department for culture, media and sport

SPORT ENGLAND

Taking Part and Active People Surveys: an independent review

Kerry Sproston, National Centre for Social Research and Susan Purdon, Bryson Purdon Social Research.

August 2010

improving the quality of life for all

Our aim is to improve the quality of life for all through cultural and sporting activities, support the pursuit of excellence, and champion the tourism, creative and leisure industries.

Disclaimer: this review was undertaken during 2009, at a time when PSA21 was still active. The review therefore gives consideration to the measurement of targets under the previous spending review period. In taking forward the findings and recommendations of this review, the changing PSA landscape will need to be reflected.

Contents

Foreword 4
Executive Summary 5
Chapter 1: Introduction 7
Chapter 2: Options for the future 9
Chapter 3: Options considered, but ruled out 19
Chapter 4: Data on children 21
Appendix A: The statistical arguments for and against selecting multiple adults per household 24

Foreword

The review of Taking Part and Active People was jointly commissioned by the Department for Culture Media and Sport (DCMS) and Sport England. Their responses to NatCen's recommendations are below.

DCMS response

This publication of this review is timely and comes at a point when preparatory work for the next iteration of the Taking Part survey is about to begin. DCMS are committed to using resources in the most efficient way and, to this end, will use the review to inform our discussions on the future of the survey and the development of the Taking Part tender specification.

In particular, we will ask all bidders for any future Taking Part survey to provide fully costed options for: conducting interviews with all adults in each household; a survey using a reduced core sample; boosting the child sample; and providing a longitudinal element to the survey. The technical pros and cons of each of these options are also being explored.

Taking Part is a well established and highly regarded survey that provides the most cohesive view available across our culture and sport sectors. While we will work hard to generate efficiencies in carrying out the survey, the quality and robustness of the survey will remain foremost in our minds. As a National Statistic, any significant changes to the Taking Part methodology and range of outputs will be subject to user consultation.

If you would like to be notified of ongoing developments and potential changes to the Taking Part survey, please email us at: TakingPart@culture.gsi.gov.uk. We also welcome other feedback and comments at this address.

Taking Part team, DCMS

Sport England response

Sport England welcomes this review. The Active People Survey plays a crucial role in providing the evidence underpinning our strategic objective to grow and sustain participation in sport. The results from the survey inform performance frameworks linked to each of our funded national governing bodies of sport as well as providing a key sport and recreation measure, National Indicator 8 (participation in sport and active recreation) for local government. It is vital that in any measurement approach consistency is maintained to ensure that we can track trends over time.

Sport England has only recently completed a competitive procurement process that has resulted in the appointment of TNS-BMRB to carry out Active People 5 commencing mid October 2010. This has resulted in significant economies and we will continue to drive savings into the survey process and respond in flexible ways to opportunities for improvements in survey design whilst supporting our strategic priorities.

Research and Evaluation team, Sport England

Executive Summary

DCMS and Sport England commissioned NatCen to undertake an independent review of the Taking Part (TP) and Active People (AP) surveys. The aim of the review was to scope the options for collecting national and local level data on participation in culture and sport, and to consider how the two surveys might be merged or modified in order to reduce costs, while still meeting the requirements of stakeholders.

This review was primarily a desk research exercise. In addition, feedback was sought from key stakeholders across a number of government departments.

A number of options were drawn up and presented to DCMS and Sport England. Some of the options were ruled out. These options, and the reasons why they were ruled out, are covered in Chapter 3. Four potential options for the future with associated strengths and weaknesses are described in Chapter 2, and summarised below.

1. **Substitute some of the Active People sample with Taking Part cases.** This option would use the Taking Part sample to provide partial sample for each local authority (LA) for Active People, thereby reducing the size of the Active People sample. Each annual Taking Part survey covers all 354 (lower tier) local authorities in England. The latest Taking Part sample size for adults (aged 16 and over) was 14,452. Divided by 354 LAs, this would provide an average of 41 respondents per LA. This represents about 8% of the minimum necessary Active People sample (N=500) per year.

   The main advantage of this approach is that it substantially reduces the costs of the Active People fieldwork; though combining the surveys would make the generation of LA-level statistics more complex, and the additional costs needed to handle this would slightly offset the savings.

2. **Move Taking Part from a 'one adult per household' design to an 'all adults in household' design.** One relatively straightforward way of reducing the cost of the TP adult survey would be to move to a design where, instead of selecting just one adult per household, two, or even all, adults are selected. Given that the average number of adults per household is around 1.8, in order to achieve a sample of around 15,000 individuals, as few as 14,000 households would need to be selected (compared with 25,000 for a 'one adult per household' model). This would substantially reduce fieldwork costs.

3. **Have a smaller core Taking Part sample with a full interview, boosted by a sample with shorter interviews.** This would involve having a core national TP sample (say around 5,000 cases). This core sample would be given the full TP face-to-face questionnaire. An 'extra' sample of 10,000 respondents would be given a much shorter questionnaire, with just the key participation variables (to monitor, for example, progress against national indicators). This would afford significant cost savings, while still providing the precision required to measure subtle national changes in levels of participation. More detailed policy-driven analysis (e.g. of background data and motivations for participating/not participating) could still be done using the core sample of

5,000, though for some small subgroups of the population this would probably require pooling data over more than one year.

