Time Use Feasibility Study:
Topic 3 - New Dimensions in the Diaries

Contents

Table of figures

Table of tables

Executive Summary

Introduction

Time Use studies are carried out across the world and provide valuable data which can be used by social scientists, economists, policy makers, health researchers, psychologists and others to investigate a whole host of issues and to make a wide range of cross-national comparisons.

The key objective of this report was to deliver recommendations on the new well-being dimensions for the 2020 HETUS drawing on findings from:
1. Literature and best practices review on the impact of inclusion of such additional variables on response rate, burden and comparability of survey results.
2. Testing of various scales for the subjective well-being variables.
3. Cognitive testing of the general subjective well-being variables and variables related to the activities.

Literature Summary

Time use [research] studies (TUS) can help us understand what activities people do from day to day, as well as a wealth of other pertinent contextual information (e.g. where they are, who they are with, devices used) that can provide the basis for systematic integration of various well-being measures (Merz, 2009).

Time use data and well-being measures have previously been collected using questions in stylized surveys, time diaries, and the experience sampling method (ESM). Stylised survey questions ask for instance, how often do you engage in [activities]? Time diaries are a log of respondents' activities divided into fixed time intervals. In ESM respondents may be prompted at random instants through the day/week, to describe their current activities and well-being. While there are benefits and limitations associated with each approach, diaries are typically more valid than stylized questions, roughly comparable to ESM (yet much less expensive), and provide researchers the opportunity to collect associated contextual information (Baghal et al., 2014; Bolger et al., 2003; Harvey & Royal, 2000). Furthermore, it is widely anticipated that future diary design developments that harness modern technology (e.g. GPS for location data, online diaries/apps) will further expand the range of details that time diaries collect (Kelly et al., 2015; Fisher et al., 2015).

Time diaries have typically been validated in previous research using focus groups and cognitive testing as part of their design, development and evaluation. Several international TUS utilize time diaries including the Harmonized European Time Use Study (HETUS), which is a project coordinated by Eurostat with the aim of creating comparable and standardized time use statistics across European countries; and the American Time Use Survey (ATUS) which has collected a large number of time diaries each month from a sample designed to be broadly representative of the American population.

The UK government has prioritized the use of subjective well-being questions in policy research (ONS, 2018). Policy implications of studying well-being in TUS include evaluation of policy measures based on emotional responses to them, promoting greater well-being and reducing harm and suffering (Fisher et al., 2016). Despite being an ambiguous concept that lacks a universally-accepted

operational definition, and that cannot be directly observed/measured, the Stiglitz commission has further identified TUS as a key way to inform on well-being and progress.

Fisher et al (2016) explain that the more complex an instrument becomes the more likely people are to refuse to participate. Non-response in time use research may occur to a number of factors including features of the sampling design (Phipps & Vernon, 2009) and demographic characteristics of respondents (Abraham, Maitland & Bianchi, 2005). Relatively low response rates of 56%-58% in the ATUS have previously raised concerns there might be bias in time use estimates (O'Neill & Dixon, 2005). Therefore, it might be expected that adding well-being components to an instrument could increase its complexity and have a negative effect on response. However, there is some [albeit currently limited] evidence that including well-being questions had a positive (Fisher et al., 2016) or neutral (Fisher et al., 2016; OECD, 2013) effect on response rates.

HETUS guidelines recommend minimum 10-minute time slot "grains" for time diaries (Eurostat, 2001) and activity episode analysis emphasizes the context (why, where, when, with whom, duration, sequence) of activity participation. However, previous inconsistencies and deficiencies in the recording of time use, may be preventing a full understanding of how people actually use their time in their everyday lives (Kenyon, 2010), as multitasked or co-occurring activities like childcare, other caring activities, housework, communication activities, online activities, leisure etc., can have a significant impact quality of life, well-being and life chances.

In conclusion, the diary study methodology in TUS can provide researchers and governments with valuable information to inform social policy as well as in drawing conclusions for, and making comparisons between, the well-being of societies, populations and individuals throughout the world.

Research

Alongside the three studies, an expert panel was established via email that included people with varying expertise in time use or well-being measurement around the world. At various stages of the project, their input was sought to inform the methodologies implemented and selected for the present study. Their input and recommendations have fed into the research and the recommendations for HETUS 2020.

