (DQoL – a tool for measuring quality in mild and moderate cases of dementia), The Quality of Life Assessment Schedule (QOLAS), The Geriatric Depression Scale (GDS) and Barthel Index

other theoretically-related measures, was also good. The initial testing did note that the presence of depressive symptoms in the caregiver were related to lower quality of life scores, but suggested that more work was needed to understand this finding.

Logsdon et al. (2002) further tested the reliability and validity of the QoL-AD with 177 patient/carer dyads. A key finding was that all of the 22 people with dementia who could not complete the QoL-AD scored lower than ten on the MMSE, leading to advice noted earlier that the QoL-AD is not suitable for administering to people with more severe dementia (in these cases it should be administered to the caregiver only as treated as a proxy tool). The study also confirmed the earlier findings of Logsdon et al (1999), that the reliability and validity of the QoL-AD were good.

Hoe et al. (2007) and Crespo et al. (2012), however, suggest that there are some problems with using the QoL-AD as a proxy measure. By comparing the scores of 102 residents with dementia with those of the carers and staff at a nursing home, Crespo et al (2012) found that whilst the scores of carers and staff tended to correlate with each other, they did not correlate well with resident scores. The study also found that the residents rated their quality of life significantly higher than the proxies rated the resident's quality of life. Hoe et al (2007) compared the factors that were perceived to affect the quality of life for people with dementia from both the perspective of the person with dementia and their caregivers in 191 dyads and noted that different factors were reflected in the quality of life scores produced by QoL-AD by each respondent. According to carer's reports, predictors of a higher quality of life were less depressive symptoms, less irritability, less apathy and less daily living impairment in the person with dementia and living at home¹³. Reports from people with dementia, on the other hand, stated that a higher quality of life was associated with not only having less depressive symptoms, but also the taking of acetyl cholinesterase-inhibitors¹⁴. The two papers come to the same conclusion: proxy ratings do not measure the same things as self-reports, that each person provides a different perspective and that proxy ratings using QoL-AD should not be treated the same as self-ratings.

Quality of Life Assessment Scale (QOLAS)

QOLAS is a tool for measuring the quality of life of people with dementia. Unlike the other tools presented here it combines both qualitative and quantitative approaches. In a semi-structured interview with the person with dementia (or the proxy, if that approach is being taken), the interviewee identifies what is important to them with respect to their quality of life (or what the proxy thinks is important to the person they care for in the case of the

¹³ Information on these factors was gathered by using a combination of tools in the study to supplement QoL-AD. These included MMSE, The Cornell Scale for Depression in Dementia, The Neuropsychiatric Inventory.

¹⁴ Medication to treat Alzheimer's disease.

questions being answered by a proxy) by picking two issues or “constructs” for each of the following domains: physical, psychological, social/family, usual activities and cognitive functioning. For each issue the interviewee has to provide a rating of how much of a problem is currently being experienced within that domain. In the proxy version the proxy has to report how much of a problem they think the person with dementia experiences within each domain. The scale runs from 0 (no problem) to 5 (it could not be worse). With the domains/issues combined a quality of life score ranging from 0-50 is produced with higher scores representing a poorer quality of life.

Selai et al (2001) outline the development and testing of the tool which was based on the Quality of Life Assessment by Construct Analysis (QoLASCA) a that measures the quality of life amongst epilepsy sufferers but was modified to study people with dementia.. The pilot study of this revised version involved ten patient/carer dyads and showed that, whilst patients with mild to moderate dementia could understand the questions, they did need some prompting to actually rate their quality of life. The pilot study also resulted in some small changes to the tool, including the replacement of the work/economic functioning domain with daily activities domain, and was then field tested. Thirty-seven patients/carer dyads were recruited although, due to severity of dementia, only 22 of these dyads could be included in the study. All of those who participated in the study had a MMSE score of between 11 and 30. The carers and patients were interviewed separately. Internal consistency of the scale, as measured by coefficient alpha was .78 for patients and caregivers, whilst construct validity was demonstrated by significantly ($p < 0.5$) higher patient reported quality of life in patients with lower levels of disability than in patients with higher levels of disability. The study found that there was quite a lot of overlap between the constructs chosen by carers and patients¹⁵. It also found that for each domain carers rated the patients as having a worse quality of life than the patients rated themselves. However, there was good agreement between carers and patients in all the domains except daily activities and cognitive functioning. The study also found that carer/patient agreement was stronger when the patient had only mild, rather than moderate, dementia.

Evaluation of Quality of Life Instrument (EQLI)

The EQLI is a tool for measuring the quality of life of adults with learning disabilities who live in residential settings. The English version of the tool has a section where a member of staff rates the residents’ level of disability in ten domains (see box below). Disability is rated between zero (no difficulty) and four (ability only occasionally present).

¹⁵ What is not so clear is if overlap is at the level of the individual dyad.

Box 8: EQLI Domains

Learning and applying knowledge
Communication
Mobility
Self-care
Domestic life
Interpersonal interactions and relationships
General tasks and demands (carrying out single or multiple task, organising routines and handling stress)
Major life areas (engage in education, work and employment)
Community, social and civic life

The rest of the tool consists of 14 items that ask the staff member to rate satisfaction they think the resident feels with an aspect of their lives (see below for sample questions). These items cover issues such as the accommodation, staff interaction, privacy, activities and social interaction and each item is rated between 1 (not at all satisfied) and 5 (very much satisfied). There are also options for the staff member to report either that the question is not relevant or that they cannot answer it. As the sample questions illustrate, the questions are evaluation-based.

Box 9: Sample items from EQLI

The hygiene and tidiness of the Centre/Institution?

The time devoted to him/her by the staff members?

The type of recreational activity he/she carries out?

The opportunity to increase the number of social interactions outside the Centre/Institution?

Nota et al. (2006) is the key English language article regarding the development and testing of EQLI¹⁶. During the first stage of the development, 50 items derived from consultation with experts were created and then piloted with 100 adults with varying levels of intellectual disability in northern Italy. The focus of this test was to look at both the psychometric properties of the tool and to gauge its intelligibility. Following this pilot 16 items were removed for either low intelligibility or distributional properties. The next stage of development saw the remaining 34 items administered to 305 adults with intellectual disabilities who attended day centres in Northern Italy and 305 corresponding members of

¹⁶ Two earlier Italian articles outline the early development of EQLI, although the process is summarised in Nota et al. Nota L, Soresi S & Perry J (2006) Quality of life in adults with an intellectual disability: the Evaluation of Quality of Life Instrument, *Journal of Intellectual Disability Research*, 50, 371-385..

staff who had known the person for more than one year. An item and factor analysis removed sixteen items due to either low discriminant ability or loading on more than one factor, and the result was the 18-item, three-domain tool (as outlined above). Nota et al. (2006) also presented details of further testing of the EQLI with staff working with 248 participants with intellectual disabilities in Northern Italy. The article presents findings on the EQLI's internal consistency, convergent and discriminant validity from data collected from staff who had worked with the person with an intellectual disability for over a year. The tool was shown to have high internal consistency (0.90) and the items in the scale were shown to be related to one another. Neither test-retest nor inter-rater reliability were tested.

Quality of Life Interview Schedule (QUOLIS)

The Quality of Life Schedule (QUOLIS) was developed in the late 1980s to fill what its creators thought was a gap in available tools to measure the quality of life experienced by adults with learning disabilities. It is based on the idea that disability entails a need for either support or assistance, and reflects this in its measurement of quality of life. The tool takes the form of a semi-structured interview with a proxy, who should be somebody who knows the person with the intellectual disability well (e.g. a close family member or friend, or a support worker who has worked with them for some time), and consists of questions covering twelve domains which include income maintenance, education/employment, social/recreational and advocacy. In each of the domains the proxy is asked about four dimensions — availability of support, accessibility of support, chosen level of participation for or by the disabled individual, and the disabled person's 'apparent' level of contentment with the current situation — and their responses are rated by the interviewer on a scale of one to seven. It is unclear from the documentation whether guidance was given to the interviewer to aid their coding of the proxy's responses. The actual development of the tool is not documented in published literature.

Ouellette-Kuntz (1990) outlines the testing of the tool's test-retest reliability and inter-rater reliability using ten adults with a diagnosis of 'mental retardation' and a willing proxy. Three people rated the responses given by the proxy: two were present at the interview, whilst the other rated from a video recording of the interview. One of the raters present at the interview also conducted a second rating from the video recording at least two weeks after the initial interview. The study found very high levels of both test-retest reliability - and inter-rater reliability. For example, 73% of intra-class correlation coefficients for test-retest agreement were found to be either almost perfect (above 0.80) or substantial (0.60 or above), the level for intra-class correlation coefficients for inter-rater agreement was 65%. However, as Ouellette-Kuntz (1990) points out, all the raters were involved in the development of the tool and agreement might not be so high in raters who were less familiar with the tool and the principles that lie behind it.

The Comprehensive Quality of Life Scale – Intellectual Disabilities (COM QOL-ID)

The Comprehensive Quality of Life Scale – Intellectual Disabilities (COM QOL-ID) is an adaptation of the general Comprehensive Quality of Life Scale. It contains seven domains, (material well-being, health, productivity, intimacy, safety, place in community and emotional well-being), which were drawn from reviews of relevant literature, and are measured both ‘objectively’ and ‘subjectively’.

Three items are combined to form the ‘objective’ measure for each domain. In the material well-being domain, for example, the ‘objective’ measure draws on questions about level of income, type of accommodation and personal possessions (see Box 10). The ‘objective’ questions are sometimes highly factual (e.g. where do you live) but are mainly perception-based. However, some questions, such as ‘how safe do you feel’, are best characterised as evaluation-based measures. The idea behind the ‘objective’ section of the instrument is that the person with intellectual disabilities will be able to answer the questions, albeit with the help of their carer.

The ‘subjective’ section of the instrument has versions for both the person with intellectual disabilities and their close informal carer, who acts as a proxy, to complete. Both versions ask about the importance of each domain (which feeds into the weighting of the domain in the final score) and satisfaction with that area of life. There are significant differences between how the questions are worded in each version: the version for people with intellectual disabilities spends quite a while determining the respondents’ ability to comprehend the scales and uses either a two, three or five point scales accordingly. The proxy version, in contrast, always uses a five-point scale. The questions, as shown below, are evaluation-based measures.

Box 10: Sample questions from COM QOL-ID

Material Well Being

Objective Questions

1 (a) Where do you live?

- A house
- A flat or apartment
- A room (e.g. in a hostel)

Do you own the place where you live or do you rent?

- Own
- Rent

(b) How many personal possessions do you have compared with other people? (prompt: think about the things you own, like your clothes, furniture etc.)

- More than almost anyone
- More than most people
- About average
- Less than most people
- Less than almost anyone

(c) What is your personal or household (whichever is most relevant to you) gross annual income before tax?

- Less than \$10,999
- \$11,000 - \$25,999
- \$26,000 - \$40,999
- \$41,000 - \$55,999
- More than \$56,000

Subjective Proxy Questions

1. How important to (client) are the things he/she owns?

- Could not be more important
- Very important
- Somewhat important
- Slightly important
- Not important at all

1. How satisfied is (client) with the things he/she owns?

- Delighted
- Pleased
- Mixed
- Unhappy
- Terrible

Cummins (1991) and Cummins (1997) outline the background to the tool and present information upon its psychometric properties. Both articles concentrate upon its use with people with intellectual disabilities, where it is shown to have some validity, but not its use with proxies. Cummins (1997) actually gives a warning in the case of use with proxies. Drawing on material presented in a conference paper, he suggests that his data and also the bulk of data found in other studies show that interpretation of proxy data should be carried out cautiously. He continues by suggesting that even proxies who are close to the person they are rating do not manage to supply accurate ratings on subjective measures of the other persons' quality of life. Despite this, COM QOL-ID does have a proxy tool designed for carers. Moreover, it has been used in more recent work looking at proxies, although the results are not presented (Perry and Felce, 2002).

Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39)

The Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39) is an adaptation of the SS-QOL (see below) designed specifically for individuals with post-stroke aphasia. It consists of 39 items, grouped into four domains (physical, communication, psychosocial and energy) and rated on a five-point Likert scale, running from one ('could not do at all'/'definitely yes') to five ('no trouble at all'/'definitely no'). Overall and subdomain scores run from one to five. It is primarily designed to be a self-report instrument and it is claimed by Hilari et al. (2003) that as almost 90 per cent of post-stroke aphasia patients are able to self-report in an interview there is minimal need for a proxy respondent. Nonetheless, there is a proxy version of the tool.

The use of the proxy tool requires no training and is designed to be completed by somebody who sees the service user at least twice a week. Box 11 below shows some sample questions from the proxy version and the original version from which they were adapted. It shows that the questions asked of the proxy are both evaluation based and ask the proxy to answer from the perspective of the service user. The published work involving the proxy version discusses stroke survivor/proxy agreement.