4. **Move Taking Part to a part rotating panel design.** One of the perceived shortcomings of the TP repeat cross-sectional design is that it does not give any longitudinal data on individuals. So, change over time can be measured in net terms, but the gross change (that is the positive and negative components of change) in participation is unmeasured. Furthermore, the survey does not collect information on individuals' reasons for changes in participation over time. One way of addressing these concerns, would involve introducing longitudinal data collection to the survey. On the other hand, the most important requirement for TP is that it needs to be able to generate robust cross-sectional statistics each year on a large sample of the general population. So a purely longitudinal design would not be acceptable. One option would be a design that includes a panel *element* (so that some individuals are re-interviewed a small number of times) but that also incorporates some fresh sampling, so that new people are added to the study each year. The simplest design that meets these criteria would be a one year rotating panel.

Chapter 1: Introduction

DCMS and Sport England commissioned NatCen to undertake an independent review of the Taking Part and Active People surveys. The aim of the review was to scope the options for collecting national and local level participation data, and to consider how the two surveys might be merged or modified in order to reduce costs, while still meeting the requirements of stakeholders. The contracts for both surveys finish in early 2011. In advance of that time, DCMS and Sport England wish to consider the design of both surveys. In order to help this process, they have funded this review.

Taking Part has been funded by DCMS, Sport England, Arts Council England, the Museums, Libraries and Archives Council, and English Heritage since 2005. It is a face-to-face continuous survey of adults aged 16 and over, and children aged 5-15. The interview lasts around 45 minutes, and information is collected on participation in sport, arts, museums and galleries, libraries, archives, and heritage. Until recently, its key objective has been to provide robust measurement for DCMS's Public Service Agreement (PSA) and Departmental Strategic Objective (DSO) targets (on increasing participation in culture and sport)[1]. With a large sample size (14,452 adults and 2,500 children aged 11 to 15 in 2008-2009[2]) it allows detection of small changes in the population's participation levels. As well as providing data on participation, it also provides information on non-participators, and improves understanding on the barriers to participation in culture and sport. Taking Part also collects data on social capital, volunteering and community cohesion. The large sample size enables ad hoc analysis of relatively small subgroups of the population (e.g. ethnicity, religion, social class etc).

Active People is a short (around 18 minutes) telephone survey funded by Sport England. The survey provides local authority level data on sports participation, with a current sample size of around 500 in each of the 354 local authorities in England. Most of the Active People questions are also included in Taking Part, though the different mode means that the questions are worded and structured differently[3]. Since 2008, Active People has also included some questions on cultural participation and from January 2009 additional physical activity questions funded by the Department of Health. Active People is used to measure public targets, for example National Indicator 8 (participation in sport and active recreation), Sport England's '1 million' Strategy target and the wider 2 million sport and physical activity Olympic Legacy target. It is also used to measure National Governing Body of Sport targets for growing participation. Since 2008, Active People has also included some questions on cultural participation (funded partly by the Taking Part budget).

This review was primarily a desk research exercise. In addition, feedback was sought from key stakeholders across a number of government departments.

[1] PSAs and DSOs are now in the process of being phased out.

[2] The sample size varies year on year, and in previous years has been in the region of 24,000 adults and 3,000 children.

[3] Taking Part uses showcards to list the activities, while Active People asks open questions and the interviewer codes responses.

Based on the desk research and consultation with stakeholders, a number of options were drawn up. These options were presented to DCMS and Sport England, and a meeting was held to discuss each in turn. At this point, some of the options were ruled out. These options, and the reasons why they were ruled out, are covered in Chapter 3. Others were highlighted as potential candidates for the future, and these (along with their implications) are written up in Chapter 2.

Of course, one option would simply be to make no changes to the current design of the two surveys. Based on our stakeholder consultation, it seems that the two surveys are fulfilling their purpose and satisfying different needs (broadly speaking, TP tracks national targets, providing the national picture of participation in culture and sport. It also generates data for policy evaluation and appraisal. AP monitors local performance) and measures prevalence of participation in different sports at levels of precision suitable for tracking change over time. Any change to the current design may undermine one or other of these requirements. However the current design does not generate any longitudinal information on individuals. One of the options presented in Chapter 2 is an attempt to address this; the others suggest ways of maximising the efficiency of the surveys.

One issue with the current design is that the two surveys generate slightly different estimates of key participation statistics. It might be possible to reduce this problem by modelling the difference between the two surveys and then applying survey weights to AP to bring it into line (nationally) with TP on a number of key participation and demographic variables. This would however slow down the production of AP statistics and is likely to lead to objections that the AP data is being 'manipulated'. On discussion we concluded that this would not be an acceptable way forward.

Part of the remit of the study was to consider needs for data on participation among children. This is discussed in Chapter 4.

Chapter 2: Options for the future

Option 1: Substitute some of the Active People sample with Taking Part cases to reduce the size of the Active People-specific sample.

Each annual Taking Part survey covers all 354 (lower tier) local authorities in England and could in future years regularise this so that the numbers per LA are roughly equal. This would generate only a small loss in precision for Taking Part national estimates.

This option would use the Taking Part sample to provide partial sample for each LA for Active People, thereby reducing the size of the Active People sample.

The 2008/09 Taking Part sample size for adults (aged 16 and over) was 14,452[4]. This would provide an average of 41 respondents in each of the 354 English local authorities. This represents about 8% of the minimum necessary Active People sample (N=500) per year.