Three studies were completed as part of this report:
1. a focus group with researchers
2. a diary study with two types of diaries (measuring conceptions of "enjoyment" and "satisfaction")
3. cognitive interview follow-ups with diary respondents

Focus Group

The first study was a preliminary focus group[1] including 16 researchers (4 male, 12 female) from relevant fields to discuss important issues in the design of diary study instruments that incorporate well-being measurement. This included type of well-being measurement (e.g. enjoyment, pleasantness), rating scale issues (e.g. number/wording of scale points) and diary format. Key results included:

- The term 'enjoyable' was an easier and more effective way to measure quality of time than other terms.
- A potential alternative useful term would be 'satisfaction'.
- 7-point uni-polar[2] rating scales due to the greater response flexibility they provide.
- Respondents should be able to circle numbers, allowing for a visual comparison and easier differentiation between ratings

Diary Study

The second study was the diary study which was the primary focus of this research project. Diaries were designed to measure either satisfaction or enjoyment. Activities were recorded in ten-minute slots ranging from early morning to evening that provided space to record primary/secondary activities and other contextual details. Activity and contextual data was coded using a simplified version of the HETUS activity codes as part of an iterative process by coders trained in the design and analysis of time use diaries.

899 respondents who indicated in the Opinion and Lifestyles Survey (OPN) that they would be willing to participate in future research were offered the diary. 378 respondents consented to complete a diary and 159 respondents returned completed diaries to ONS, containing 2432 episodes. A total of 2418 activity episodes from 156 diaries were included in subsequent analyses (1175 from enjoyed diaries and 1243 from satisfied diaries). The overall response rate was 41.3%. Key results[3] included:

- Small but significant correlations between diary study well-being ratings and well-being reported in the respondents OPN survey
- All scale points were utilised on both well-being scales but a clear trend of selecting the upper end of both scales more was observed. A similar, yet not identical, bias was also observed in diary study respondent' ratings for well-being questions in the OPN survey

---

[1] Focus groups ensure comprehensive coverage of relevant issues while also facilitating rich and constructive discourse from a diverse range of perspectives, potentially including those previously unconsidered by the research team.

[2] Uni-polar scales range from 0 to a positive number. Bi-polar scales range from a negative number to a positive number (with zero typically representing the mid-point of the rating scale).

[3] It is important to note before reading and interpreting results from this study, that it was a pilot study, designed specifically to examine the functionality of a diary instrument containing different well-being measures, and in particular to highlight the kinds of analysis that might be achieved with such an instrument. There was no attempt to obtain a particularly large or representative sample and there was no intention in this study to draw any firm time use conclusions from the data or to generalise any findings beyond this particular sample.

- The well-being scales proved to be a useful addition to the diary instrument with regards to the types of activity analysis their inclusion permitted. Specifically, they permitted straightforward analysis of:
    o The effects of contextual details (e.g. type of day, smart device use, location, mode of transport, other people) on well-being ratings
    o The most and least enjoyable and satisfying main and secondary activities
    o Notable differences in enjoyment and satisfaction ratings for particular activities (e.g. the mean enjoyed and satisfied ratings for employment as a main activity were 5.5 and 4.6 respectively)
    o Well-being ratings for activities (e.g. housework, childcare) and contexts (e.g. location) associated with co-occurring and singularly-occurring activities.
    o Identification of the main activities most and least associated with secondary activities and vice-versa.

Cognitive Interviews

A sample of diary study respondents was selected to participate in telephone based cognitive interviews to fully evaluate the functionality of the diary and to identify what worked and what didn't work, as well as any potential refinements that could be implemented to improve its overall functioning in time use research.

The cognitive interviews centered on four primary themes: diary completion impressions, questions about the respondents' day yesterday[4], ratings and their understanding of rating scale terms (see appendix N).

In total, 22 participants (8 male, 14 female, mean age=59.3 years) from a range of socio-economic positions took part in this study. 8 of these participants had completed enjoyed diaries and 14 had completed satisfied diaries. Key results included:
- Most participants initially reported that they found it easy to complete the diaries but there were a few difficulties reported.
- A third of participant responses indicated that they found it easy to determine the main activity.
- Over half of their responses indicated they found it easy to distinguish main and secondary activities.
- When asked about their day yesterday:
    ▪ The activities participants found most satisfying included house/garden repair, housework and food preparation/consumption
    ▪ The activities participants found most enjoyable[5] included work, family time and hobbies/voluntary work
- A common theme across both diary types was that participants typically found it easier to rate novel activities and more difficult to rate ordinary, mundane activities. This 'novelty factor' also affected the classification of well-being ratings for "middle of the road" activities (i.e. those that were neither the most or least enjoyable/satisfying).