Box 11: Sample questions from SAQOL

SAQOL-39 Proxy version	SAQOL-39
<p>We would like to know how _____ is doing with activities or feelings that can sometimes be affected by stroke. Each question will ask about a specific activity or feeling. For each question, think about how that activity or that feeling has been in the past week.</p> <p>PLEASE ANSWER EACH QUESTION FROM _____ PERSPECTIVE, i.e. AS YOU THINK _____ WOULD.</p>	<p>We would like to know how you are doing with activities or feelings that can sometimes be affected by stroke. Each question will ask about a specific activity or feeling. For each question, think about how that activity or that feeling has been in the past week</p>
<p>How much trouble did s/he have: Preparing food?</p> <ul style="list-style-type: none"> • Couldn't do at all • A lot of trouble • Some trouble • A little trouble • No trouble at all 	<p>How much trouble did you have: Preparing food?</p> <ul style="list-style-type: none"> • Couldn't do at all • A lot of trouble • Some trouble • A little trouble • No trouble at all

SAQOL-39 Proxy version	SAQOL-39
<p>How much trouble did s/he have: Getting other people to understand him/her?</p> <ul style="list-style-type: none"> • Couldn't do at all • A lot of trouble • Some trouble • A little trouble • No trouble at all <p>Did s/he: Feel discouraged about his/her future?</p> <ul style="list-style-type: none"> • Definitely yes • Mostly yes • Not sure • Mostly no • Definitely no <p>Did s/he: Go out less often than s/he would like?</p> <ul style="list-style-type: none"> • Definitely yes • Mostly yes • Not sure • Mostly no • Definitely no 	<p>How much trouble did you have: Getting other people to understand you?</p> <ul style="list-style-type: none"> • Couldn't do at all • A lot of trouble • Some trouble • A little trouble • No trouble at all <p>Did you: Feel discouraged about the future?</p> <ul style="list-style-type: none"> • Definitely yes • Mostly yes • Not sure • Mostly no • Definitely no <p>Did you: Get out less often than you would like?</p> <ul style="list-style-type: none"> • Definitely yes • Mostly yes • Not sure • Mostly no • Definitely no

Hilari et al (2007) presents the findings of a study looking at proxy/self-report agreement on SAQOL-39. Following assessment of aphasia using the Frenchay Aphasia Screening Tool (FAST), 50 people with aphasia¹⁷ and their proxies, defined specifically as someone who saw the person with aphasia at least twice a week, were recruited into the study. . The aphasiac respondents completed SAQOL-39 in a face-to-face interview with a speech and language therapist with expertise in aphasia, whereas the proxies completed the tool via a telephone interview. Agreement between these groups of respondents, as measured by intra-class correlation coefficients, was found to differ by domain with physical being excellent (0.80), psychosocial and communication being good (0.70) and energy being fair (0.50). Agreement on the overall scale was excellent (0.80) though proxies tended to report a lower quality of life for people with aphasia than the people with aphasia reported for themselves.

¹⁷ Only people with a FAST score of 7/15 we consider able to communicate sufficiently to take part in the study.

Stroke-Specific Quality of Life (SS-QOL)

The Stroke-Specific Quality of Life (SS-QOL) measure is a quality of life measure designed specifically for adults who have suffered a stroke and consists of 49 items grouped into 12 domains which cover such areas as family and social roles, energy, language, self-care and mood. Each item is measured on a five-point scale which runs from one (total help/couldn't do at all/strongly agree) to five (no help needed/no trouble at all/strongly disagree). Whilst originally developed and tested for self-report (Williams et al., 1999), a proxy version was created but details of how it differs from the original version or how it was developed are not available in the public domain. However, the implication from the original is that the questions in the proxy version are evaluation-based measures. Details of studies looking at proxy and stroke survivor agreement are, however, in the public domain.

Williams et al. (2006) look at the agreement between ratings of quality of life using SS-QOL given by stroke survivors and family carers. The study used the responses of 225 patient/carer dyads drawn from people enrolled in a clinical trial for post-stroke depression and found that proxies rated the quality of life of patients slightly worse than the patients rated their own quality of life. The greatest disparity was found in the following domains: mood, energy and thinking. Using paired t-tests and intra-class correlation coefficients to measure agreement, the authors concluded that agreement was modest at best. The study also found that both patient depression and the proxy's perception of the burden of caring impacted upon agreement – with agreement being higher where patient depression scores were high and lower when the proxy reported higher levels of caregiver burden. In conclusion, Williams et al (2006) suggested that the differences they found between the scores of people post-stroke and proxies were large enough to impact upon the findings of clinical trials.

Summary

This section has reviewed some of the key tools for measuring quality of life via a proxy respondent. We found that almost all of the tools were adapted or developed for specific client groups, primarily people with dementia or learning disabilities or who had suffered a stroke, and that none of the tools we found were designed for use with a range of client groups. As we have already stated, given the difficulties we had conducting this review and identifying quality of life instruments for proxy respondents, we cannot be certain that we have exhaustively covered all the available tools. Nevertheless, we feel this review gives a feeling for the types of tools that are available and the way that researchers use and have approached the development of quality of life tools for proxies.

As **Table 3** summarises the majority of the tools use evaluation-based questions. The researchers appear not have considered the theoretical problems associated with this choice and largely formulated questions that ask proxies to answer either as if they were the recipient or as how they think the recipient would have answered had he/she been able.

It is probably that such an approach is extremely challenging and prone to error as the proxy is unlikely to utilise the same judgement criteria as the recipient would have used. The ADRQoL instrument is an exception in this respect since it uses perception-based questions. The researchers appear to have made a decision to develop questions that proxies could answer easily based on their observations of the recipient – a strategy that is likely to yield more consistent responses, at least as predicted by theory (Schwartz and Rapkin, 2004).

Table 3: Summary of the question types and proxy requirements of the reviewed tools

Instrument name	Type of question	Differences to self-report version	Restrictions around use of proxy
Alzheimer’s Disease-Related Quality of Life (ADRQL)	Perception	No self-report version	Must be a caregiver (formal or informal/paid or unpaid)
Camberwell Assessment of Need for the Elderly (CANE)	Evaluation	None	None Used staff, informal caregiver
Euroqol (EQ-5D)	Evaluation	Yes, verb changed person	None Used formal/informal caregiver
DEMQOL-Proxy	Evaluation	Yes, different number of items, changed tense	Informal carer Not for use without self-report version
Quality of Life in Alzheimer’s disease (QoL-AD)	Evaluation	Yes, verb changed person	Family caregiver Proxy only version if MMSE<10
Quality of Life Assessment Scale (QOLAS)	Evaluation	Yes, verb changed person	None
Evaluation of Quality of Life Instrument (EQLI)	Evaluation	No self-report version	Staff
Quality of Life Interview Schedule (QUOLIS)	Evaluation	No self-report version	Family, close friend or support worker, Must know person well

Instrument name	Type of question	Differences to self-report version	Restrictions around use of proxy
The Comprehensive Quality of Life Scale – Intellectual Disabilities (COM QOL-ID)	Factual, perception and evaluation	Yes, verb changed person and different response options	Close informal carer
Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39)	Evaluation	Yes, verb changed person	Sees person at least 2 times/wk
Stroke-Specific Quality of Life (SS-QOL)	Evaluation	Details not published	Used with carers who saw person at least 2 times/wk

The lack of consideration given to the format of the questions is reflected in the development of the instruments as shown in **Table 4**. Very few of the instruments underwent any significant development work with proxy respondents. The ADRQoL and DEMQOL are exceptions since both were designed with proxy input and were extensively tested with proxies. The ADRQoL is particularly noteworthy in this respect since they used focus groups with proxies to establish the content and wording of the items for the areas of quality of life that were of interest. Some other instruments were piloted with proxy respondents, but in too many cases there was no development with proxies. Indeed, proxy versions of tools often seem like an afterthought of the researchers and there is little consideration of how appropriate they may be for proxy respondents to answer.

For the instruments studied in this review, researchers had in many cases given some thought to who is the best person to act as a proxy, although the rationale behind the choice is rarely made explicit. Restrictions around the choice of proxy depend very much on the instrument with some stipulating that the proxy must be a caregiver and others that the proxy must be a relation. Some tools goes further and make restrictions around how well or how long the person has known the recipient for, or how frequently the person sees the recipient (see **Table 3**). It is not clear whether these decisions are based on research or on philosophical positions. The QOL-AD was interesting in that it based the decision on when a proxy should be used (when the MMSE score is less than 10) on research that demonstrated that people with MMSE scores <10 did not seem to understand the self-report version of the instrument. This is rare; QOL-AD is the only instrument to consider when a proxy response should be sought.

Table 4: Summary of the question types and proxy requirements of the reviewed tools

Instrument name	Method used	Details of development	Details of testing
Alzheimer’s Disease-Related Quality of Life (ADRQL)	Face-to-face structured interview	Development of items with caregivers and professionals Cognitive testing of items with caregivers Scale development	Construct and concurrent validity assessed Sensitivity assessed No inter-rater reliability
Camberwell Assessment of Need for the Elderly (CANE)	Face-to-face structured interview	Development with professionals Piloted with older people, caregivers and staff	Face, content , construct and concurrent validity Test-retest reliability Proxy-subject agreement assessed
Euroqol (EQ-5D)	Self-completion	None reported	Proxy-subject agreement assessed
DEMQOL-Proxy	Face-to-face structured interview	Development of items with caregivers, as well as patients and experts Scale development Piloted with caregivers and patients	Construct and concurrent validity assessed Sensitivity assessed Proxy-subject agreement Test-retest reliability
Quality of Life in Alzheimer’s disease (QoL-AD)	Self-completion	Draft reviewed by caregivers, as well as patients and experts Piloted with patients and caregivers Scale development	Test-retest reliability Concurrent and construct validity assessed Proxy-subject agreement

Instrument name	Method used	Details of development	Details of testing
Quality of Life Assessment Scale (QOLAS)	Semi-structured interview	Piloted with patients and caregivers	Construct validity Sensitivity assessed Proxy-subject agreement
Evaluation of Quality of Life Instrument (EQLI)	Self-completion	Consultation with experts and piloting Scale development	No test-retest or inter-rater reliability Convergent and discriminant validity
Quality of Life Interview Schedule (QUOLIS)	Semi-structured interview	Not documented	Test-retest reliability Inter-rater reliability
The Comprehensive Quality of Life Scale – Intellectual Disabilities (COM QOL-ID)	Structured interview	not documented for proxy version	Only self-report version tested Implied proxy-subject agreement assessed
Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39)	Telephone interview	not documented for proxy version	Proxy-subject agreement assessed
Stroke-Specific Quality of Life (SS-QOL)	Not reported	not documented for proxy version	Proxy-subject agreement assessed

The versions designed specifically for proxy respondents were assessed for validity, reliability and, in some cases, sensitivity, and were treated as standalone measures of quality of life as summarised in **Table 4**. Researchers frequently examined validity by correlating scores generated by their tools against those from other theoretically relevant or conceptually similar instruments. They also frequently assessed test-retest and, where relevant, inter-rater reliability, and in the instances of the ADRQoL and the DEMQOL-proxies, subjected the instruments to scale development, thus creating in the case of DEMQOL a different self-report and proxy version. For many instruments, however, the main focus was on proxy-recipient agreement.

A clear picture emerges from this literature that there are differences between proxy and self-reports, and that proxy reports should not be seen as substitutes for self-report. The creators of some tools, such as the DEMQOL/DEMQOL-Proxy, acknowledge this and make it clear that the proxy should not be used as a substitute for the service user and that both proxy and self-report instruments should be administered. However, this strategy seems difficult to maintain where the recipient lacks the capacity to answer the questionnaire, as is clearly the case for people with severe cognitive impairments or intellectual disabilities. What is interesting is that similar findings emerge from the studies: in particular, that the extent of proxy-recipient agreement depends on the characteristics of the proxy, the characteristics of the user and can vary according to the subject of the question. Thus, family members are usually found to report lower quality of life than is self-reported whereas professionals often report higher levels of quality of life than is self-reported. It has also been noted that people with mild or moderate dementia show more agreement with proxy reports than those with severe dementia. Although this picture is relatively consistent, the literature we have found does not explore whether the effects of proxy characteristics are mediated by other factors, such as the length of time for which the proxy has known the recipient or how frequently the proxy sees the recipient. While studies have looked at factors that may be associated with higher levels of disagreement between proxies and service users (carer burden and carer depression being two examples), the body of work presented here does not, in any systematic way, consider whether agreement would be better were the questions phrased as perception-based rather than evaluation-based responses, as theory would predict may be the case (Schwartz and Rapkin, 2004).

Studies of service user/proxy agreement are often presented as assessments of the reliability of the proxy measure. Our opinion is that this approach is misguided since it portrays one of the versions (self-report or proxy-report) as the accurate version. Particularly where questions are evaluation-based both reports must be considered to be accurate reflections of that individual's evaluation of the facts that they possess. The issue is whether either version accurately captures what the researcher intended to measure. A proxy-user agreement study could not answer this question. Moreover, studies looking at proxy/service user agreement are often, by their nature and purpose, carried out with

service users whose memory and cognitive abilities mean that they would not in a normal study have their quality of life rated by only a proxy. Such studies can therefore say nothing about the validity or reliability of proxy reports for severely cognitively impaired or intellectually disabled people. A far more fruitful approach is found in the ADRQOL development where researchers analysed the data to see whether the proxy's characteristics were associated with their responses. They found no evidence of association and concluded therefore that the proxy reports were not biased. This approach combined with an assessment of the reliability and validity of the instrument on its own terms, against, for example, measurements gained from observational tools seem to be a more sensible strategy for testing the reliability and validity of proxy reports.