The main advantage of this approach is that it substantially reduces the costs of the Active People fieldwork; though combining the surveys would make the generation of LA-level statistics more complex, and the additional costs needed to handle this would slightly offset the savings. There would also be initial one-off set up costs. Three key barriers are:

a. **The differences in the TP and AP questionnaires**[5] would mean that generating combined variables could potentially be tricky. Certainly secondary users of either survey would find handling data from two different questionnaires problematic. One way around this would be to change the two questionnaires so that they match exactly on key questions[6]. In practice, only Active People questions would be suitable for both face-to-face and telephone interviewing[7]. This would mean replacing some TP questions with their AP equivalents. The detail of the Taking Part questions could be retained, but where lists of activities are presented on showcards, these would need to be turned into open-ended questions (which the interviewer would code).There would be obvious implications in the time trends for TP estimates in doing this.

b. **The Taking Part sample is clustered** within postcode sectors where the selected addresses within a 'cluster' define an interviewer's allocation of addresses for a month. In any LA the 41 TP respondents would therefore

[4] Between mid April 2008 and mid April 2009.
http://www.culture.gov.uk/images/research/PSA21-FinalBaselinereport.pdf

[5] Taking Part uses showcards to list activities, while Active People asks open questions, the responses to which are coded by the interviewer.

[6] Where two questions are similar, they would need to be replaced by one consistent version, as it would not be efficient to include two questions on the same issue.

[7] Telephone surveys cannot use showcards, and so the Active People format would need to be followed.

be clustered within a very small number of postcode sectors, with the interviews being carried out over a short period of time (a few months). This is in contrast to the AP sample which is completely unclustered, and where there is very tight control over the distribution of the sample across each week of the year. It is important to note that the effect of the TP clustering should not introduce bias if added to the AP sample, but the effect of the geographical and time clustering would increase the standard errors of AP estimates very marginally[8].

One possible 'solution' would be to move TP to an unclustered design. This would significantly increase the costs of TP, because of the need for interviewers to do additional travelling between sampled addresses. As far as we are aware, with the exception of the extremely large ONS Integrated Household Survey (IHS), no government face-to-face interview surveys are unclustered. And even the IHS uses clustering by month (so that all those sampled in a particular area are interviewed in the same month). The costs of unclustered samples are simply too great.

c. **The fieldwork period for TP is longer than for AP**. TP interviewers will try and make contact at addresses for longer than AP interviewers will try to make contact by phone. This, coupled with the fact that the TP interview is longer and more complex than AP, means that TP data is likely to take longer to generate and prepare than AP data. The implication is that the production of LA-level estimates is likely to take longer if the two surveys are combined (because Sport England will need to wait for the TP data to be available). However, preliminary discussion with the survey organisation suggests that any extra time lags would be minimal.

If the cost savings to AP are considered to be sufficient to make overcoming these barriers worthwhile we recommend that procedures are established for merging the two datasets. Without this there is a significant risk that there will be delays in the production of local level statistics, at least in the first year. In addition some assessment should be made of the scale of change in LA estimates that might be introduced purely as a result of the design change. A dummy run would be possible using existing AP and TP data. This dry run would ideally cover:

- The generation of macros to merge the datasets and produce common variables;
- An assessment of how much the addition of TP cases changes the AP local estimates;
- An assessment of the change in standard errors associated with the inclusion of the TP cases.

In practice we would expect the change in LA estimates to be small, simply because TP will contribute such a small percentage of the total sample size each year. But demonstrating that this is the case would be sensible – if only to have a ready answer to questions being raised in the future as to whether observed change over time might be attributable to the design change.

[8] The clustering would only affect 8% of the total sample per LA, so any impact on standard errors would be minimal.

Option 2: Move Taking Part from a 'one adult per household' design to an 'all adult per household' design

At present the Taking Part design involves the selection of:

- One adult per household (with that adult being selected at random)
- One child aged 11-15 (with that child being selected at random)
- One children aged 5-10 (with that child being selected at random, and then interviewed by proxy via the adult)

Under this design, in order to achieve interviews with 15,000 adults, a similar number of households have to be persuaded to take part. This means starting with a sample size of around 25,000 households. One relatively straightforward way of reducing the cost of the adult survey would be to move to a design where, instead of selecting just one adult per household, two, or even all, adults are selected. Given that the average number of adults per household is around 1.8 as few as 14,000 households would be needed for the survey rather than the 25,000.

However, in practice the arguments for and against selecting multiple adults per household are not so straightforward. The **arguments for retaining the one adult per household model** include:

a. **The statistical argument**. Interviewing all adults per household introduces within-household clustering effects. These effects reduce the precision of survey estimates and standard errors are inflated. But this argument is usually countered with the fact that selecting one adult per household also leads to a loss of precision, because the data has to be weighted to compensate for the under-representation of adults from multi-adult households. And in most instances the impact of weighting is far more detrimental than the impact of within-household clustering. Furthermore, for analysis by sex, the within-household cluster size will be, on average, very small so any effects of clustering will be minimal for this analysis. So, on balance, the statistical argument is probably in favour of selecting multiple adults per household. These arguments are set out in more detail in Appendix A.

b. **The burden and subsequent non-response argument**. Another argument for selecting just one adult per household is that it keeps the survey burden on households down. This, in turn, helps to maintain high survey response rates. This is the reason why most UK government surveys adopt the 'one adult per household' design, although there are notable exceptions such as the Health Surveys for England and Scotland, the National Travel Survey, and the UK Household Longitudinal Survey (UKHLS). In practice, if the response rate was to drop significantly then more addresses would need to be issued to compensate (slightly offsetting the apparent cost saving) and, there would of course, be increased risk of non-response bias. The TP interview is already long, and, we understand, relatively hard to sell on the doorstep. Preliminary discussion with the survey organisation suggests that response rates might not be detrimentally affected in households with more than one adult (but no children), since once one interview has been done, it is relatively easy to persuade another adult to take part. In households with children, however (where there is the possibility of doing an additional 10-15 minute interview with a parent of a child aged 5-10, and a 20-25 minute interview direct with a child aged 11-15) the burden may be too great and have an adverse

effect on overall response. This could, of course, be trialled before being rolled out.