---

[4] The researchers deemed it most appropriate to ask respondents questions from the diary about their day yesterday rather than relying on their recall of their diary responses, given that this was a telephone interview and the duration between diary completion and subsequent interview may have been several months.

[5] Please note that participants **either** rated their enjoyment or satisfaction with yesterday's activities.

- Respondents provided words to represent what various numbers along the 1 – 7 scale meant to them. These "scale point" descriptor terms were categorized into factors or "components" of enjoyment and satisfaction. In total there were 6 factors, of these there were two satisfied 'factors' (i.e. task factors, individual factors) and three enjoyed factors (i.e. enjoyment factors, situational factors, task factors) that represented five or more continuous points on the respective scales.
    - Satisfied scale words included: "Basic task achievement", "Fairly routine".
    - Enjoyed scale words included: "Tedious/emotionless", "Novel activity".
    - Words for enjoyment and satisfied scales included: "Task completion", "Really enjoyed task".

  The upper scale-points of both scales were frequently represented by task process, outcome, and timing/novelty factors. These factors were also frequently used to denote scale midpoints, as were ordinal-level factors (for example: "ok", "neutral", "just under"/ "over halfway").

  Other factors represented scale points more irregularly, with only the satisfaction scale having terms that covered the full range of the scale (1 - 7).

- Participants most frequently defined both satisfaction and enjoyment in terms of:
    1. Task factors (i.e. achievement/completion/outcome)
    2. Emotional characteristics related to enjoyment (e.g. happiness, pleasure)
- Proposed synonyms included:

  Satisfied
    - Nice/Good (38%)
    - Adequate/Reasonable (38%)
    - Agreeable/Acceptable (25%)

  Enjoyed
    - Pleasure (20%)
    - Like (20%)
    - Content (20%)

- Most participants indicated that it would be easier to rate enjoyment of activities than satisfaction. Specifically:
    - Half of those who completed a satisfied diary indicated they would find enjoyment easier to rate than satisfaction.

Just under half of those who completed an enjoyed diary indicated that satisfaction would be harder to rate.

Recommendations
HETUS 2020

1. The literature suggests that the well-being element of time use surveys produces results policy-makers find more directly relevant. The literature also indicates that respondents find this element of the survey particularly relevant in recent time use survey implementations. This dimension would be used widely if collected, adding urgency for adding this field.
2. An increasing number of national time use surveys incorporate a well-being element. HETUS 2020 remains cutting edge by including this component in diary design recommendations. Missing this element out misses a major advantage of the effort and cost of conducting a time use survey.
3. The concept of enjoyment should be used to measure well-being in HETUS 2020 as it was more meaningful to respondents than the concept "satisfied". The enjoyment column produced a wider range of responses across both activities and context variables (mode of transport, location, who else was present), produced more variation across days of the

week and times of day, and produced more policy-relevant activity distinctions than asking about satisfaction. Alternative terms reviewed by this project were deemed less appropriate to a multi-purpose survey capturing a range of national statistics.

4. Respondents in the pilot diary and the cognitive interviews reported finding the satisfaction as opposed to enjoyment more relevant for reporting how they experienced routine housework, commuting and daily living tasks. People may not enjoy doing the ironing, brushing their teeth, or cleaning the bathroom, but they appreciate that such a task was performed well (or feel less happy if the task was not done well or they did not have time to properly complete the task). While this point is worth noting in terms of the limitations of choosing a single well-being measure, and this knowledge is relevant to understanding the experience of housework and other routine activities, the more blunt the information captured about the range of other activities, context of activities, and times of day make this a less useful option for measuring well-being. Capturing satisfaction is not relevant to measuring the volume and assigning economic value to the range of unpaid production activities more often performed by women, and contributes only modestly to this central function of time use surveys.

5. 1-7 uni-polar scales provide respondents with flexibility of various ratings whilst remaining generally comparable to other means of well-being collection (e.g. survey questions on well-being).

6. Diaries containing full scales with numbers that respondents can circle or otherwise mark allow respondents to visualize the scale to compare ratings. Enabling visualisation of scales will be easier once diaries are collected online or via apps as space restrictions would be less applicable.