Proxy tools do offer a way of gaining insight into the quality of life experienced by service users whose impairments would make participation in a postal survey impossible. It is clearly important to recognise that proxy reports can never be a substitute for self-reports and that this limitation has implications for the analysis and interpretation of the data gathered. However, this should not be seen as a barrier to adopting proxy tools as a way of gathering the experiences of people who are unable to provide their own views in a structured format. A more important consideration is how to implement proxy tools within a postal questionnaire. The proxy tools identified in this review used either face-to-face interviews or handed a self-completion questionnaire directly to the proxy to complete. Identifying when proxy versions are required, choosing an appropriate proxy respondent and further engaging the proxy to participate in the process (so overcoming many of the problems associated with nonresponse identified in earlier sections of this report) are practical challenges to implementing such an approach within the ASCS.

Observational Tools

Observational methods are another way of eliciting information about social care outcomes from people who are unable to answer survey questions (Mansell, 2011) and use trained observers to provide the quality of life assessment. The observer usually studies the behaviour of specific individuals but, in some cases, may also observe the behaviour of others interacting with that individual and / or the environment. Our interest here is in tools that allow for structured observation and can therefore be used to produce a consistent and replicable measurement of quality of life through observation. Our key concern is to understand the processes researchers have gone through to develop structured observational instruments. In particular, we are interested in:

- Are there any restrictions around who can be an observer/assessor, including training requirements?
- What is the focus of the observation (behaviours or environmental surroundings)?
- How observations are structured, in terms of time periods for observation, classification of behaviours/environmental surrounds?

- Are there any restrictions on what can be observed?
- The level of detail and method for presenting instructions to map observations to produce scores?
- How the researchers have assessed the validity and reliability of the instrument?

The methods used to identify observational tools for this review are outlined in Appendix 2. Similar to the review of proxy tools, we found it difficult to identify observational tools using standard systematic review methods. We therefore used a more ad-hoc approach and relied upon the advice of experts in observational methods. For this reason, we cannot be certain that we have identified all the observational tools available to measure quality of life. Nevertheless, we feel that our review gives a flavour of the approach taken by researchers in developing observational tools and the usage of such tools.

Our review identified 5 quality of life instruments. Not surprisingly, all of the tools we found are designed for use within residential settings, such as care homes, nursing homes and sheltered housing. Observational work in these settings is well-established (Clark, 2007), although much of it is inspired by a more ethnographic and qualitative approach. Since we are specifically interested in the tools that allow for structured observation, such work is not included within this review. Although all the tools were designed for use in either care or nursing homes, there has been some work on adapting them for use in either sheltered housing or hospital settings. Two of the tools were designed specifically for use with people with dementia¹⁸, whereas the other three were designed more generally for either populations within care homes or people who were unable to report their quality of life via an interview or survey.

¹⁸ DCM, whilst developed for use with people with Dementia has been tested with other groups such as adults with learning disabilities and older adults with physical impairments.

Table 5: A summary of the characteristics of the reviewed observational tools

Instrument name	Client group	Setting	Main usage	Other uses
Dementia care mapping (DCM)	Dementia, but tested with physically impaired and learning disabled	Care home, but tested in acute and community hospitals	Intervention for improving practice	Research Shorter version (SOFI) for inspection (see below)
Short observational framework for inspectors (SOFI)	People unable to participate in interview	Care home	Inspection	None
Therapeutic Environment Screening Survey for Nursing Homes (TESS-NH)	Dementia	Care home	Research	None
Observable Indicators of Nursing Home Quality Instrument (OIQ)	Not specific	Nursing home Adapted for residential care and assisted living	Research	Measuring quality for any purpose
Adult Social Care Outcomes Toolkit (ASCOT)	Not specific	Care home	Research	Improving practice

For each of the instruments outlined in **Table 5** we describe the features of the tool and its development, focusing on addressing the six questions set out above: that is, the structure and focus of the observations, the choice of and restrictions surrounding who is an observer, instructions regarding the mapping of observations for scoring purposes and the validity and reliability of the measure. Details for some of the measures are sparse and we were not able to find information regarding the development of the tool or obtain access to the tool, despite sending requests to the corresponding author. The availability of evidence for each of the tools is summarised in **Table 6**.

Table 6: Summary of the types of documentation available for each instrument

Instrument name	Copy of instrument obtained	Details of development and testing available
Dementia care mapping (DCM)	No	Limited details, development largely undocumented
Short observational framework for inspectors (SOFI)	No	Very limited details
Therapeutic Environment Screening Survey for Nursing Homes (TESS-NH)	Yes	Yes, quite detailed
Observable Indicators of Nursing Home Quality Instrument (OIQ)	Yes	Yes, quite detailed
Adult Social Care Outcomes Toolkit (ASCOT)	Yes	Yes, quite detailed

Dementia Care Mapping (DCM)

Dementia Care Mapping (DCM) is a tool that was originally developed as an intervention to promote person-centred care for people with dementia. However, it has been used reasonably extensively in research (Sloane et al., 2007), including in cross sectional studies of quality of life and in evaluations of interventions to measure the quality of life of people with dementia (Brooker, 2005). Drawing on the work of Kitwood, the DCM tool aims to make the observer see the world from the point of view of the person with dementia (Brooker, 2005; Edvardsson and Innes, 2010; Kuhn et al., 2000). It is currently in its eighth incarnation (Brooker and Surr, 2006). People wishing to use the tool have to attend a four day course and pass a written examination to ensure their competence in carrying out the mapping process and their understanding of its philosophical basis. Reflecting its origin as an intervention, the tool is aimed very much at practitioners, although as noted above, it has also been used by researchers as a data gathering method.

Using the DCM tool, the observer continuously observes (or ‘maps’) five people in a communal area of a residential establishment over, what should be, a representative time period. The associated guidance suggests that the full DCM observation consist of six hours on two consecutive days so that the bulk of the waking day is observed. For example, 2.00pm to 8.00pm on day one, followed by 8.00am to 2.00pm on day two (Innes 2002). Each participant is coded on two scales, capturing behaviour and mood respectively, every five minutes (the time frame used by the tool).

The scale capturing behaviours is called the Behavioural Category Code (BCC) and describes what the participant is primarily doing during the past five minutes. The tool contains 24 BCCs, which cover a range of activities from sleeping to interacting with others verbally. Box 12 shows the behaviour codes and descriptions for DCM 8 in order to illustrate the range of behaviours DCM attempts to record.

Box 12: DCM8 behaviour categories and codes

General description of category and code
Interacting with others verbally or otherwise [A]
Being engaged but passively (watching) [B]
Being disengaged, withdrawn [C]
Self-care [D]
Expressive or creative activities [E]
Eating or drinking [F]
Reminiscence and life review [G]
Prioritising the use of intellectual abilities [I]
Exercise or physical sport [J]
Walking, standing or moving independently [K]
Leisure, fun and recreational activities [L]
Sleeping, dozing [N]
Receiving practical, physical or personal care [P]
Engaging in a religious activity [R]
Sexual expression [S]

General description of category and code

Displaying attachment to or relating to inanimate objects [O]

Direct engagement of the senses [T]

Attempting to communicate without receiving a response [U]

Work or work-like activity [V]

Repetitive self-stimulation of a sustained nature [W]

Episodes related to excretion [X]

Interaction in the absence of any observable other [Y]

Fits none of existing categories [Z]

(adapted from Brooker and Surr 2006)

The scale capturing mood is called the mood/engagement value (ME value, previously known as the well or ill-being value (WIB))¹⁹ and runs from +5 (very happy, cheerful, very high positive mood/very absorbed, deeply engrossed/engaged) to -5 (very distressed, very great signs of negative mood/withdrawn and out of contact). The ME scores can be aggregated over the observation period to create an overall ME value score.

In addition to these two scales, personal distractions (PDs), staff behaviours which undermine the personhood of participant, and positive events (PEs), events that enhance the personhood of the participant, are recorded as they occur during the total observation period. Box 13 provides some example PDs and PEs.

Box 13: DCM8 personal detraction and enhancer types

Personal Detraction type	Personal Enhancer type
Comfort <ul style="list-style-type: none">• Intimidation Identity <ul style="list-style-type: none">• Infantilization	Comfort <ul style="list-style-type: none">• Warmth Identity <ul style="list-style-type: none">• Respect

¹⁹ As well as renaming the WIB as ME, version 8 simplifies this aspect of the DCM tool, which had been critiqued in evaluations of the psychometric properties of DCM.

Personal Detraction type	Personal Enhancer type
Attachment <ul style="list-style-type: none"> • Accusation Occupation <ul style="list-style-type: none"> • Objectification Inclusion <ul style="list-style-type: none"> • Mockery 	Attachment <ul style="list-style-type: none"> • Acknowledgement Occupation <ul style="list-style-type: none"> • Collaboration Inclusion <ul style="list-style-type: none"> • Fun

The origin of DCM as an intervention has implications for both its documentation and its use as a tool to measure outcomes. The initial development of DCM is not documented in the public domain and the early years of the tool saw very little published work on its psychometric properties compared with its widespread usage. This situation has now changed and details of the testing of version eight of the DCM tool (DCM8) are outlined. Nevertheless, as Edvardsson and Innes (2010) argue, the dual usage of DCM, both as an intervention and measure of outcome, is problematic since measurement and intervention are not separate or, as they argue the, “technique/process is the same as the outcome” (844).

Sloane et al. (2007) re-analyse various data sets from a number of studies to examine the psychometric properties of the seventh version of DCM (DCM7). Two studies reviewed explored inter-rater reliability using ‘Well and Ill-being’ (WIB) scores which, the authors felt, were the best indication of quality of life. They found only a moderate correlation ($r=0.32$) between the scores of different pairs of observers and thus concluded that WIBs had ‘limited potential’ as a global measure of quality of life. Concurrent validity was tested by looking at data from a study where proxy-based quality of life tools were used alongside DCM. The study found there to be only moderate correlations ($r=0.28$ to 0.40) between WIB and DCM scores and the scores from the proxy tools QOL-AD and ADRQOL. Even lower correlations were found when resident-reported QOL-AD scores were compared with DCM scores. Tests of construct validity, on the other hand, showed that whilst resident cognitive impairment, functional impairment, social withdrawal, agitation, depressive symptoms and a number of comorbid conditions had significant negative associations with WIB scores, age, gender and race did not. The authors of the review conclude that although there is low correlation between DCM and other measures of quality of life in dementia they still feel that the DCM has potential as a research tool. They argue, for instance, that the philosophical underpinnings of the tool may mean it gets closer to seeing quality of life from the perspective of the person with dementia than other tools. The authors also note that since the collection of these data, the DCM has gone a major update (from version seven to

version eight) which addresses some of the concerns they raise. We were, however, unable to find any studies looking at the validity and reliability of the new tool.

Adaptations of Dementia Care Mapping

Although developed as a tool for use with people with dementia in residential setting, the DCM has been adapted both for other settings and for different groups of people.

Woolley et al. (2005) reported on a study which looked at the feasibility of using DCM8 with physically rather than cognitively impaired people within a hospital setting. The mapping was carried out in the elderly care wards of three general hospital and the day areas of two community hospitals. The study found that there was a high level of missing data (27%), the primary reason being that the participant was not in the mapping area during that time-frame. This was most acute in community hospitals, which has 40% missing data. However, the study did find that the coding framework of DCM8 worked very well with the activities they observed, with observers being unable to code the observed activity in only two time frames (0.06%). The study concluded that it may be feasible to use DCM in hospitals with older adults with physical needs but that it would only be practical in wards with multi-occupancy bays — and not single rooms — so that patients could be observed. The authors also suggested that further work was needed, particularly around the practical and cultural issues of observations in hospital wards, and that minor adaptations to the manual were needed to include examples of physical illness.

Persaud and Jaycock (2001) looked at an application of the DCM tool to assess the quality of life of 22 adults with learning disabilities in three residential services . All the participants were aged between 20-63 years of age and were classified as having severe mental impairment, but not dementia. Two observers made a total of 32 hours of observations. The study was largely positive about the use of DCM in this setting but did find small problems with the DCM tool codes. While most of the codes used in the tool were found to be relevant to the lives of adults with learning disabilities, Persaud and Jaycock (2001) found that there were issues with the interpretation of some codes, owing to the extended range of behaviours displayed by this group which, they argued, could have a different meanings compared to those with dementia. For example, whilst participants with learning disabilities were observed walking, as described in the DCM manual, they were, in fact, pacing in a self-stimulatory manner. In dementia, walking is regarded as a positive sign of the ability to be mobile, and is coded thus in the DCM. Despite these problems, the authors thought DCM provided rich and useful data from which not only could they produce scores but also they could draw meaningful conclusions about the quantity and quality of behaviours found amongst adults with learning disability. They also noted the ability of DCM to feed into positive changes within organisations that care for adults with learning disabilities and subsequently for the lives of the adults with learning disabilities themselves.