c. **The impact on the child sample**. Probably the strongest argument for the one adult per household model is that having a large number of households in the survey generates an adequate child sample size. If, instead, multiple adults were to be selected then this would reduce the child sample to around 60% of its current size. For the 11-15 year olds, the sample might be partially increased back up from this 60% level by selecting all 11-15 year olds for interview rather than just one, but this would certainly not get the sample size back to current levels. In addition this would further increase the burden on large households.

   So, under the 'all adult per household' model, in order to maintain the current sample size of children, an additional boost sample of children would be needed. The approach used on the Health Survey for England (in survey years when a large child sample is required) is to select 'main survey addresses' and 'boost survey addresses'. At the 'boost' addresses a short doorstep screening interview is carried out to establish whether there are children at the address. Where there are, interviews with (or about) those children are carried out. Adopting this approach for TP would mean dividing the current sample size of addresses per year into two: around 60% main addresses and 40% screening addresses.

   This design is probably still marginally less expensive than the current design, but the saving will be much smaller than could be achieved for a purely adult survey.

In contrast, the **arguments for moving to an 'all, or multiple, adults per household' design** include the following. (We have not repeated the statistical arguments here.)

a. **Cost.** Selecting all adults per household rather than one would reduce the survey costs, although, for the reasons set out above, perhaps not very greatly[9]. One element of the cost saving is that household information does not have to be collected from all those interviewed – it can be collected just once from a suitable 'household informant'

b. **Analytic advantages.** Collecting data on multiple adults in a household allows for analysis of the degree to which household members follow similar patterns of participation. Understanding whether participation is a household level decision or an individual level decision has potential implications for policy, so a survey design that allowed for this type of analysis could be more powerful than the current design.

The arguments presented above apply to the Taking Part survey only. We do not recommend that Active People moves to an 'interview all adults' model because of the probable impact on the response rate for a telephone survey. (Much longer interviews are more tolerated face-to-face than by telephone, and the initial negotiation about what taking part in the survey will entail is far easier face-to-face.)

[9] This could be explored with the survey organisation.

There is no reason to expect that the change in sampling design described above would affect time series data. This could, of course, be tested by doing a split-run experiment in the first year (where one half of the sample remains one adult only, and the other half moves to the new design).

Option 3: Have a smaller core Taking Part sample with a full interview, boosted by a sample with shorter interviews

This would involve having a core national Taking Part sample (say around 5,000 cases). This core sample would be given the full TP face-to-face questionnaire. An 'extra' sample of (let's say) 10,000 respondents would be given a much shorter questionnaire, with just the key participation variables (e.g. to monitor the PSA21 Indicator 6). Active People would remain as it is.

The 'extra' TP sample could be obtained either face-to-face, or, in principle, by telephone. However, a telephone approach would have important disadvantages in that it would introduce two modes into the survey, and thus, potentially, mode effects. This would mean that the two parts of the survey could generate slightly different participation estimates (as is the case, to an extent, between Taking Part and Active People). This could, in turn, undermine the reliability of the survey results. This approach would also affect the time series, i.e. not all results would be comparable with data obtained in previous years.

Moreover, a telephone survey would have a considerably lower response rate (Active People obtains a response rate of around 28%, compared with around 60% for Taking Part). We estimate, with a face-to-face survey, that a response rate of 60-62% could be obtained for the boost sample (with the short questionnaire), and 57% for the core sample. Furthermore, the increasing number of mobile phone only households makes obtaining a fully representative sample difficult.

A face-to-face approach for the boost would still allow significant cost savings, (the survey organisation has estimated savings in the region of 15%), while still providing the precision required to measure subtle national changes (e.g. for PSA21 Indicator 6). More detailed policy-driven analysis (eg of background data and motivations for participating/not participating) could still be done using the core sample of 5,000, though for some small subgroups of the population this would probably require pooling data over more than one year.

Further cost savings could be made by interviewing all adults in a household, although all the arguments for and against raised above would apply.

Option 4: Move Taking Part to a part rotating panel design

One of the perceived shortcomings of the current repeat cross-sectional Taking Part survey design is that it does not give any longitudinal data on individuals. Consultation with stakeholders revealed that there was a strong interest in longitudinal information[10]. The current design measures change over time in net terms, but the gross change (that is the positive and negative components of change) in participation is unmeasured. Nor does the survey collect information on individuals' reasons for

[10] In contrast, consultation around the Active People survey suggested that there was no requirement for local level longitudinal data.

changes in participation over time[11]. There are a number of designs that could address these concerns, all of which involve introducing an element of longitudinal data collection to the survey (where, by 'longitudinal' we mean that individuals are approached for interview on more than one occasion). However the main constraining factor is that the most important requirement for TP is that it needs to be able to generate robust cross-sectional statistics each year on a large sample of the general population. So any design that seriously compromised this would not be acceptable.

This requirement rules out moving TP to a simple panel design, where a sample of individuals is selected at Time 1 and then the same individuals are re-interviewed repeatedly. The attrition rates over time on studies using that design tend to accumulate, so deriving cross-sectional estimates after a number of waves gets increasingly problematic.