7. For the period where paper diaries remain a mode of diary collection, a more flexible who else was present column with a tick box for alone and a box to fill in others present reduces the complexity of the diary grid and could create more space for an enjoyment section. The simplified, open-ended self-completion box captures a wider range of persons present than the existing HETUS design, and may prove easier to complete as respondents record these categories of people who are relevant to their daily activity account, rather than trying to make sense of the HETUS categories if these are not applicable.

8. Our effort to pilot a reduced day diary was not successful. Multiple respondents tried to fit in activities which clearly took place over the whole day into the 6 one-hour blocks. A limited consecutive observation window which allows respondents to fill in a complete story about a part of their day might prove easier to complete than a selection of hours over different times of the day.

9. Further testing of understandings and alternative definitions, conceptualizations, synonyms and opposites of enjoyment to:
   - Examine their relative suitability/appropriateness as alternative measures of enjoyment or as complementary measures as part of a multi-dimensional enjoyment scale
   - Examine their association with particular activities or activity types with a view to explain the using different terminology to explain enjoyment in different activity contexts
   - Particular consideration will need to be made to the concept that best translates across European languages for the 2020 HETUS survey.

Literature Review

Introduction

This literature review begins by introducing the concept of time use and of time use research. It will show how a variety of methods including surveys with stylized questions, time diaries as well as other methods have previously been used to collect time use data, and will discuss the advantages and disadvantages of these methods. Then, to provide the reader with contextual detail for the methodology and results presented in section 3.4 and chapter 4 of this report, the review considers in detail, the utility of cognitive interviewing techniques in validating diary studies. The literature review then outlines large national and international time use studies (e.g. American Time Use Survey, Harmonized European Time Use Study) that feature heavily in the research subsequently discussed.

The literature review then goes on to consider the type of data to collect in time use research. Given that the focus of the present study was to examine investigate the use of well-being ratings in time use diary-based research, there follows a detailed section outlining definitions, determinants and correlates of well-being that goes on to discuss the measurement of well-being, potential research outputs and policy implications of considering well-being in time use research. There are also many sections throughout this report that have varying degrees of association with well-being, and these are referenced throughout this review.

Discussion on the collection of activity data follows, including an in depth look at the benefits and drawbacks of collecting multitasking and co-occurring activity data. The review then goes on to consider other data that has been collected in previous time use research and that should be collected in future time use research. The next section considers the methodological implications of time use research with particular emphasis on response rate in general, and in the collection of well-being data in time use research. The final section considers the future of the diary method in time use research in light of recent technological developments (e.g. apps and smart phones/tablets).

Time use

Time use describes the allocation of time among various circumstances and subjective states. Gershuny (2011) views it as a key social indicator that can be used to understand the material welfare and well-being of individuals in societies. We can also learn a lot about society on the whole by finding out how people spend their time during a typical day (Kenyon, 2010), so in order to do this, it is centrally important at individual, societal and economic levels to collect statistical knowledge about the use of time for all conceivable activities. Merz (2009) refers to time as the compound dimension encompassing resource of individual activities and living arrangements. He explains that in general, any characteristic or information is only complete, when time is considered additional to the factual/socio-economic and geographic attributes. Time use research can include varying levels of detail. On a basic level, in understanding what activities people do from week to week or day to day, or on a more detailed level, showing what people are doing, where they are, who they are with and how they feel from minute to minute (Merz, 2009). In short, the aim of time use research is to be able to examine how people spend their time on a typical day and be able to explore differences between groups of people. According to a United Nations report:

"Time use statistics offer a unique tool for exploring a wide range of policy concerns including social change, division of labour, allocation of time for household work, the estimation of the value of household production, transportation, leisure and recreation, pension plans and health-care programmes, among others" (United Nations, 2009).