Later work on adults with learning disabilities and dementia was also very positive about DCM applicability to this field (Finnamore and Lord, 2007).

Short Observation Framework for Inspectors (SOFI)

DCM has also been adapted specifically for use as an inspection tool in care homes. The Short Observation Framework for Inspectors (SOFI) was developed by the Bradford Dementia Group and the Commission for Social Care Inspection (now the Care Quality Commission) (Brooker, 2007; Commission for Social Care Inspection, 2008) and is a shortened version of the DCM tool that is less resource intensive for inspectors to use. The observation tool is just one part of the inspector's toolkit and is used when care home residents are unable to verbally describe their experiences. The SOFI looks at and rates mood, engagement and staff interaction during an observation period that last a couple of hours and it usually done over a lunchtime period. During the observation period five residents and their activities are recorded. It is only available to the statutory social care inspection bodies of England, Scotland and Wales.

Brooker (2007) outlines the development of the SOFI. The initial tool was composed by drawing on the experiences of the DCM team, inspectors, providers and care home residents. Seventeen inspectors were trained to use the initial SOFI tool and it was used in fourteen inspections. Feedback from these inspections, both from the inspectors themselves and the care-providers, was used to shape the final tool.

The Therapeutic Environment Screening Survey for Nursing Homes (TESS-NH)

The Therapeutic Environment Screening Survey for Nursing Homes (TESS-NH) is an observational tool for gathering data on the physical environment of a long-term care facility. So while it is not a tool that directly explores quality of life, it comprises 80 items which capture environmental characteristics that are thought to be of importance to people with dementia. Data from 18 items in the TESS-NH tool can be used to calculate the Special Care Unit Environmental Quality Scale (SCUESQS), a measure of environmental quality from the perspective of dementia care. TESS-NH (and SCUESQS) does not measure environmental quality for an individual residents but the environmental quality of the care facility.

The tool is based on highly structured observations that are completed in a 15-45 minute walk through of the nursing home, carried out between 9.00 and 17.00 on a weekday (but not during a mealtime). The highly structured nature of the tool means that training is not necessary and guidance on the rating of items is provided in the manual. The items in the tool, a sample of which are shown in box 14, are grouped under twelve domains (see box 15) including those that are generally found in quality of life measures, such as safety and cleanliness, and others which relate much more to environment, such as lighting and exit control. All of the items are also mapped on to a therapeutic goal — safety/security/health, orientation, privacy/control/autonomy and social milieu — that reflects quality of life.

Box 14: Sample items from TESS-NH

Cleanliness

Rate the general cleanliness of each of the following areas

- a. Shared social spaces
 - Very clean
 - Moderately clean
 - Poor level of cleanliness
- b. Halls
 - Very clean
 - Moderately clean
 - Poor level of cleanliness

Safety

To what extent are handrails present on the unit

- a. Hallways
 - Extensively
 - Somewhat
 - Little or none
- b. Bathrooms
 - Extensively
 - Somewhat
 - Little or none

Orientation/cueing

1. Residents rooms
 - a. Doors routinely left open
 - Yes
 - No
 - b. Residents names on/near door
 - Yes
 - No

Box 15: TESS-NH Domains

Unit autonomy	Space / seating
Exit control	Familiarity / homelikeness
Maintenance	Visual / tactile stimulation
Cleanliness	Access to outdoors
Safety	Orientation / cueing
Lighting	Privacy

Sloane et al. (2002) charts the development of the TESS-NH tool, which is based on an exploration of dementia special care units in the United States conducted by the National Institute on Aging (NIA) that noted a need for a tool to look specifically at the care environment experienced by people with dementia. An assembled workgroup identified the goals of the physical environment in long-term care, which along with a literature review formed the basis of the initial tool called TESS-2+. This tool was assessed by experts for face validity, but only modest pilot testing and no testing of its psychometric properties were carried out. Therefore, the TESS-2+ was further tested by NIA in studies of 263 special care units. Each data collection consisted of a 30 to 45 minute daytime walkthrough inspection and the analysis resulted in a few revisions and a new tool, the TESS-NH. As well as undergoing scale development, this new tool was tested for inter-rater reliability, test-retest reliability and validity: Inter-rater reliability was found to be high with inter-rater Kappa statistics for around a third of the items being above 0.60. Whilst for items included in SCUESQS, test-retest reliability was found to be 0.88. It should, however, be noted that the authors were cautious about these findings as very few raters were used in the studies. Concurrent validity was also found to be high with TESS-NH scores being strongly correlated with both light meter levels ($r=.29-.38$) and the scores of the Professional Environmental Assessment Protocol (PEAP)²⁰ ($r=.52, <.01$).

Observable Indicators of Nursing Home Quality Instrument (OIQ)

The Observable Indicators of Nursing Home Quality Instrument (OIQ) comprises of 30 items²¹ and is designed to measure the quality of care within nursing homes. The IQI instrument was developed for use by health care professionals, nursing homes, social care staff, regulators and researchers and, unlike other tools, explicitly suggests that it may be

²⁰ The Professional Environmental Assessment Protocol (PEAP) standardized method for the expert evaluation of special care units in nursing homes for people with dementia. The PEAP provides a global assessment of the quality of dementia care environments

²¹ This is correct for version 7. Earlier version consisted of more items.

used by friends and family of a resident or by the resident him/herself. It is a highly structured tool and does not require training prior to use, although the user guide provides further instructions on how to rate certain items and how to carry out the observation more generally.

OIQ contains thirty items which cover both environment and staff behaviour. The items can be used to calculate five subscales: communication, care delivery, grooming, odour and environment²², and two overall scales which assess process and structural aspects of quality using Donabedian's (1980) terminology. Each item is measured on a five-point Likert scale where 5 indicates the highest quality and 1 indicates the lowest quality. Whilst there are specific versions of the OIQ for the different types of observer (e.g. researcher, staff, family), the observational items are consistent across the versions. The researcher version, however, differs in the detail around creating sub-scales to judge quality in different domains whilst the consumer version concentrates upon an overall score of quality. A sample of items from the researcher version of OIQ is shown in box 16. In common with the TESS-NH, it does not measure the quality of care experienced by an individual but rather the quality of care at the level of the home.

Box 16: Sample items from OIQ

Did staff appear caring (compassionate, warm, kind)?

- Most did not
- A few did
- Some did
- Many did
- Most did

Did staff help residents with food or fluids?

- Rarely seen
- Occasionally
- Sometimes
- Often
- Very often
-

²² The environment scale comprises three separate scales: basic, access and homelike.

Were residents rooms, hallways and common areas clean?

- Dirty
- Somewhat dirty
- More or less clean
- Clean
- Very clean

Assessment of quality takes place over a 20 to 30 minute observation period where the observer walks through the facility and, like other observational tools, it is suggested that observation is done around a mealtime and between the hours of 10.00 and 16.00.

The QIQ tool was developed via a series of studies which sought to understand and measure the various dimensions of nursing home quality and, although developed in the United States, has been used in both Canada and Iceland (Rantz et al., 2002). Rantz et al. (2006) describes the final stage of field testing and refinement of the OIQ through a study of 407 nursing homes in three states (Missouri, Minnesota and Wisconsin) each of which was visited a least once by two observers drawn from a large pool of observers, consisting of registered nurses, 'consumers' and retired regulators. Following an exploratory factor analyses and a confirmatory factor analysis, the original 47 items of the QIQ were reduced to 30 items. The study found that this 30-item tool had both acceptable inter-rater reliability (0.76) and acceptable test-retest reliability (0.77), as tested by weighted Kappa coefficients, as well as a very strong internal consistency (0.91). The data from this study was also used to create guidelines about how to interpret the overall scores created by the tool, allowing users to easily rate a home as a good quality facility, a typical facility or a facility with quality problems.

Adaptations of OIQ

As noted above, the QIQ instrument was designed to measure quality of care in nursing homes but has been adapted for use in both residential care and assisted living facilities.

Aud et al. (2004) outlines the development of the residential care version of OIQ which began with an informal test of the nursing home version in a residential care setting. From this initial test, a residential version was created that contained 34 items and retained the five-level response format of the nursing home version. As well as reducing the number of items²³, several items were reworded and one new item was added. At this stage experts were asked to review the tool for content validity by rating each item and each subscale out

²³ Please note that when the residential version was created the nursing home version had 47 items

of four for relevance. Inter-rater and test-retest reliability were also tested and found to be acceptable (.73 for inter-rater reliability and .94 for test-retest reliability). In the light of this development work, one item about residents being out of their room was removed and two items on plants and pets were combined, resulting in a final tool of 32 items. A final focus group with residential care administrators uncovered the need to add items around food choice and access to telephone and email.

Aud et al. (2007) describes the development of an assisted living version of OIQ instrument. As in the development of the residential care version of OIQ, a panel of experts was assembled to assess the relevance of the items to an assisted living environment and, whilst all the items were rated as relevant, items were identified and added to the tool for testing. The data from this study was collected from 216 assisted living facilities in Missouri and Wisconsin and observers, who were registered nurses, made their rating on a 20-30 minute walk through. An exploratory factor analysis of the data collected resulted in a 34-item instrument with a six-factor structure (homelike, caring, access and choice and outdoor spaces). Further testing noted that this instrument had acceptable inter-rater and test-retest reliability, and strong internal consistency for the whole instrument and the individual scales (with the exception of outdoor spaces). Unfortunately, the paper does not provide the statistical results to support these conclusions. Finally guidelines on interpreting scores were created and were based upon those created for the original tool.

The Adult Social Care Outcomes Toolkit – Care Homes version (ASCOT-CH3)

The Adult Social Care Outcomes Toolkit (ASCOT) is a set of tools for measuring social care outcomes, or what is defined as social care-related quality of life (SCRQoL), across the full range of social care user groups and care settings. The collection of tools includes a self-completion and interview version of ASCOT, and also a tool designed to measure SCRQoL in care homes (ASCOT-CH3) for use with the residents of care homes who may struggle to share their experiences in either of the other formats. Unlike some of the other care homes tools mentioned in this report, ASCOT-CH3 measures outcomes for the individual resident and not the home as a whole.²⁴

The care home version of ASCOT primarily uses structured observation but is supplemented by interviews with residents, staff and family where possible. The ASCOT-CH3 is not as structured as other tools reviewed here and, although there is extensive guidance about

²⁴ The ASCOT team are in the early stages of developing a shortened version of ASCOT that measures outcomes at the level of the home rather than the individual. It is also worth noting that with thoughtful sampling and care, individual ASCOT scores could be aggregated to reflect a score for a care home.

what to look for when observing, it is advised that training by the ASCOT team is undertaken before embarking on observations within care home.

Like other tools in the ASCOT, the care home version of ASCOT is designed for use by researchers, staff and commissioners of services, and can be used to observe five residents in a two hour period. It is advised that the observational period includes a mealtime and is restricted to public areas. The focus of the observation is on resident behaviour and activities, staff and social interactions, the mood of the resident and the environment.

ASCOT focuses on eight domains of social care related quality of life (SCRQoL): control over daily life, personal cleanliness and comfort, food and drink, personal safety, social participation and involvement, occupation, accommodation cleanliness, and comfort and dignity. During the observational period, observers structure their notes under the following headings: engagement in meaningful activities, support quality and user experience, opportunity and support for choice and autonomy, physical wellbeing, emotional wellbeing, and environment. At the end of the observational period, the observational material and any other supplementary evidence (e.g. interviews with residents, family and staff) are combined in accordance with the ASCOT guidance to create a score for each domain. Each domain is measured on a three-point scale of no needs, some needs, high needs, and a sample scoring schema can be found below.

Box 17: Sample ASCOT ratings guidance

1. Control over daily life - presence of services

Key indicators for expected needs in the presence of services		Please tick one
<p>No needs The individual has no or the type of temporary trivial needs that would be expected in this area of life of someone with no impairments.</p>	<p>If the person is more able, s/he may be in control of the things that happen in their lives. You would expect to see them make regular choices about everyday things, come and go as much as they like and generally seem content with life. They would not constantly ask to do other things, like go out or do an activity. It would be very rare to see this in someone who is in a care home, but you do occasionally find people with very mild ID living with people with more severe disabilities when they could actually be living quite independently.</p> <p>For the less able majority of residents, you would see staff assisting the person to make choices using adapted communication, special equipment etc; offering choices in a facilitative way, respecting those choices; involving residents in decisions around the home and respecting those decisions.</p>	
<p>Some needs Some needs are distinguished from no needs by being sufficiently important or frequent to affect an individual's quality of life</p>	<p>Resident may be offered choices by staff for certain activities, but many things are decided by the service agenda. A more able person may be prevented from doing things because staff are risk-avoidant. Those who are more able are supported safely to do some activities and access some areas of the home, but not a full range because of the perceived risk: e.g. support provided to go into the community but not to help in the kitchen.</p> <p>Rather than assisting the person, the service might be creating a need. For example, a resident might not be permitted to make choices, whereas without services s/he would be making many of his/her own decisions. For example, the person cannot go out to the shops without staff support because service is too far from shops to walk and staff can't take them right now.</p>	

<p>High needs High needs are distinguished from low-level needs by having mental or physical health implications if they are not met over a period of time. This may be because of severity or number.</p>	<p>Person is simply processed with no regard for their sense of control. When they do try to make choices, these choices are not noticed or not respected. Person is completely disempowered by staff or by lack of support for choice making.</p>	
<p>EVIDENCE:</p>		

Unlike other tools, not only does the ASCOT attempt to measure a person’s current quality of life, but it also asks observers to make a judgement about the quality of a resident’s life if services were not helping that individual. Therefore, the ratings for each individual can be used to create three ASCOT scores: *current SCRQoL* (the resident’s observed quality of life), *expected SCRQoL* (the resident’s hypothetical quality of life if services were removed), *SCRQoL gain* (the difference between current and expected SCRQoL, and an indication of the impact the care home is having on the residents SCRQoL). To generate the final SCRQoL score, the scores for each domain are weighted using values from a study of service user and general population preferences (Netten et al., 2012).