What is not ruled out, however, is a design that includes a panel element (so that some individuals are re-interviewed a small number of times) but that also incorporates some fresh sampling each year, so that new people are added to the study each year. The simplest design that meets these criteria would be a one year rotating panel, which would work as follows:

- Year 1 would be a straightforward cross-sectional sample (where Year 1 could, in principle, be the current TP year, if sufficient contact information has been collected)
- In Year 2 a proportion (one half or less) of the Year 1 respondents would be selected for re-interview. The remainder of the required sample size for Year 2 would be selected as a new sample.
- In Year 3, all or a proportion of the new Year 2 sample would be selected for re-interview. The remainder of the required sample size for Year 3 would be selected as a new sample.
- And so on.

Under this design, individuals would be selected for interview either once or twice. Nobody would be selected three times. The table below shows an illustration of a design that aims to achieve 1,000 interviews per year. Fresh samples are shown in blue; follow-up interviews are shown in red. The TP design might, in principle, be designed as a scaled-up version of this model. The assumptions in the table are that the response rate for a fresh sample will be around 57%, that 90%[12] of the follow-up sample agree to be recontacted the following year, and that 80% of those are re-interviewed (an overall response rate of 70% for the follow-up sample).

Select Y1	Respond	Select Y2	Respond Y2	Select Y3	Respond Y3	Select Y4	Respond Y4

[11] While a limited amount longitudinal data will be available from 'Understanding Society' in 2014, consultation with stakeholders suggested that there was an interest in obtaining more detailed data before then.

[12] Currently a general question about recontact (at some time in the future) is asked at the end of the TP survey. 80% agree to this. With a more focused question we believe that this could be increased to around 90%

(base year)	Y1						
880	500						
880	500	500	350				
		1140	650	650	450		
				965	550	50	380
						1090	620

Note that with this model it is important that no more than half of the sample each year is selected for follow-up. If more than this are selected then the pool of new respondents in the follow-up year would be relatively small, and this would restrict the numbers available for follow-up in the subsequent year.

One considerable disadvantage of this approach is that the time interval between interviews is very short – just one year. This means, of course, that data will be generated that aids understanding of the nature and drivers of short term changes in participation behaviour, but will not contribute to an understanding of the longer-term trends for individuals. In principle the design illustrated above could be modified to partially deal with this:

- Individuals could be approached for interview more than twice. Two follow-up interviews would probably not be too detrimental to response rates; or
- The interval between the first and second interview could be set at, say, two years rather than one. Having a longer interval would reduce the response rate to the follow-up survey because more movers would have to be traced, but a reasonably high response rate should still be achievable. A three year interval would probably be too long.

Illustrative designs for these two options are given in the tables below. In practice there are several ways that the samples could be allocated each year between repeat interviews and new interviews, so these should by no means be taken as recommendations. They merely illustrate some possibilities.

Illustrative design for a three wave rotating panel:

Select Y1 (base year)	Respond Y1	Select Y2	Respond Y2	Select Y3	Respond Y3	Select Y4	Respond Y4
1050	600						
700	400	400	280	280	200		
		1260	720	400	280	280	225
				900	510	400	280
						880	500

Illustrative design for a two wave rotating panel with a two year interval:

Select Y1 (base year)	Respond Y1	Select Y2	Respond Y2	Select Y3	Respond Y3	Select Y4	Respond Y4	Select Y5	Respond Y5
877	500								
877	500			500	325				
		877	500						
		877	500			500	325		
				1190	675			500	325
						1190	675		

Moving TP to a rotating panel design would have a number of implications for the statistical precision of survey estimates. The main ones are as follows:

- The follow-up sample would be subject to two (or more) levels of attrition which would increase the risk of non-response bias. This could be (at least partly) offset by the fact that we would have a good deal of information on non-responders to inform non-response analysis and weighting.
- Any change over time statistics that involve comparing some of the same people over time should be measured with more precision than can be achieved with the current repeat cross-sectional design. For example, under the rotating panel where a proportion of the sample is re-interviewed after a year, change over time over the space of a year will be measured more precisely. But change over time over two years will not be measured with more precision because, over a two year interval, there will be no overlap in the samples. The improvement in precision for year on year change is unlikely to be very large (because only part of the two samples being compared will be overlapping) but it should, nevertheless, be significant (assuming that individual participation is correlated year on year). Some of the potential improvement in precision will be lost because the panel element of the study will need more complex non-response weights applied (to account for the attrition) but we would not anticipate this counter-effect to be large. In practice, it is very difficult to estimate what the precision gains for a rotating panel will be, so the anticipation of gains should not drive the decision. Rather, we believe, that the decision should be made on the consideration of the analytic advantages that a panel element introduces.
- Although change in time statistics should be measured more precisely, any cross-sectional statistics that rely on combining two or more years of TP data will be measured less precisely. This is because the combined sample size of individuals will be smaller than the combined sample of completed interviews – and it is the sample size of individuals that drives the statistical precision. We understand that, for the most part, there is very little analysis that relies on combining two or more years of data so this may not be a major problem. But it may be an issue for analysis of small sub-groups, where being able to accumulate several years of data might be useful. It is worth noting here that other major government surveys, such as the Health Survey for England, have not adopted a rotating panel design precisely because to do so would reduce the

capacity for generating rolling averages across multiple years of data for small sub-groups (such as geographical areas smaller than GOR).