People's time allocation has been studied since the beginning of the 20th century (e.g. Bevans, 1913; Pembers-Reeves, 1913; Jahoda et al., 1975). Since these early studies, time use research has been used to collect data about a wide array of policy questions and social issues. Several studies have even highlighted the role of time use research in shaping social policy in many countries including Norway (Gross & Swirsky, 2002; Aslaksen & Koran, 1996); Canada (Canadian Women Making an Impact, 1998); South Africa (Budlender, 2002) and the Netherlands (UNECE, 2013). It has been argued that tracking long-term shifts in the daily behaviours of populations is one of the most useful research possibilities that time use research has to offer (Fisher et al., 2016; Abraham et al., 2005). While the scope and purpose of time use research varies enormously, a frequent aim of these studies is to provide better information on work performed by different categories of people (male and female in particular), with particular emphasis on unpaid work which is usually excluded from national accounts that underlie GDP (Elson, 2002; Aslaksen & Koren, 1996; Ironmonger, 1996; Cloud & Garrett, 1996); the ways in which domestic work is divided between the genders and family context (e.g. ICLS, 2013; UNECE-2013; Hook & Chalasani, 2008) – given that the biggest change in individuals' time use occurs with the arrival of children (OECD, 2012); healthy (e.g. leisure activities, family time and other quality of life behaviours) and unhealthy (e.g. extended static periods) activities (Phipps & Vernon, 2009); transport patterns (to name only a few – for a detailed review of more information collected by time use research, also see UNECE, 2013; Gershuny, 2011; Phipps & Vernon, 2009).

Exactly *when* these activities take place, during the day, week and year, is also significant for understanding well-being. A primary motivation of this literature review is to consider the instrument and comparability level effect of asking respondents about subjective well-being. Since all human states and activities occupy time, an appropriately designed TUS instrument could therefore in principle, provide a comprehensive account of rhythm and balance among all the conditions and circumstances of daily life (Gershuny, 2011). As such, time use accounts provide the basis for systematic integration of various measures of well-being. To this end, it is particularly important for studies such as this and future research investigating well-being in the context of time use research to understand the most appropriate methods of understanding this complex phenomenon.

Collecting time use data

Surveys with stylized questions

Survey design normally poses a problem of balance between information quality and information burden (Gershuny, 2011). The most familiar measures of time use are stylized time use questions in conventional questionnaires which ask questions like: How often do you engage in [activities]? Who usually does the [routine/domestic work]? Or, how much time do you usually spend doing [activities]? Stylised surveys have also been used to understand well-being using quality of life measures (such as earnings) and by asking people a single question related to life happiness and satisfaction. Stylized time use questions provide the lowest burden to respondents but can also provide very poor-quality information due to a range of problems associated with this approach.

Firstly, respondents may lack the ability to accurately recall activity details in this way. Sunnenberg et al (2012) point out that the extent to which time use as retrospectively reported in a questionnaire reflects people's actual behavior, is somewhat questionable. Stylized questions can also lack clarity about the inclusiveness of activity categories or descriptions (Gershuny, 2011) and respondents may also be unclear about the specified reference period, particularly in terms of whether successfully recalled events occurred within it. Actual or usual duration questionnaire items do not tell us *when* during the day and the week paid work is undertaken (and therefore lack evidence of atypical or antisocial hours) or the *duration of work spells* (and so lack evidence of work stress – Gershuny, 2011). They also tell us nothing of whether spouses or other household members are simultaneously working or taking leisure (hence they miss evidence of unsociable hours). Yet arguably, the most important impact of paid work changes on well-being relate to exactly these issues of daily and weekly work rhythms (Sunnenberg et al., 2012). In short, stylised questions do not allow researchers to capture the range of contextual detail associated with many activities.

Like with other surveys and questionnaires, a range of potential respondent biases can further threaten the validity of using questionnaires with stylized questions to investigate time use. Firstly, there is the possibility that respondents will be affected by social desirability bias, whereby they give less honest answers as they want to be viewed in a more favourable light (e.g. see Kretuer et al., 2008; Grimm, 2010; Krumpal, 2013). Questions concerning issues like food consumption, exercise, socializing etc., are particularly prone to social desirability bias. Related to this, Gershuny (2011) points out that for some stylized questions (such as those relevant to health-related behavior), respondents may seek to not only mislead others, but also mislead themselves, by upwardly adjusting perceptions of their own participation in healthy behaviours and downwardly adjusting unhealthy ones. Similarly, overestimation bias (i.e. a tendency to overestimate the duration of more desirable activities and underestimate the duration of least-loved activities) can affect the validity of stylized time use responses (e.g. see Otterbach et al., 2010).

Moreover, in surveys using stylized questions, estimates on time use are collected by asking respondents how much time they 'normally' spend on given activities in a certain time interval (Juster et al., 2003). However, no firm conclusions can be drawn regarding the course of the respondents' days or the sequencing of tasks. It also usually remains unclear what is meant exactly by the given activities (e.g. housework) or by the time intervals (e.g. normal day) (Robinson, 1999). As outlined above in section 2, an important application of time use research is to understand the working behavior of a population. It could therefore, be argued that national labour force surveys (LFS) could also be designed (perhaps with the addition of stylized questions) to collect some of this information. However, LFS are a continuous, very large exercise that can produce unreliable estimates of actual work durations, specific hours of work and other work rhythms (Gershuny, 2011). Furthermore, the potential applications of time use data reach much further than employment (e.g. in the measurement of well-being in daily activities).