Netten et al. (2010) describes the development and initial testing of the care home version of ASCOT. A total of 173 care homes (83 for older people and 90 for adults with learning disabilities) were recruited to the study and 5 residents were observed over a period of at least two hours in each. In addition to ASCOT-CH3, a range of data was collected during the research including information of ADLs, scores from the Minimum Data Set Cognitive Performance Scale (MDS CPS) and, importantly, data from a range of observational tools, such as the engagement in meaningful activities and relationship (EMACR) (Mansell et al., 2005)²⁵, the Active Support Measure (ASM) (Mansell et al., 2005)²⁶, the OIQ-NS and pain measurement tools. The data collected in the study showed that, overall, the ASCOT measure had good internal reliability (0.71 for current SCRQoL and 0.95 for expected SCRQoL). In order to establish concurrent validity, CH3 data was compared with data collected using other observational tools and a very strong correlation between CH3 and ASM data, both overall and for each domain, was found, but given the links between these two toolkits, this was not surprising. Correlations were also found between overall and domain specific data generated by the CH3 and OIQ-NS. Inter-rater reliability was tested using data on 113 residents and Kappa coefficients were found to be 0.47 for current

²⁵ EMACR is not included in this review of tools because it looks at only some aspects of quality of life namely activities and social life.

²⁶ ASM is not included in this review of tools as it concentrates upon the support provided by staff.

SCRQoL and 0.57 for expected SCRQoL. This was not as high as the authors would have liked and further work is planned to explore these issues further.

Section summary

We identified five tools that can be used to assess the quality of life of people using observational methods. As summarised in **Table 5**, a key characteristic that unites these tools is that they are all designed to be used in residential or institutional settings. There has been some work looking at the use of the DCM in day hospitals but the authors note this is more difficult and gives rise to quite significant amounts of missing data when people are not found in the defined observational area (Brooker, 2005; Woolley et al., 2008). There is some developmental work being done regarding observation in the community but it is still at an early stage and the tools focus predominately on aspects of quality of life, such as social interaction (Bigby and Wiesel, 2011). In terms of the suitability of observational methods for the ASCS, this implies that such an approach would be of limited use with people who cannot participate in a survey but live in a care home. It may, however, be possible to explore the feasibility of developing observational methods for people living in their own homes drawing on the experience of Bigby & Wiesel (2011).

Table 7: Summary of the characteristics of the observational tools, in terms of requirements around usage and testing of validity, reliability and sensitivity

Instrument name	Observer requirements	Structure of observations	Focus of observations	Details of testing
Dementia care mapping (DCM)	Practitioner or researcher Extensive training required	5 residents, for 6hrs over 2 days Observe bulk of waking day	Resident mood and behaviours Positive and negative staff interaction with resident	Inter-rater reliability Validation against other proxy respondent scales Construct validation No sensitivity testing
Short observational framework for inspectors (SOFI)	Inspectors Training required	5 residents, for 2hrs Observe lunchtime period	Resident mood and behaviours Staff interaction with resident	No formal testing, as far as we could ascertain
Therapeutic Environment Screening Survey for Nursing Homes (TESS-NH)	No details about who should be the observer No training required, follow detailed instructions with tool and in manual	15-45 minutes Observe mealtime, weekday	Environment	Face and concurrent validity, assessing against other measures of physical characteristics Inter-rater and test-retest validity Scale development No sensitivity testing
Observable Indicators of Nursing Home Quality Instrument (OIQ)	Anyone can be the observer, including resident No training required, follow detailed instructions with tool and in manual	20-30 minutes Observe over a mealtime and between hours of 10am to 4pm	Environment Staff behaviour (?toward residents)	Inter-rater and test-retest reliability Scale development No validity or sensitivity testing
Adult Social Care Outcomes Toolkit (ASCOT)	Researcher, practitioner or commissioner Training required	5 residents, for 2hrs Observe mealtime Observe public areas only	Resident behaviour/activity/mood Staff interaction Environment	Inter-rater reliability Validation against other observational tools capturing domains of ASCOT No sensitivity testing

Another key characteristic of all the observational tools reviewed here is that they are primarily designed for use with people with dementia (see **Table 5**). Whilst the DCM has been adapted for use with other groups (e.g. adults with learning disabilities and older adults with physical impairments), the authors of these studies report the need for caution when interpreting behaviours as they may have different purposes and meaning for different client groups (Persaud and Jaycock, 2001; Woolley et al., 2008). This implies that observational tools need to be designed for specific groups or that users of the tools need training or instruction on how to interpret the behaviours of different types of people. We found two tools designed for use in care homes across multiple client groups: the ASCOT, which is designed for use in care home for older people and learning disabilities, and the SOFI (the Short Observational Framework for Inspectors, an adaptation of DCM) which is also designed for use in care homes for older people with dementia and learning disabilities. Both of these tools require observers to attend training.

The tools summarised in **Table 7** have different foci and place different requirements on the observers. This is partly reflected by variations in the length, focus and detail of the observations. The TESS-NH and the OIQ are both short observational tools and generate ratings of the quality of the home. They both have a greater focus on environmental characteristics and do not examine resident's behaviours. In contrast, the ASCOT and the DCM/SOFI tools generate measures of the quality of life of individual residents and, as a result, require longer periods of observation (especially with the DCM/SOFI) and focus more on the behaviours of both residents and staff. Moreover, DCM/SOFI and ASCOT either require or advise a training course before their tools are used in the field. The lack of need for training for TESS-NH and OIQ may be explained by the fact that they are both highly structured, almost questionnaire like tools, whilst ASCOT and DCM are semi-structured tools that require the observer to interpret their observations into ratings on quite extensive scales.

It is clear that any tool used for the ASCS would need to be short and require no, or at least minimal, training. Since observation is a much more expensive activity than sending out questionnaires, the most appropriate measures for the ASCS would be the shortest and least demanding tools that focus on the care environment. These, however, would only be applicable for generating ratings in care homes or other institutional settings and would not be suitable for people living in their own homes. There is also limited evidence about the relationship between structural aspects of quality, such as the care environment, and the quality of life of care home residents (Spilsbury et al., 2011). Tools that focus on the behaviours of the resident and staff seem more appropriate and would be more easily adapted for use in people's own homes. However, only a small number of people can be observed at any one time and it seems that even where shorter, 2 hour observational periods are possible some training is required to ensure reliability between raters.

Despite the requirements around training to ensure reliability of ratings of behaviours, both ASCOT and DCM had low inter-rater reliability, particularly when compared to the OIQ and TESS-NH. Whether this reflects the focus of the ASCOT and DCM on behaviours or the less structured nature of the observational schedule is not clear, but it seems likely that both are contributory factors. Since neither the TESS-NH or the OIQ compare the scores with other quality of life indicators it is not clear how well either of these measures capture the contribution of the care home to their resident's quality of life. Indeed, and as we have already suggested, because the TESS-NH and OIQ focus on environmental characteristics there is reason to believe that they are unlikely to be good measures of quality of life. In contrast both DCM and ASCOT have demonstrable associations with other measures of resident quality of life, or aspects of quality of life.

Discussion and recommendations for the Adult Social Care Survey

This review has identified a variety of reasons why people may not respond to postal surveys and the potential methods for improving response rates. Some of these are summarised in Figure 1 and, in the context of the ASCS, are of varying importance. Given the survey population of the ASCS, disability is likely to be a key factor influencing the inability to respond. Whilst the studies in the literature review suggest that disability has more of an impact on survey response for some groups than others, there is little evidence from the ASCS directly. Data about the impairments of respondents (and nonrespondents) in the ASCS is poor, so researchers have not been able to investigate in any real detail the types of people who do not respond to the ASCS, except within very broad categories (Malley and Fernandez, 2012). For example, the category of 'mental health' includes not only people with depression and schizophrenia, but may also include people with dementia. These broad categories are not particularly useful in an analysis of nonresponse or as a way of understanding how response may be improved.

Despite this, it seems likely that service users with cognitive impairments are underrepresented in the ASCS. Not only are they more likely to be excluded from the sampling frame in the ASCS survey on the grounds of their mental capacity, but there is a strong possibility of interception by gatekeepers. Unfortunately, research does not tend to directly address, at least with regard to postal surveys, ways of improving participation among people with cognitive impairments. In the case of dementia, there tends to be an assumption that people with dementia cannot complete a structured questionnaire and that alternative methods of gathering data about their outcomes should be sought.

The most common alternative method for eliciting outcomes from people thought not to be able to take part in a postal survey is the use of a proxy respondent to report on behalf of the service users, although another possible method is to adopt an observational approach to gathering outcome data. Of course, neither approach really provides the service users'

views of either their quality of life or the quality of the services they receive. Instead these methods reflect the views of the assessor. Whilst using a proxy is clearly cheaper than using an observational approach there are challenges around identifying when it is appropriate to use a proxy respondent to talk on behalf of a service user, who is best placed to be the proxy and also engaging the proxy in the survey. Nonetheless, people with dementia — and, in particular, people with moderate dementia — may have capacity to consent to and take part in research directly, although they may not be able to answer survey questions without support. Developmental work on the ASCS survey suggested that using a semi-structured approach when interviewing may allow some people with dementia to express their views directly (Malley et al., 2010).

Evidence from the analysis of the 2010-11 ASCS suggests that PWID are not actually underrepresented in the survey compared to other client groups, but that there are still issues around how this group responds to the ASCS (Malley and Fernandez, 2012). Firstly, it is not clear whether all subgroups within the very broad grouping of PWID are equally likely to respond. Secondly, what is evident from the initial development of the Easy Read version of the ASCS is that most PWID needed help to answer the questions (Malley et al., 2010). This finding is also borne out by analysis of the 2010-11 ASCS (Apps and Malley, 2012, in prep; Malley et al., 2012, in prep) where 58% of the respondents to the community LD questionnaire reported that somebody ‘translated the questionnaire for me’. This compares with only 7% of the respondents to the standard community questionnaire and 10% of the respondents to the care homes version of the questionnaire. This suggests that further development of the Easy Read version of the ASCS is needed. In particular, to create a version that requires less ‘translating’ by a third party and enables PWID to understand and respond to the survey on their own or with more minimal help.

There is very little direct evidence around how sensory impairments impact upon response to postal surveys, although there is some evidence that those with visual impairments may have a lower response rate than those without visual impairments. This review though found no research that specifically addressed how to improve response rates to postal surveys from the visually impaired beyond using alternative formats or methods, or offering support.

Evidence around service users with physical impairments was even less clear. There are some reports of people with physical impairments struggling to respond to postal surveys and it is suggested that alternative methods, such as face-to-face interviews, may enable greater levels of participation in a survey such as the ASCS. We were, however, unable to find any quantitative evidence to support these suggestions.

Nevertheless, it is clear that many people do need support to participate in a postal survey. This was found in analysis of the User Experience Surveys (the precursor of ASCS) as well as the analysis of data from the first ASCS (Apps and Malley, 2012, in prep; Francis and Netten,

2004; Malley et al., 2012, in prep) and the development work on the ASCS (Malley et al., 2010). Evidence from both the development and analysis of the ASCS indicates that the type of support received is variable (from answering questions on behalf of somebody through to just posting the questionnaire back on someone else's behalf) and that the type of support also varies with regard to the characteristics of the respondent (Apps and Malley, 2012, in prep; Malley et al., 2012, in prep). Beyond this though we have very little understanding of how support, in its many guises, affects response rates, but, given that development work has seen that some people cannot participate in a postal survey without support and help, it is fairly safe to assume that help and support does improve the response rate to surveys. Of course this draws attention to the fact that the availability of support is unevenly distributed and raises concerns about the participation of people who have no support other than care workers. The analysis of the ASCS (Apps and Malley, 2012, in prep; Malley et al., 2012, in prep) and other user survey research (Francis and Netten, 2004) clearly shows that responses to the survey do alter depending on who is providing the support to complete it. Given these concerns and the overall lack of evidence of the impact of support, better evidence is needed on how support works and its effects on response rates and responses to a survey.

There is evidence that minority ethnic groups do have lower response rates compared to white respondents (Malley and Fernandez, 2012). However, while it could be argued that this may be down to language difficulties, ethnicity is not actually a good indicator of ability to understand the English language. Therefore, other reasons for lower response rates may need to be considered. There is clear evidence, for example, that generic approaches to improving response rate have been effective with some minority ethnic groups. Even so, making appropriately translated versions of surveys available is suggested as good practice for enabling the participation of those who struggle with English language, even though there is very little evidence of the effectiveness of providing translated survey questionnaire and material on response rates.