- In contrast rolling averages calculated within a single 12 month period of TP would not suffer a loss of precision, because in any 12 month period no individual would be interviewed twice. This is on the assumption that those re-interviewed would be approached around 12 months after their first interview. In general, we would recommend that all follow-up interviews take place in the same month as the first interview, so that seasonal effects in change over time estimates for individuals are controlled for.

A part rotating panel approach may involve some cost savings in that half of the interviews would not involve 'cold calling', and interviews at these households would be somewhat easier to obtain. But this would be, at least partly, offset by the need to trace those who had moved (and travel to interview them) within the last year (around 10% of the sample). As a result, the survey organisation estimate that cost savings would be negligible.

Since telephone numbers could be collected at the first interview, subsequent panel interviews could be done by telephone, and this would lead to substantial cost savings. However, there are important disadvantages with this mixed mode approach in terms of low response rates for telephone surveys, the need for a much shorter interview, and mode effects.

Whichever mode is chosen, the first year would involve extra outlay in terms of set up costs: piloting, changing booking in and data management systems to incorporate a follow-up element, as well as designing and programming the longitudinal module(s) of questions.

We anticipate that the extra module would only be asked of those whose level of participation had changed (either up or down) since (the same time) the previous year, but this in itself would involve some relatively complicated programming and data management, to ensure that relevant information was fed forward.

Moreover, a module would (presumably) be required for each of the six areas of participation (sport, arts, museums and galleries, libraries, archives, and heritage). This module would include a number of questions around the reasons behind any change and, given that these are likely to vary between types of activity, it will probably need to be somewhat 'area-specific'. Finally, since the motivations behind any increase in activity will differ from reasons behind a decrease, there will need to be two sets of questions/responses for each of the six activity areas. It should be noted that a quantitative survey would be limited to pre-coded responses to these questions (perhaps with an option for an 'other reason' which could be written in and coded later)[13]. Therefore, information on reasons behind change would be somewhat broad[14]. Nevertheless, a longitudinal approach would provide data on the magnitude of change, and the main reasons behind it.

[13] We would recommend a qualitative study to explore reasons behind change in participation, (across the six areas), to inform the questionnaire development.

[14] If detailed information is required on subtle psychological/behavioural motivations behind change in participation, then a qualitative study would be more appropriate.

As well as the questionnaire design issues discussed above, there would inevitably be a number of other practical details to deal with. One decision would be whether to randomly select individuals for follow-up from within the sampling clusters (that is the postcode sectors selected for the study) or whether to randomly assign whole clusters to follow-up. The latter approach would be less expensive, but longitudinal estimates would have slightly higher standard errors (because the panel element would be more clustered). If TP were to move to a model where all adults per household are interviewed, a decision would also be needed on whether to follow whole households or just randomly selected individuals within households. The former approach would be less costly.

If only adults were followed up then at the follow-up interview the panel element of the survey would not include any 16 year olds. This is easily overcome by ensuring that 15 year olds, from the original sample, are also followed-up. Potentially more problematic is the fact that if all adults and children are followed-up, then the follow-up survey would not include five year olds (because the lowest age in TP at present is five years). The simplest solution would be to weight up the five year olds in the fresh sample. Related to this coverage of age issue, the panel element would not include any households new to GB. Again, the new households in the fresh sample can be weighted up to deal with this if a question is included that identifies new immigrant households, though the percentage of new households is so small (less than 1% of the population[15)] this adjustment may not be merited.

Another issue is around ethics. Most ethics committees now require that, if there is an intention to follow-up respondents in the future, then this must be made clear in the initial approach (e.g. the advance letter). While the Taking Part survey does not require ethical approval, nevertheless good practice in this area should be followed. This could, potentially, reduce initial response rates slightly for the longitudinal part of the panel (though this could be explored through piloting). It would also rule out using the current year's survey as the starting point for a future panel.

A final potential complication is that the TP sample size has tended to vary from year to year, with, we understand, the decision on sample size being taken on an annual basis. This doesn't necessarily marry well with the pre-planning required for a panel survey. But, assuming that all those asked to take part in the survey are alerted to the fact that they may be interviewed again, there do not appear to any major reasons why flexibility around sample size cannot be factored in. It simply means that decisions on sample size in any one year have to be made taking into account the implications of those decisions on the future panel design.

There is no reason to expect that moving to a rotating panel design would have any impact on time series data.

[15] http://www.statistics.gov.uk/CCI/nugget.asp?id=260

Chapter 3: Options considered, but ruled out

A number of options were initially drawn up, presented to stakeholders for discussion, and subsequently ruled out. These options are discussed below.

Option 1: Change Active People from a telephone to a postal survey

Active People is currently a telephone survey. Changing the mode to a postal survey would reduce costs since postal surveys are considerably cheaper to run. The National Indicators that rely on surveys tend to use postal rather than telephone methods (e.g. the Place Survey and the Services for Disabled Children Survey).

Though the costs savings would be significant, after discussion with the Steering Group this option was ruled out for a number of important reasons:

1. No sampling frame exists with addresses for named individuals. The only potential sampling frame would be the Postcode Address File, and this does not hold any information on individuals. Therefore, random selection of an adult within a household would not be possible and it would be very difficult to control who completes the questionnaire.

2. The AP questionnaire is currently CATI (computer assisted telephone interview) which allows complex routing. The questionnaire structure involves the respondent generating a list of the sports that they do. Then, for each one, they are routed through questions about the number of times per week, length of time per session, and (for those activities that aren't considered automatically moderate or vigorous) whether it 'raised breathing rate' and whether 'out of breath or sweaty'. (Without CATI, these intensity questions would need to be asked of each activity.) This kind of routing would be very difficult to present in a postal questionnaire, especially given that respondents will vary in terms of the number of activities done. Data quality would be compromised due to completion errors by respondents, and the questionnaire would look very thick, deterring participation.