Time diaries

Szalai (1972) was the first to ask these questions in an international research and policy comparative context a number of conventions about the methodology of time use research and has been referred to as the bible of time use research. Szalai (1972) led the UNESCO-sponsored Multi-national comparative time budget research project, which was the first cross-national time use project to harmonise data collection methods. This project set the blueprint for data collection that, even today, is still used in most time use research. It proposed the utilisation of a time diary to capture all related and contextual activity and time use information during a 24-hour period of a respondents' life.

Consequently, substantial records of the development and collection of representative national, annual, time use diary samples by academic researchers and national statistics institutes are now available for time use researchers (Gershuny, 2011). In particular, two methodological approaches to TUS have now become commonplace. In the first, respondents are asked to log their activities throughout the day in a diary divided into fixed time intervals. The other main method is to conduct a recall interview about yesterday's activities. Hence diary records may be called *time driven* while interviews are *activity driven* (Bittman, 2000). The former TUS data collection method will be discussed in detail in this review, though implementations of the latter will also be outlined due to its utilisation in the American Time Use Survey (ATUS) upon which a great deal of research has been conducted (see section 4.1 for the ATUS).

Unlike survey questions, time use diaries typically collect data for just one or two days (Harvey, 1993). However, more detailed data are captured for this short period of time as respondents are asked to fill in time diaries dividing the day into (usually 15 minute) blocks. Time diaries therefore, acquire information on, not only the amount of time spent on different activities on that day, but also the duration of that activity, the frequency and the sequence of different activities. Apart from this detailed information, primary and secondary activities can also be covered using diary methods. For these reasons and because respondents are requested to fill in the diary immediately when one activity ends and another begins, time diaries are considered by many to be the most exhaustive, as well as the most elaborate methods of surveying peoples time use (Kitterod, 2001; Niemi, 1993; Robinson, 1999). Time diaries are therefore, typically seen as an accurate representation of people's behavior.

While outside the scope of the present review, many authors have charted and outlined the history and development of time use diaries (e.g. Sunnenberg et al., 2012; Gershuny, 2011; Merz, 2009). Researchers use these diaries to record what respondents do during the day, map the entire spectrum of issues relevant to measuring balance of needs in life and for the development of work-life balance policies (Fisher & Layte, 2004). The diary is literally a memory device, making it possible to organize recall around activities whose time structuring function is particularly strong (e.g. the time the train leaves, the time work starts or that the TV show starts). This in turn considerably reduces the overestimation bias affecting stated durations of more desirable activities and the underestimation bias for least loved activities, because the enquiry is not focused on any particular activity and the respondent does not have to calculate activity durations (Chenu & Lesnard, 2006). Questionnaire items concerning food consumption are also unreliable, reflecting the usual problems of recall but also distortions resulting from a potential for socially desirable responses. However, diaries provide plausible evidence of the incidence of each eating episode. They allow us to count meals, whether in private homes or pubs/restaurants, establish their durations, discover their relationship to snacking (visible in diaries where eating appears as a secondary activity) and how they align with eating patterns of other household members (Schulz & Grunow, 2011).

With regard to their superiority over surveys with stylized questions for understanding work specific time use, Merz (2009) points out that the information provided by diaries also allows particular analysis of labour markets which are not available by other labour market surveys (e.g. the sequencing, timing and fragmentation of daily working hour arrangements, which is important for new forms of labour contracts within labour market flexibility developments); unpaid work and nonmarket activities (which illustrate the importance of the informal economy and emphasise women's economic importance and gender approaches in particular) and leisure activities with social networking and volunteer work, family interaction, media use, culture, sports and genuine leisure (these are important in many aspects for the understanding of economic, social individual

and societal living conditions). Bonke (2005) also revealed complex differences between the data obtained by survey questions and time diaries. People who spend many hours on paid work seem to report more working hours when responding to survey questions than time diaries whereas those less engaged in labour market activities tend to underestimate the actual time worked. Other studies have replicated this finding too (e.g. Otterbach & Sousa-Puza, 2010).