Many of the methods used to improve response rates outlined in this report are generic and are not targeted at specific hard to reach groups. Incentives have been shown to improve response rates if they are pre-paid or non-conditional. Moreover, they have been shown, as noted above, to be successful in helping improve the response rates of some hard-to-reach groups. It is unclear how an incentive system would work in a survey such as the ASCS, particularly as each CASSR carries out the survey. There are also issues around incentives being a subtle form of coercion, which are magnified for those who take part in the ASCS since the population is more vulnerable than the general population. Following up nonrespondents has also been shown to work as a generic approach to improving response rates and, indeed, there is clear evidence that CASSRs who do not follow up nonrespondents have lower response rates than those that have the one follow-up allowed by the research ethics committee as the maximum number of follow-ups (Malley and Fernandez, 2012). It

is, though, a practice that is poorly observed amongst CASSRs in the ASCS despite the fact that it is a relatively inexpensive way of ensuring better response rates and minimising the danger of response bias.

The ONS review of the ASCS (Williams and Betts, 2010) and other research (Dillman et al., 2009; Fowler Jr., 2009) clearly shows that offering alternative methods of survey administration — in particular, face-to-face or telephone interviews — can improve response rates from particular groups, including those with visual or physical impairments, and some people with dementia and, as a result, may be a better way to approach data collection with PWID. Such alternative methods may be particularly useful where potential respondents do not have informal carers or family members who can help them complete the survey. However, this approach is not without its drawbacks. Both types of interview are more expensive and time consuming than other modes of survey administration and, in light of resource constraints, are not suitable as the main method of data collection in the ASCS. Moreover, even if interviews are only used to follow up nonrespondents, there are issues around mode effects. Unfortunately, what effect different modes of survey administration would have on the responses given is currently unclear and further research into mode effects in the ASCS would be required before the use of interviews could be expanded.

The findings of this review clearly suggest that response rates are not solely due to the characteristics of service users but are also affected by the characteristics of organisations carrying out the survey, in this case the CASSRs. The review showed that trust in the organisation carrying out the survey was one of the key issues affecting whether somebody felt encouraged to participate in the survey. Most of the evidence and recommendations comes from work carried out in a commercial context and is thus of limited applicability to the ASCS. Despite this shortfall, developing ways of improving trust between local authorities and service users in the context of the ASC survey may be an area that it is worth exploring in order to improve response rates.

There are a number of recommendations for improving nonresponse and minimising response bias that draw on the work reviewed in this report.

1. Improve the data held by CASSRs on disability of clients.

This would enable research to explore the response rates for people with different levels of cognitive impairment (e.g. mild vs. moderate vs. severe) and to investigate the impact of more specific conditions on survey response, for example sensory impairments (particularly visual) and learning disabilities. If these data were sufficiently comprehensive and detailed it would allow for some distinction between types of learning disability. This would give a better picture and improve our understanding of the degree of underrepresentation within the very broad client group categories. The

collection of these data would also allow CASSRs to better target alternative versions of the survey and to provide support that was a better match to individual service users' needs.

2. The focus of future ASCS development should be on methods for gathering data from people with cognitive impairments

Given what we know about the characteristics of nonrespondents in the ASCS so far (e.g. that high numbers do have cognitive impairment), the focus of future development should concentrate upon identifying and developing appropriate methods of data collection. One potential method identified in this report was the use of proxies to capture some information on the quality of care and outcomes for people who cannot participate in the survey directly, although it is important to recognise that this would not enable one to capture the views of service users, but give another perspective upon a service user's quality of life. We would recommend that any work developing a proxy version should also explore the point at which a person, for example, with dementia, would find it impossible to complete a questionnaire and thus require a proxy to represent them. This would help ensure proxy versions are distributed in a similar way across geographical areas and minimise variability.

3. Revise the Easy Read version of the ASCS

The Easy Read version of the ASCS was created during the initial development of the ASCS but has not been fully tested and validated. Improvements could be made to the Easy Read questionnaire to ensure validity of responses for PWID. As the Easy Read version of the ASCS was initially developed to maintain compatibility with standard version, there are some long response options and, as a consequence, this format and some questions may be difficult for PWID. Whilst acknowledging the need to retain consistency in the meaning of questions, it has been suggested that the Easy Read version should be revised so that it can be completed with a minimum of support, even if this means many fewer response options and some loss of compatibility with standard. Redevelopment of the Easy Read version should also include a few additional questions which, given the very positive question responses found in the Easy Read ASCS so far, would test for acquiescence (Malley et al., 2012, in prep).

4. Develop a strategy to provide support to service users to complete the ASCS

In addition to developing alternative methods for gathering outcomes data (such as proxy versions, Easy Read version and observational approaches), work on developing the ASCS for people with cognitive impairments should consider how to develop strategies to provide support for people who can answer the questionnaire but require some help to ensure consistency of application across areas, and to ensure the participation of people who do not have family or friends to help them. Research that is

currently being conducted into the effect on responses of having support to answer the ASCS postal questionnaire will make a useful contribution to our very limited understanding of the effect of having help has on survey response (Apps and Malley, 2012, in prep). Further research into the effect of availability of help on response rates would also be extremely valuable.

5. Improve the quality of the translated versions of the ASCS

Improvements to the translated versions of the ASCS are recommended to follow best practice of forward and back-translation. Moreover, these translated versions should be favoured over interpreters to ensure consistency of translation. It is also suggested that CASSRs do more to target translated version of the ASCS, perhaps by improving records of service users for whom English is not their first language.

6. The introduction of stricter guidelines around following up nonrespondents

Given the very clear evidence around the effectiveness of following up nonrespondents, we suggest that stricter guidelines are drawn up to reflect this evidence and methods are put in place to ensure that the guidelines are adhered to.

7. Consider greater use of alternative modes of data collection, with the caveat that extensive use be introduced only following a randomised trial to investigate the effect of mode on responses.

We hesitate to fully recommend the increased use of alternative modes of data collection because of the additional time and expense they entail. However, it is clear that interviews provide better response rates and are easier for certain groups of people with disabilities. We would also recommend caution is exercised before extending the use of alternative modes since different modes are known to have an influence over responses. A randomised trial would be necessary to investigate the effect of mode on responses before alternative modes are used more widely.

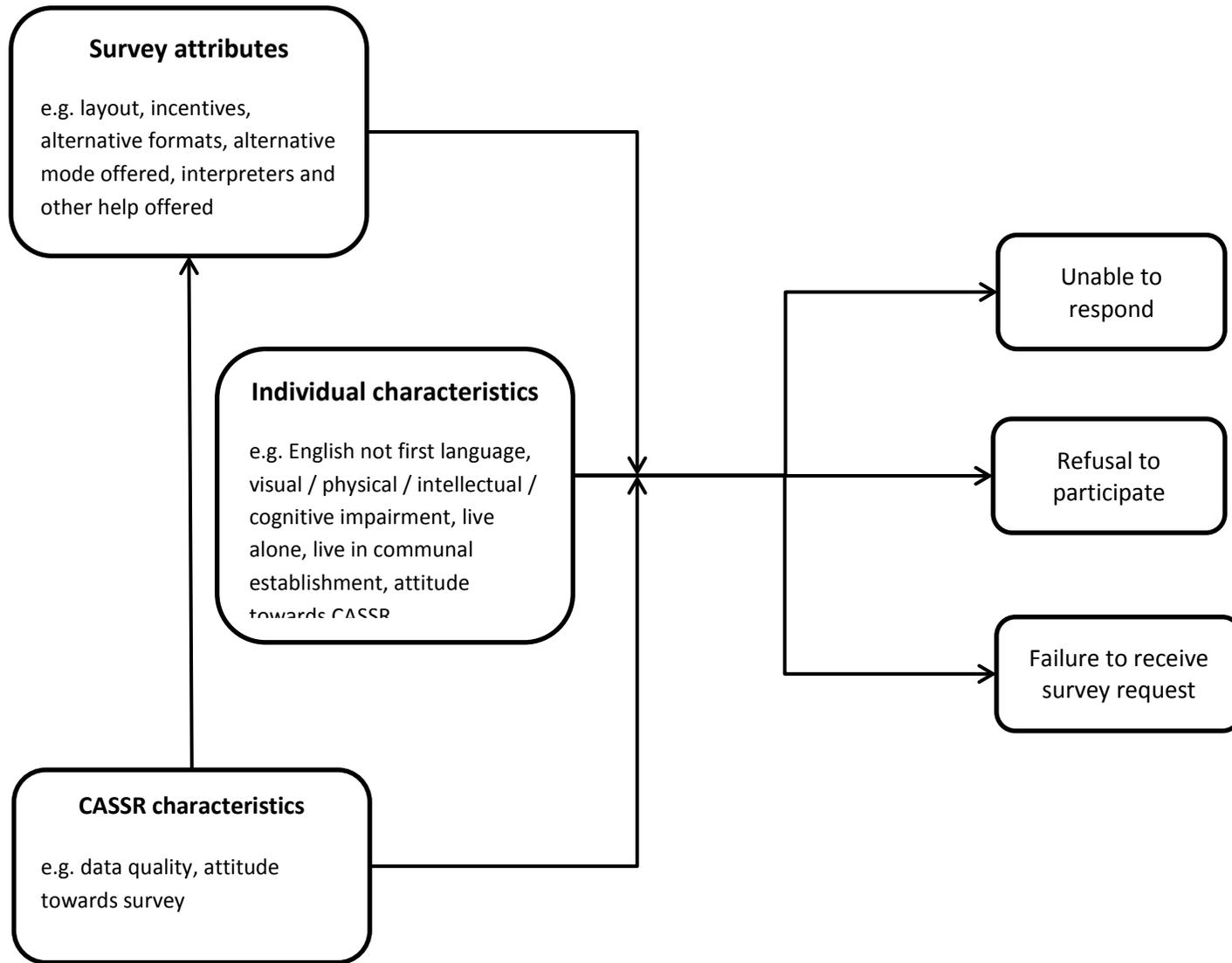
8. CASSRs should use their position to improve response rates

As CASSRs have considerable control over the administration of the ASCS, they are in a good position to take measures to improve response rates. Two key issues for CASSRs to focus on are improving trust with service users around their use of ASCS data and confidentiality, and on improving the quality of their data on service users, so they can better target alternative modes, questionnaire versions and methods.

9. Consider the greater use of incentives, with the caveat that extensive use be introduced only following a randomised trial to investigate the effect on responses and response rates.

Whilst the use of incentives, particularly pre-paid options, has consistently been shown to be effective in improving response rates, there is no proof that such measures encourage the completion of the ASCS. As there is also the issue of coercion we hesitate to recommend incentives and suggest that their use is expanded with caution and only following a randomised trial to ensure there are no adverse effects on responses and response rates.

Figure 1 Factors affecting nonresponse to postal surveys



Appendix 1: Methods for the systematic review of proxy tools

The primary aim of this systematic review was to gather information about proxy tools and their use in research. We carried out a search with the aid of a library and search specialist who searched two databases (Psychinfo, Social Science Citation Index) using key terms *proxies, proxy, develop*, tool, survey*. We focussed specifically on articles about the development of proxy tools for adults in health and social care settings that were published in English during the last 10 years and identified 51 papers.

A quick review of the abstracts revealed that almost all of these 51 articles were not relevant to the aim of our report and confirmed the view of the search specialist that articles about methodological issues were difficult to identify in searches. Therefore, the search for tools and information about their development and testing was carried out in a much more ad hoc way. A review carried out by the Quality and Outcomes Research Unit (Beadle-Brown et al., 2012) had identified a number of proxy tools, and coupled with others found via conversations with experienced researchers provided us with an initial list to search for additional information. We carried out searches using the names of individual tools in the Social Science Citation Index and Google Scholar. We scanned all returned papers for both information and references on the development of the tools. General web searches were also used to identify any training manuals or support websites that existed, and, in several cases, those who had developed the tools were contacted and asked to provide access to the tools, support manuals and other material.

Appendix 2: Methods for the systematic review of observational tools

The primary aim of this systematic review was, as for the proxy review, to gather information about observational tools and their use in research. Following our experience of search for proxy tools a more ad-hoc approach was adopted. We consulted experts on observational methods about relevant tools and searched directly for papers which outlined their development and usage. References to other tools made in these papers were also followed up and, if considered to be relevant, included in this review. We found and rejected a few tools developed for commercial purposes which did not have extensive documentation or published work on either its development or testing.

As with the proxy search, where the tools themselves were not easily accessible in the public domain we contacted those who had developed them for copies.