3. Postal surveys tend to achieve lower response rates than telephone. The current response rate for AP is about 25%. A postal survey with a short, well designed questionnaire could, in principle, achieve this level of response. However, the nature of the questionnaire content does not lend itself to a simple, short questionnaire.

4. The mode change would also have implications for the analysis of time series data.

Option 2: Change Taking Part to a cheaper mode (postal or telephone)

Taking Part is a face-to-face survey. Significant cost savings could be made by changing the mode to postal or telephone. A postal mode would not be feasible

because the questionnaire is long and complicated, and there would be similar problems around lack of sampling frame. A telephone survey could be considered, though cuts would have to be made to the interview length. Currently the TP survey length is around 45 minutes. An ideal length for a telephone survey would be around 20 to 25 minutes. This option was ruled out for a number of reasons:

1. A significant advantage of the current TP questionnaire is that it covers six areas of participation (sport, arts, museums and galleries, libraries, archives, and heritage), allowing analysis of the association of participation across these areas. It also covers issues around motivations, and barriers, to participation. Halving the questionnaire length would obviously reduce the usefulness of the data in this respect, and the survey would, essentially, become a smaller version of Active People.

2. Response rates from a telephone survey would be significantly lower than face-to-face and, as previously mentioned, the increasing number of mobile phone only households makes obtaining a fully representative sample difficult.
3. The mode change would also need to be considered in any analysis of time series data.

Option 3: Reduce the TP sample size.

The 2008/09 annual sample size for Taking Part was 14,452 adults. This is already considerably smaller than when the survey first began in 2005 (around 27,000 adults). Cutting the sample size would, obviously, produce cost savings. Even halving it would still allow for monitoring of reasonably subtle changes in the population overall, and analysis of small subgroups could be done (if necessary) every two to four years (depending on the size of the sub-group). This option was ruled out for a number of reasons:

1. PSA21 Indicator 6 requires a 'statistically significant increase in two or more different cultural or sports sectors'. Currently, the survey is able to pick up statistically significant differences between years of less than 2 percentage points. Discussion with stakeholders suggests that this level of precision is necessary, particularly with Communities and Local Government's (the lead department for PSA21) requirement for quarterly data, so there is a need for a large sample size at least on the variables that contribute to PSA21.

2. While PSA21 is the indicator for the current spending review, policy needs will change in the future and TP needs to be 'future-proof'. For example, in the past, TP was used to monitor PSA3, which stipulated increase (of 2-3 pps) in particular subgroups of the population (eg BME groups). Clearly, this requires a larger sample size than monitoring change overall. Meanwhile, the new government is moving away from such indicators all together,

3. Requests for analysis by small subgroups (e.g. disability, religion and ethnicity) are common, and TP currently provides an invaluable source for this kind of analysis. Reducing the sample size would undermine the survey's ability to produce this kind of data.

Chapter 4: Data on children

Taking Part includes a sample of around 2,500 children aged 11-15. From 2008/09, a sample of children aged 5-10 was included, with the information collected by proxy through parents/guardians. Parents of children aged 5-10 are asked only about activity that took place out of school while children aged 11-15 are asked about their in and out of school activity.

The Taking Part sample size is too small to provide local level estimates[16]. Some stakeholders expressed a requirement for data on children at a local level. In particular, they were interested in the proportion of children doing at least five hours of physical education and sports per week in and out of school.

Obtaining local level data on children would involve adding children to the Active People survey. As well as the cost implications, including children could have a detrimental effect on response rates, though this could be offset with an appropriate incentive. This could be piloted in 100 households in each local authority. Within each of the 100 households data on children's participation would be collected from one randomly selected child per household. Data might be collected a) directly from the child if aged 11-15, or b) by a parent/guardian on behalf of the child if aged <11. It should be noted when calculating a likely achieved sample that around 25% of households contain children – this means than the annual sample size of children per LA would probably be no greater than 125, unless boosting was introduced.

There are a number of options for generating a boost sample of children. The main ones are described below. It is difficult to make a recommendation on which would be best. A final decision would depend on whether a postal/self-completion survey would be a reasonable alternative to a telephone survey (which in turn depends on a clear articulation of what data is required) and whether self-completion by children is adequate for the younger children (rather than parental report). In addition, access to some of the sampling frames (Child Benefit Records and the National Pupil Database) would depend on agreement being reached with the government department responsible for the data.

Option 1: Telephone screening based on a Random Digit Dialling Sample

The only method that would allow for telephone interviews with, or about, children appears to be screening of an RDD sample. That is, household telephone lines would be selected via RDD (as in Active People). The interviewer would then establish whether there are any children at the household, and, if so, an interview would take place. Around 25% of households have children so, allowing for refusals, around 500-600 households would need to be contacted and screened to generate 100 child interviews. There are, however, likely to be considerable problems with this approach. On the face of it, cold-calling households to ask about children would generate a lot of anxiety amongst parents, so, although not ruled out, some means of being able to

[16] DH obtain data on children's participation through HSE, which provides GOR level results

reassure parents that the research request is legitimate would be needed. We could not recommend this approach without it being thoroughly tested.