Schulz & Grunow (2011) compared time use surveys and diary studies and found that time use estimates may differ noticeably according to the measurement technique used. On the aggregate level of total time for housework, both diary and survey estimates produced significantly different means for several subgroups. The authors therefore, concluded that for obtaining accurate descriptive statistics of absolute time budgets within different populations, there still seems to be no alternative to time diaries. Indeed, many studies indicate that survey questions provide less accurate data than time diaries. In particular, survey questions appear to produce invalid data for less distinct or externally structured activities (Marini & Shelton, 1993; Niemi, 1993; Juster et al., 2003; Schulz & Grunow, 2007). For example, Niemi (1993) found high correspondences of hours reported on paid work, unpaid work and leisure collected by survey questions and by time diaries, while this was not true for occasional or less externally set activities. Specifically, she found that women seem to over-estimate the amount of physical exercise they complete in surveys compared to time diaries.

However, use of the diary methodology in the study of time use is not without its limitations either. Completing the diary requires considerable time-consuming effort, particularly as people are requested to provide this information immediately to avoid retrospective response biases and since time diaries often only cover one day, they typically do not account for time use differences between work days and weekends on an intra-individual level (Sunnenberg et al., 2012). This could have serious implications for the study of well-being in time use research as it might be reasonably be expected to vary significantly between weekdays and weekends (e.g. due to increased family exposure or fewer employment-related commitments on the weekend). It has also been shown that the amount of time reported to be spent on different activities varies according to data collection procedures used in time diaries (Kitterod, 2001; Harvey, 1993). For example, response patterns of main activities differ according to whether or not parallel secondary activities can be reported (Kitterod, 2001) and previous empirical research shows that failure to ask about concurrent activities results in an over estimation of certain activities like housework (Kan, 2007; Kitterod, 2001; Kitterod & Lyngstad, 2005; Robinson, 2005). Some authors consider time diaries to produce systematic incorrect data with regard to unpaid work in comparison to paid market work, as short breaks from work are considered as leisure in unpaid work, while short breaks are incorporated in paid work (Rydenstam, 2001). Therefore, time diaries seem to systematically underestimate hours spent on housework.

Furthermore, time use diary data provide only revealed activity episode patterns. The generation and scheduling process that determines the revealed episode patterns can only be understood if additional data on the internal mechanism leading up to the revealed episode patterns is collected (Bhat & Koppelman, 1999). Such data may be obtained in experimental (but fairly realistic) settings using "think aloud" protocols or by collecting information on intended episode patterns at several points in the day, comparing these patterns with actual revealed patterns and asking respondents to identify reasons for updating/revising their patterns. Moreover, Bhat & Koppelman (1999) also point to a substantial void in our knowledge about the fundamental time use decision mechanism underlying activity episode patterns, which is unlikely to be filled using diary-based research methods alone. Filling this void will provide insight into refining empirical episode generation and scheduling models.

While time use research does and should utilise all kinds of time use data, and both methods have benefits and limitations, the diary is typically the preferred method of sampling, followed by stylized data (Harvey, 1999). Schulz & Grunow (2011) propose that rather than simply applying available survey questions on time use in new studies, it would be worthwhile to invest in obtaining empirical data to clarify why survey and diary estimates differ so markedly, and what can be done to improve the quality and power of both techniques.

Validation of time diaries

Focus groups

Focus groups are a type of group interview. Instead of each member of a group being interviewed in turn however, focus groups capitalize on discussion between research participants to generate data as they are encouraged by an experienced facilitator to talk to each other, ask each other questions, share anecdotes, and comment on each other's' perspectives (Kitzinger, 1995). Therefore, while this is not always necessary (e.g. in the design, development and evaluation of commercial products), depending on the purpose of the focus group, it is usually advisable for the sample to comprise subject matter experts. Focus groups have been described in detail in the qualitative research literature including best practice guidelines concerning their design and implementation. Many research studies concerned with the design/ development (e.g. Foasberg, 2014, Stopher & Wilmot, 2006; Horrigan & Herz, 2004; Schwartz, 2002) and evaluation (e.g. Kenyon, 2004, 2008) of time use diaries have also utilized the focus group methodology in each of these stages, and it would therefore, also be a prudent initial step in the design/development of well-being measures within a diary instrument.

Cognitive interviews

Some researchers have utilised cognitive interview techniques to further validate and refine the design of diaries used in time use research. There are three key assumptions related to surveys and diary studies (and the questions they include). These are that (1) respondents are able to understand questions being asked, (2) Questions are understood in the same way by all respondents and (3) respondents are willing and able to answer such questions.