References

- Aldridge J (2007) Picture this: the use of participatory photographic research methods with people with learning disabilities, *Disability & Society*, 22, Jan-17.
- Allison T, Ahmad T, Brammah T, Symmons D & Urwin M (2003) Can findings from postal questionnaires be combined with interview results to improve the response rate among ethnic minority populations?, *Ethnicity & Health*, 8, 63-69.
- Andrews J (2005) Wheeling uphill? Reflections of practical and methodological difficulties encountered in researching the experiences of disabled volunteers, *Disability & Society*, 20, 201-212.
- Apps J & Malley J (2012, in prep) The effect of having help to answer the questionnaire on survey respondent's views of their quality of life: evidence from the Adult Social Care Survey.
- Atkinson R & Flint J (2001) Accessing hidden and hard-to-reach populations: snowball research strategies, *Social Research Update*, 33.
- Aud M, Rantz M, Zwygart-Stauffascher M & Flesner M (2007) Measuring Quality of Care in Assisted Living: A new tool for providers, consumers and researchers, *Journal of Nursing Care Quality*, 22, 4-7.
- Aud M, Rantz M, Zwygart-Stauffer M & Manion P (2004) Developing a Residential Care Version of the Observable Indicators of Nursing Home Care Quality Instrument, *Journal of Nursing Care Quality*, 19, 48-57.
- Auster J & Janda M (2009) Recruiting older adults to health research studies: A systematic review, *Australasian journal on ageing*, 28, 149-151.
- Beadle-Brown J, Ryan S, Windle K, Holder J, Turnpenny A, Smith N, Richardson L & Whelton B (2012) Engagement of people with long term conditions in health and social care research: Barriers and facilitators to capturing the views of seldom-heard populations. Canterbury: Quality and Outcomes of Person-Centred Care policy Research Unit.
- Bigby C & Wiesel I (2011) Encounter as a dimension of social inclusion for people with intellectual disability: beyond and between community presence and participation, *Journal of Intellectual and Developmental Disability*, December 2011, 1-5., December 2011, 1-5.
- Black B
(2012) http://www.hopkinsmedicine.org/psychiatry/specialty_areas/geriatric_psychiatry/research/adrql.html [Online]. [Accessed 12/03/2012].
- Bootsma-van der Wiel A, van Exel E, de Craen A J M, Gussekloo J, Lagaay A M, Knook D L & Westendorp R G J (2002) A high response is not essential to prevent selection bias: Results from the Leiden 85-plus study, *Journal of Clinical Epidemiology*, 55, 1119-1125.
- Brewster S J (2004) Putting words in their mouths? Interviewing people with learning disabilities and little/no speech, *British Journal of Learning Disabilities*.
- Brooker D (2005) Dementia Care Mapping: A review of the research literature, *The Gerontologist*, 45, 1-18.
- Brooker D (2007) Introducing SOFI: a new tool for inspection of care homes, *Journal of Dementia Care*, July/August, 22.
- Brooker D & Surr C (2006) Dementia Care Mapping: Initial validation of DCM8 in UK field trials, *International Journal of Geriatric Psychiatry*, 21, 1018-1025.
- Brown P & Scullion L (2010) 'Doing research' with Gypsy-Travellers in England: reflections on experience and practice, *Community Development Journal*, 45, 169-185.
- Bryan S, Hardyman W, Bentham P, Buckley A & Laight A (2005) Proxy completion of EQ-5D in patients with dementia, *Quality of Life Research*, 14, 107-118.
- Buckley B, Murphy A W, Byrne M & Glynn L (2007) Selection bias resulting from the requirement for prior consent in observational research: a community cohort of people with ischaemic heart disease, *Heart*.

- Cambridge P & McCarthy M (2001) User focus groups and best value in services for people with learning disabilities., *Health and Social Care in the Community.*, 9, 476-489.
- Clark A (2007) The Potential of Observation Methods for Gerontology. In: Clark A (ed.) *Making Observations: The Potential of Observation Methods for Gerontology* London: Centre for Policy on Ageing.
- Commission for Social Care Inspection (2008) See Me, Not Just the Dementia: Understanding people's experiences of living in a care home. London: Commission for Social Care Inspection.
- Coucill W, Bryan S, Bentham P, Buckley A & Laight A (2001) EQ-5D in patients with dementia: An investigation of inter-rater agreement, *Medical Care*, 39, 760-771.
- Crespo M, Quiros M B d, Gomez M & Hornillos C (2012) Quality of life of Nursing Homes Residents With Dementia: a comparison of perspectives of residents, family and staff, *The Gerontologist*, 52, 56-65.
- Cummins R (1991) The Comprehensive Quality of Life Scale - Intellectual Disability: An instrument under development, *Australia and New Zealand Journal of Development Disabilities*, 17, 259-264.
- Cummins R (1997) Comprehensive Quality of Life Scale - Intellectual/Cognitive Disability 5th Edition: manual. School of Psychology, Deakin University.
- Department of Health (2001) Personal Social Services User Experience Survey 2000-2001. Bulletin 2001/30. London: The Stationery Office.
- Department of Health (2002) Personal Social Services User Experience Survey 2001-2002. Bulletin 2002/28. London: The Stationery Office.
- Department of Health (2003) Personal Social Services Survey of Home Care Users in England Aged 65 or Over: 2002-03. Bulletin 2003/26. In: Department of Health (ed.). London: The Stationery Office.
- Department of Health (2004) Personal Social Services Survey of Physically Disabled and Sensory Impaired Users in England Aged 18-64: 2003-04, Bulletin 2004/23. London: The Stationery Office.
- DeVellis R (2003) *Scale Development: theory and applications*, London, Sage.
- Dillman D A, Smyth J D & Christian L M (2009) *Internet, mail and mixed-mode surveys: the tailored design method*, New York, John Wiley and sons.
- Donabedian A (1980) *Explorations in Quality Assessment and Monitoring. Volume I: The Definition of Quality and Approaches to its Assessment*, Ann Arbor, MI, Health Administration Press.
- Edvardsson D & Innes A (2010) Measuring Person-Centred Care: A critical comparative review of published tools, *The Gerontologist*, 50, 834-846.
- Edwards P, Roberts I, Clarke M, DiGiuseppi C, Wentz R, Kwan I, Cooper R, Felix L & Pratap S (2009) Methods to increase response to postal and electronic questionnaires. Cochrane Review.
- Finlay W M L & Lyons E (2002) Acquiescence in interviews with people who have mental retardation, *Mental Retardation*, 40, 14-29.
- Finlay W M L L, E. (2001) Methodological issues in interviewing and using self-report questionnaires with people with mental retardation., *Psychological Assessment*, 13, 319-335.
- Finnamore T & Lord S (2007) The Use of Dementia Care Mapping in People With Learning Disabilities and Dementia, *Journal of Intellectual Disabilities*, 11, 157-165.
- Fowler Jr. J (2009) *Survey Research Methods*, , LONDON, Sage.
- Francis J & Netten A (2004) Raising the quality of home care: a study of service users' views, *Social Policy and Administration*, 38, 290-305.
- Fraser M & Fraser A (2001) Are people with learning disabilities able to contribute to focus groups on health promotion?, *Journal of Advanced Nursing*, 33, 225-233.
- Garland J, Spalek B & Chakraborti N (2006) Hearing lost voices: issues in researching 'hidden' minority ethnic communities, *British Journal of Criminology*, 46(3), May 2006, pp.423-437.

- Gilbert T (2004) Involving people with learning disabilities in research: issues and possibilities, *Health & Social Care in the Community*, 12, 298-308.
- Gordon M S, Shevlin M, Tierney K J, Bunting B & Trimble T (2007) Correspondence between self-ratings and key-workers' ratings of depression in adults with mild learning disabilities, *British Journal of Clinical Psychology*, 46, 491-495.
- Groves R (2006) Nonresponse rates and nonresponse bias in household surveys, *Public Opinion Quarterly*, 70, 646-675.
- Groves R & Heeringa S (2006) Responsive design for household surveys: tools for actively controlling survey errors and costs, *Journal of the Royal Statistical Society A*, 169, 439-457.
- Groves R, Singer E & Corning A (2000) Leverage-saliency theory of survey participation: description and an illustration, *The Public Opinion Quarterly*, 64, 299-308.
- Hancock G A, Reynolds T, Woods B, Thornicroft G & Orrell M (2003) The needs of older people with mental health problems according to the user, the carer, and the staff, *International Journal of Geriatric Psychiatry*, 18, 803-811.
- Harkins C, Shaw R, Gillies M, Sloan H, MacIntyre K, Scoular A, Morrison C, MacKay F, Cunningham H, Docherty P, MacIntyre P & Findlay I (2010) Overcoming barriers to engaging socio-economically disadvantaged populations in CHD primary prevention: a qualitative study, *BMC Public Health*, 10, 391.
- Hellstrom I, Nolan M, Nordenfelt L & Lundh U (2007) Ethical and methodological issues in interviewing persons with dementia, *Nursing Ethics*, 14, 608-619.
- Hilari K, Byng S, Lamping D & Smith S (2003) Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Evaluation of acceptability, reliability and validity, *Stroke*, 34, 1944-1950.
- Hilari K, Owen S & Farrelly S J (2007) Proxy and self-report agreement on the Stroke and Aphasia Quality of Life Scale-39, *Journal of Neurology, Neurosurgery & Psychiatry*, 78, 1072-5.
- Hoe J, Katona C, Orrell M & Livingston G (2007) Quality of life in dementia: care recipient and caregiver perceptions of quality of life in dementia: the LASER-AD study, *International Journal of Geriatric Psychiatry*, 22, 1031-1036.
- Hoffmann T, Worrall L, Eames S & Ryan A (2010) Measuring Outcomes in People Who Have Had a Stroke and Their Carers: Can the Telephone Be Used?, *Topics in Stroke Rehabilitation*, 17, 119-127.
- Howard L, de Salis I, Tomlin Z, Thornicroft G & Donovan J. (2009). Why is recruitment to trials difficult? An investigation into recruitment difficulties in an RCT of supported employment in patients with severe mental illness. *Contemporary Clinical Trials* [Online]. Available: <http://www.mrw.interscience.wiley.com/cochrane/clcmr/articles/CMR-13987/frame.html>.
- Jackle A, Roberts C & Lynn P (2006) Telephone versus Face-to-Face Interviewing: Mode effect on data quality and likely causes. Institute for Social and Economic Research Working paper.
- Kaehne A & O'Connell C (2010) Focus groups with people with learning disabilities, *Journal of Intellectual Disabilities*, 14, 133-145.
- Kasper J, Black B, Shore A & Rabins P (2009) Evaluation of the Validity and Reliability of the Alzheimer's Disease-Related Quality of Life (ADRQL) Assessment Instrument, *Alzheimer Dis Assoc Disord*, 23, 275-284.
- Kauppi M, Sokka T & Hannonen P (2005) Survey nonresponse is associated with increased mortality in patients with rheumatoid arthritis and in a community population, *The Journal of Rheumatology*, 32, 807-810.
- Kuhn D, Ortigara A & Kasayka R (2000) Dementia Care Mapping: An innovative tool to measure person-centred care, *Alzheimer's Care Quarterly*, 1, 7-15.
- Lall R, Mistry D, Bridle C & Lamb S E (2012) Telephone interviews can be used to collect follow-up data subsequent to no response to postal questionnaires in clinical trials, *Journal of Clinical Epidemiology*, 65, 90-99.

- Llewellyn P (2009) Supporting people with intellectual disabilities to take part in focus groups: Reflections on a research project, *Disability & Society*, 24, 845-856.
- Lloyd C E, Johnson M R, Mughal S, Sturt J A, Collins G S, Roy T, Bibi R & Barnett A H (2008) Securing recruitment and obtaining informed consent in minority ethnic groups in the UK, *BMC Health Services Research*, 68.
- Lloyd V, Gatherer A & Kalsy S (2006) Conducting qualitative interview research with people with expressive language difficulties, *Qualitative Health Research*, 16, 1386-1404.
- Logsdon R, Gibbons L, McCurry S & Teri L (1999) Quality of Life in Alzheimer's Disease: patient and caregiver reports, *Journal of Mental Health and Aging*, 5, 21-32.
- Logsdon R G, Gibbons L E, McCurry S M & Teri L (2002) Assessing quality of life in older adults with cognitive impairment, *Psychosomatic Medicine*, 64, 510-519.
- Malley J, Caiels J, Fox D, McCarthy M, Smith N, Beadle-Brown J, Netten A & Towers A (2010) A Report on the Development Studies for the National Social Care User Experience Survey. Canterbury: Personal Social Services Research Unit.
- Malley J & Fernandez J-L (2012) Patterns and Effects of Nonresponse in the English Adult Social Care Survey. London: Personal Social Services Research Unit, London School of Economics.
- Malley J, Fernandez J & Netten A (2012, in prep) Generating 'adjusted' indicators from social care survey data Canterbury Quality and Outcomes of Person-Centred Care policy Research Unit.
- Manfreda K L, Bosniak M, Berzelak J, Haas I & Vehovar V (2008) Web surveys versus other survey modes - A meta-analysis comparing response rates, *International Journal of Market Research*, 50, 79-104.
- Mansell J (2011) Structured observational research in services for people with learning disabilities. *Methods Reveiw 10*. London: School for Social Care Research.
- Mansell J, Elliot T & Beadle-Brown J (2005) Active Support Measure (ASM): Handbook for observers. Canterbury: Tizard Centre.
- McKee K J, Houston D M & Barnes S (2002) Methods for Assessing Quality of Life and Well-Being in Frail Older People, *Psychology & Health*, 17, 737-751.
- McKeown J, Clarke A, Ingleton C & Repper J (2010) Actively involving people with dementia in qualitative research, *Journal of Clinical Nursing*, 19, 1935 - 1943.
- McNally S (2003) A survey of self-advocacy groups for people with learning disabilities in an English region: Part II, *Journal of Learning Disabilities*, 7, 231-250.
- Mindham J & Espie C A (2003a) Glasgow Anxiety Scale for people with an Intellectual Disability (GAS-ID): development and psychometric properties of a new measure for use with people with mild intellectual disability, *Journal of intellectual disability research : JIDR*, 47, 22-30.
- Mindham J & Espie C A (2003b) Glasgow Anxiety Scale for people with an Intellectual Disability (GAS-ID): development and psychometric properties of a new measure for use with people with mild intellectual disability, *Journal of Intellectual Disability Research*, 47, 22-30.
- Netten A, Beadle-Brown J, Trukeschitz B, Towers A-M, Welsh E, Forder J, Smith J & Alden E (2010) Measuring the Outcomes of Care Homes: Final report. *PSSRU Discussion Paper 2696/2 June 2010*. Canterbury.
- Netten A, Burge P, Malley J, Potoglou D, Towers A-M, Brazier J, Flynn T, Forder J & Wall B (2012) Outcomes of social care for adults: developing a preference-weighted measure, *Health Technol Assess*, 16.
- Nind M (2009) Conducting qualitative research with people with learning, communication and other disabilities; methodological challenges, *National Centre for Review Methods*.
- Nota L, Soresi S & Perry J (2006) Quality of life in adults with an intellectual disability: the Evaluation of Quality of Life Instrument, *Journal of Intellectual Disability Research*, 50, 371-385.