Option 2: Door-step screening for children

PAF addresses could, in principle, be screened for children. At households with children a self-completion questionnaire might then be left for a randomly selected child (assuming that the cost of face-to-face interview would be prohibitive). This would be a relatively expensive option, but would get past many of the objections to Option 1 (especially given that the interviewer could present ID and an advance letter could be sent).

Option 3: Selection of children from Child Benefit records

Many surveys of children have in the past been based on samples from Child Benefit (CB) records. CB records do not have comprehensive telephone numbers so a telephone approach would be ruled out. But a postal questionnaire could be mailed.

There are, however, some problems with CB records that make it less than ideal. The main problem is that CB tends to be paid directly to bank accounts, so there is no incentive for parents to notify HMRC if they move address. Addresses are often incorrect, and are more likely to be incorrect for older children than for younger. In addition, since the highly publicised loss of CB records data, HMRC and DWP have become more reluctant to allow the use of CB records as a sampling frame for research. It is unclear how this will develop in the future, but we cannot say for certain that a sample from CB records would be possible.

Option 4: Selection of children from National Pupil Database records

Some recent surveys, such as the surveys to support the evaluation of the Free School Meals Pilots, and the DCSF's 'Parental Experience of Services for Disabled Children' national indicator surveys have used the National Pupil Database (NPD) as a sampling frame. This covers all children in mainstream, public education. Telephone numbers are not included on NPD so a telephone survey would be ruled out, but, as with CB records, a postal questionnaire approach may well be feasible. In this instance the support of the Department for Education would be needed.

Option 5: A survey of children in schools

The final method would be to adopt the approach taken by the Department for Education for the TellUs survey. This would involve pupils being selected within school, and those selected completing a self-completion questionnaire. This approach would exclude parents entirely from the process, so would only be appropriate for pupils old enough to self-report their activities (so perhaps secondary school only). Having to go via schools would make this a potentially difficult option, but once a school is engaged, the within-school response rate would be fairly high.

A concern for this approach would be that response at a school level is positively related to the amount of sporting opportunities on offer, and this would, of course, lead to inflated estimates of participation levels. This would need to be carefully considered and addressed.

Even if an appropriate sampling method were agreed, there are still a number of challenges in collecting data on participation among children. As with all surveys, obtaining information from younger children themselves is difficult and often unreliable. For children aged ten or under, such surveys tend to ask the parent to provide this information (ideally with the child present to help). This can be problematic as most policy makers are interested in the amount of exercise done both in and out of school, and this proves particularly difficult as parents often don't know what their child does at school.

Appendix A: The statistical arguments for and against selecting multiple adults per household

There are a number of possible selection strategies for adults that could be adopted on a Postcode Address File (PAF) based sample. The two considered here are:

a. Selecting all adults per household
b. Selecting one adult per household.

Other, 'in-between' models involve selecting up to two adults. And it would also be possible to adopt more complex designs, such as always selecting the main income earner and then selecting one other adult at random. Such a design would ensure that in a large adult household, the person best able to answer questions about household finance is included.

Both A and B affect the standard errors of survey estimates. It is usual to adopt the design that, for a fixed cost, gives the smallest standard errors. Or, for fixed standard error(s), has lowest cost.

The two approaches impact on standard errors in different ways. For A, the fact that questionnaires are completed by all adults in a households leads to 'within-household clustering effects'. Essentially, the greater the similarity between adults within a household on their survey responses, the greater the clustering, and the less 'information' each new household member contributes to the survey. This clustering increases standard errors. At the extreme, if all adults within a household give identical survey responses the 'effective sample size' for the survey is just the household sample size.

Design B avoids within-household clustering but, the selection of just one adult per household means that adults from households with two or more adults (the majority of households) are under-represented. To derive unbiased estimates, the respondent data has to be weighted, the weight per respondent being equal to the number of adults in the household. The subsequent weighted estimates have larger standard errors than their unweighted equivalents.

In one sense, then, the issue of which of Design A or B to adopt depends upon whether the clustering effects on the Taking Part survey are more detrimental than the effects of weighting. In the absence of any data on within-household clustering this assessment is difficult to make. (To do so formally, other survey data around participation, taken from a survey that does include multiple adults, would be needed so as to derive estimate of the intra-cluster correlation coefficient. The Health Survey for England data on sports participation might be a possible source. But arguably clustering effects will be greater on participation variables, such as arts participation, that adults are more likely to do as families.) Nevertheless, another factor to take into account is that the 'select 1 adult' design is more expensive than the clustered design because more households have to be approached to take part. Given the higher cost

of Design B, in practice the clustering effects have to be worse than the weighting effects if the extra expense is to be justified.

A final important consideration is that the negative effect of clustering on standard errors depends on the cluster size. The smaller the cluster size, the smaller the effect. For a cluster size of one there is no impact on standard errors at all (because in that scenario the sample is effectively unclustered), and the impact will be small for cluster sizes of close to one. For 'all adult' estimates the cluster size will be just less than two (because the average number of adults per household is two). But for analysis by sex the cluster size will be close to one (since the average number of adult women or men per household is close to one). So for analysis by sex the effects of clustering on standard errors will be minimal. However, under the 'select 1 adult' model, the effects of the weights are the same across almost all key sub-groups (including sex). So for a survey where most published estimates are gender-specific, the statistical arguments are almost always in favour of the 'select all adults' design. This is one of the principle reasons why the Health Surveys for Scotland and England both include all adults per household.

department for
culture, media
and sport

2-4 Cockspur Street
London SW1Y 5DH
www.culture.gov.uk