- Oliver-Africano P, Dickens S, Ahmed Z, Bouras N, Cooray S & Deb S (2010) Overcoming the barriers experienced in conducting a medication trial in adults with aggressive challenging behaviour and intellectual disabilities, *Journal of Intellectual Disability Research*, 54, 17-25.
- Ouellette-Kuntz H (1990) A Pilot Study in the Use of the Quality of Life Schedule, *Social Indicators Research*, 23, 283-298.
- Pawson N, Raghavan R, Small N, Craig S & Spencer M (2005) Social inclusion, social networks and ethnicity: The development of the social inclusion interview schedule for young people with learning disabilities, *British Journal of Learning Disabilities*, 33, 15-22.
- Perry J & Felce D (2002) Subjective and objective quality of life assessment: Responsiveness, response bias, and resident:proxy concordance, *Mental Retardation*, 40, 445-456.
- Persaud M & Jaycock S (2001) Evaluating Care Delivery: The application of Dementia Care Mapping in learning disability residential services, *Journal of Intellectual Disabilities*, 5, 345-352.
- Phelan M, Slade M & Thornicroft G (1995) The Camberwell Assessment of Need: the validity and reliability of an instrument to assess the needs of people with severe mental illness, *British Journal of Psychiatry*, 167, 589-852.
- Pickard S, Johnson J, Feeny D, Shuaib A, Carriere K & Nasser A (2004) Agreement Between Patient and Proxy Assessments of Health-Related Quality of Life After Stroke Using the EQ-5D and Health Utilities Index, *Stroke*, 35, 607-612.
- Proctor G (2001) Listening to older women with dementia: relationships, voices and power, *Disability & Society*, 16, 361-376.
- Public Works and Government Services Canada (2007) Improving Respondent Cooperation for Telephone Surveys. Ottawa: Public Works and Government Services Canada.
- Rabins P, Kasper J, Kleinman L & Black B (1999) Concepts and methods in the development of the ADRQL: an instrument for assessing health-related quality of life in persons with Alzheimer's disease, *Journal of Mental Health and Aging*, 5, 33- 48.
- Rahi J S, Manaras I, Tuomainen H & Lewando Hundt G (2004) Engaging families in health services research on childhood visual impairment: barriers to, and degree and nature of bias in, participation, *British Journal of Ophthalmology*, 88, 782-787.
- Rantz M, Jensdottir A, Hjaltadottir I, Gudmundsdottir H, Gudjonsdottir J, Brunton B & Rook M (2002) International Field Test Results of the Observable Indicators of Nursing Home Care Quality Instrument, *International Nursing Review*, 14.
- Rantz M, Zwygart-Stauffacher M, Mehr D, Petroski G, Owen S, Madsen R, Flesner M, Conn V, Bostick J, Smith R & Maas M (2006) Field Testing, Refinement and Psychometric Evaluation of a New Measure of Nursing Home Care Quality, *Journal of Nursing Measurement*, 14, 129-148.
- Rapkin B D & Schwartz C E (2004) Toward a theoretical model of quality-of-life appraisal: Implications of findings from studies of response shift, *Health Qual Life Outcomes*, 2, 14.
- Reynolds T, Thornicroft G, Abas M, Woods B, Hoe J, Leese J & Orrell M (2000) Camberwell Assessment of Need for the Elderly (CANE): Development, validity and reliability, *The British Journal of Psychiatry*, 176, 444-452.
- Rooney L K, Bhopal R, Halani L, Levy M L, Partridge M R, Netuveli G, Car J, Griffiths C, Atkinson J, Lindsay G & Sheikh A (2011) Promoting recruitment of minority ethnic groups into research: qualitative study exploring the views of South Asian people with asthma, *Journal of Public Health*.
- Rubin D B (1976) Inference and missing data, *Biometrika*, 63, 581-592.
- Rugkasa J & Canvin K (2011) Researching Mental Health in Minority Ethnic Communities: Reflections on Recruitment., *Qualitative Health Research*, 21, 132 - 143.
- Schalock R L (2004) The concept of quality of life: what we know and do not know, *Journal of Intellectual Disability Research*, 48, 203-216.

- Schmidt S, Power M, Green A, Lucas-Carrasco R, Eser E, Dragomirecka E & Fleck M (2010) Self and proxy rating of quality of life in adults with intellectual disabilities: Results from the DISQOL study, *Research in Developmental Disabilities*, 31, 1015-1026.
- Schwartz C & Rapkin B (2004) Reconsidering the psychometrics of quality of life assessment in light of response shift and appraisal, *Health and Quality of Life Outcomes*, 2, 16.
- Scott P & Edwards P (2006) Personally addressed hand-signed letters increase questionnaire response: a meta-analysis of randomised controlled trials, *BMC Health Services Research*, 6, 111.
- Selai C E, Trimble M R, Rossor M N & Harvey R J (2001) Assessing quality of life in dementia: Preliminary psychometric testing of the Quality of Life Assessment Schedule (QOLAS), *Neuropsychological Rehabilitation*, 11, 219-243.
- Seymour D G, Ball A E, Russell E M, Primrose W R, Garratt A M & Crawford J R (2001) Problems in using health survey questionnaires in older patients with physical disabilities. The reliability and validity of the SF-36 and the effect of cognitive impairment, *Journal of Evaluation in Clinical Practice*, 7, 411-418.
- Sheikh A, Halani L, Bhopal R, Netuveli G, Partridge M R, Car J, Griffiths C & Levy M. (2009). Facilitating the recruitment of minority ethnic people into research: qualitative case study of South Asians and asthma. *PLoS Medicine* [Online]. Available: <http://www.mrw.interscience.wiley.com/cochrane/clcmr/articles/CMR-15598/frame.html>.
- Simmons E & Wilmot A (2010) Incentive payments on social surveys: A literature review, *Survey Methodology Bulletin*, 53, 1-11.
- Sloane P, Brooker D, Cohen L, Douglass C, Edelman P, Fulton R, Jarrot S, Kasayka R, Kuhn D, Preisser S, C Willia s & Zimmerman S (2007) Dementia Care Mapping as a research Tool, *International Journal of Geriatric Psychiatry*, 22, 580-589.
- Sloane P, Mitchell M, Weisman G, Zimmerman S, Long Foley K, Lynn M, Calkins M, Powell Lawton M, Teresi J, Grant L, Lindeman D & Montgomery R (2002) The Therapeutic Environment Screening Survey for Nursing Homes (TESS-NH):An Observational Instrument for Assessing the Physical Environment of Institutional Settings for Persons With Dementia, *Journal of Gerontology, Psychol Sci Soc Sci*, 57, S69-S78.
- Smith S, Murray J, Banerjee S, Foley B, Cook J, Lamping D, Prince M, Harwood R, Levin E & Mann A (2005a) What constitutes health-related quality of life in dementia? Development of a conceptual framework for people with dementia and their carers, *International Journal of Geriatric Psychiatry*, 20, 889-895.
- Smith S C, Lamping D L, Banerjee S, Harwood R, Foley B, Smith P, Cook J C, Murray J, Prince M, Levin E, Mann A & Knapp M (2005b) Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology, *Health Technology Assessment*, 9, 1-+.
- Smith S C, Lamping D L, Banerjee S, Harwood R H, Foley B, Smith P, Cook J C, Murray J, Prince M, Levin E, Mann A & Knapps M (2007) Development of a new measure of health-related quality of life for people with dementia: DEMQOL, *Psychological Medicine*, 37, 737-746.
- Spilsbury K, Hewitt C, Stirk L & Bowman C (2011) The relationship between nurse staffing and quality of care in nursing homes: a systematic review, *Int J Nurs Stud*, 48, 732-50.
- Sprangers M A & Schwartz C E (1999) Integrating response shift into health-related quality of life research: a theoretical model, *Soc Sci Med*, 48, 1507-15.
- The NHS Information Centre (2006) Personal Social Services Survey of Home Care Users in England aged 65 and over, 2005-06. Leeds: The Health and Social Care Information Centre.
- The NHS Information Centre (2008) Personal Social Services Survey of Adults Receiving Community Equipment and Minor Adaptations in England, 2007-08. Leeds: The Health and Social Care Information Centre.

- The NHS Information Centre (2009) Personal Social Services Home Care Users in England aged 65 and over, 2008-09 Survey. Leeds: The Health and Social Care Information Centre.
- The NHS Information Centre (2010) Personal Social Services Survey of Adults Receiving Community Equipment and Minor Adaptations in England, 2009-10. Leeds: The Health and Social Care Information Centre.
- Tuffrey-Wijne I, Bernal J & Hollins S (2008) Doing research on people with learning disabilities, cancer and dying: ethics, possibilities and pitfalls, *British Journal of Learning Disabilities*, 36(3), September 2008, pp.185-190.
- Tuffrey-Wijne I a D, J (2007) This is my story: I've got cancer: 'The Veronica Project': An ethnographic study of the experiences of people with learning disabilities who have cancer, *British Journal of Learning Disabilities*.
- Ulivi G, Reilly J & Atkinson J M (2009) Protection or empowerment: Mental health service users' views on access and consent for non-therapeutic research, *Journal of Mental Health*, 18, 161 - 168.
- Wensing M, Mainz J, Kramme O, Jung H P & Ribacke M (1999) Effect of mailed reminders on the response rate in surveys among patients in general practice, *Journal of Clinical Epidemiology*, 52, 585-587.
- Wensing M & Schattenberg G (2005) Initial nonresponders had an increased response rate after repeated questionnaire mailings, *Journal of Clinical Epidemiology*, 58, 959-961.
- Williams B, Irvine L, McGinnis A R, McMurdo M E & Crombie I K. (2007). When "no" might not quite mean "no"; the importance of informed and meaningful non-consent: results from a survey of individuals refusing participation in a health-related research project. *BMC Health Services Research* [Online]. Available: <http://www.mrw.interscience.wiley.com/cochrane/clcmr/articles/CMR-10704/frame.html>.
- Williams L, Weinberger M, Harris L, Clark D & Biller J (1999) Development of a Stroke Specific Quality of Life Scale, *Stroke*, 30, 1362-1369.
- Williams L S, Bakas T, Brizendine E, Plue L, Tu W Z, Hendrie H & Kroenke K (2006) How valid are family proxy assessments of stroke patients' health-related quality of life?, *Stroke*, 37, 2081-2085.
- Williams S & Betts P (2010) Review of methodology for the National Social Care User Experience Survey. London: ONS Methodology Advisory Service.
- Wilson E, Pollock K & Aubeeluck A (2010) Gaining and maintaining consent when capacity can be an issue: A research study with people with huntington's disease, *Clinical Ethics*, 5, 142-147.
- Woodall A, Howard L & Morgan C (2011) Barriers to participation in mental health research: Findings from the Genetics and Psychosis (GAP) study, *International Review of Psychiatry*, 23, 31-40.
- Woolley R, Young J, Green J & Brooker D (2008) The Feasibility of Care Mapping to Improve Care for Physically Ill People in Hospital, *Age and Ageing*, 37, 390-395.
- Young A F & Chesson R A (2006) Obtaining views on health care from people with learning disabilities and severe mental health problems, *British Journal of Learning Disabilities*, 34, Nov-19.
- Zermansky A G, Alldred D P, Petty D R & Raynor D K (2007) Striving to recruit: the difficulties of conducting clinical research on elderly care home residents, *Journal of the Royal Society of Medicine*, 100, 258 - 261.