Tourangeau (1984) developed a simple yet elegant cognitive model of the survey response process that has stood the test of time. It involves comprehension of the question, retrieval of relevant information needed to answer it, a range of judgment or estimation processes that are used to integrate and edit this information and a response process in which the individuals convert their internally constructed representation of the answer to one that constitutes their answer to the question, either in spoken or written form (e.g. saying "yes" rather than providing a conversational report).

From a cognitive perspective, a survey question is a kind of instruction. Respondents are, in effect, instructed to carry out a task (i.e. to do whatever mental work is necessary to provide and answer). From this perspective, they can get into trouble completing this task in several ways. They can (1) misunderstand what they have been asked to do and therefore, carry out the wrong task, (2) understand the instruction, but find themselves unable to carry out the assigned task or (3) carry out the assigned task but find themselves unable to fit their answer in to the options provided. Each of these difficulties demands a different sort of solution. For instance, misunderstanding the task (i.e. 1

above) may be resolved by making the questions and instructions wording more clear to the respondent, task difficulty (i.e. 2 above) may be addressed by simplifying the task or response difficulties (i.e. 3 above) may be resolved by more appropriately matching response options to the way people think about the topic or question.

These problems may come to light through extensive piloting with a large pre-test field sample, but a much more cost-effective methodology for collecting quality information with a small sample of the population of interest has emerged over about the last thirty years called cognitive interviewing. Cognitive Interviews (CI) are a qualitative method for detecting a wide range of potential sources of error including specification error, measurement error related to respondents, non-response error and mode effect (DeMaio & Rothgeb, 1996; Biemer & Lyberg, 2003; Brancato et al., 2006). The main purpose of CI is in reviewing and identifying how an instrument can be improved in terms of response burden and data quality, they rationalize how respondents understood or interpreted questions and how they reached their answers according to stages of Tourangeau's (1984) cognitive model.

CI provide valuable information of the existence of problems in questions, their potential problems and their solutions (Brancato et al., 2006). Although CI only identify problems rather than guaranteeing statistical significance of the results (Willis, 2004), they look in particular at how respondents understand the question and how they arrived at their answer, and their primary goal is to enable researchers to understand the thought processes used to answer survey questions and to use this knowledge to find better ways of constructing, formulating and asking survey questions (Forsyth & Lessler, 1991).

By reducing response error, they can lead to improved reliability and validity of an instrument (Willis, 2004). Particularly important is the instrument's cognitive validity (i.e. the way people process their thoughts, emotions and experiences as they answer survey questions – Karabenick et al., 2007; Wildy & Clarke, 2009). It assesses the degree of consistency between a researcher's intended meaning and the survey user's actual interpretation of a question (Muis et al., 2014). More than twenty years of research into the effectiveness of CI has produced a large volume of experimental evidence suggesting CI are particularly effective in detecting flawed survey questions, identifying response difficulties and providing guidance for repair.

Schwartz (2002) described the use of CI in the development of the American Time Use Survey (ATUS) which began cognitive pretesting in 1998 to understand how respondents understood and interpreted the survey's concepts and questions. Their paper considered CI for development of survey questions specifically but they also wrote about future research in which they planned to apply a similar methodological approach to time use diaries too. The primary objective of this test was to evaluate the overall flow of the interview, paying particular attention to the order in which summary questions are asked. They also stated that they would assess the cognitive difficulty/burden associated with completing the time use interview and respondents' perceptions of the survey topic's intrusiveness.

Park and Park (2014) used CI on a time use survey that included a survey and a time diary in which respondents provided a sequential list of their activities over the previous two days. The time diary also included the duration of each activity and related information like concurrent activities, location and people with whom activities were conducted. They observed prior to CI that the survey previously provided respondents with a considerable amount of burden. They also experimentally reviewed the effect of instructions fidelity on response time and the number of activities recorded

throughout the day. Their results showed that respondents had difficulties understanding terms and where to write answers. They clearly didn't read time diary instructions thoroughly which suggested a need to improve instructions and present them in a more user friendly, perhaps visual way.

Other time use data collection methods

However, despite their prevalence in time use research, diaries and stylized questionnaires are not the only data collection methods utilized in contemporary time use research. One approach (greatly enabled by recent technological developments like smart phone apps) used in the literature is the experience sampling method (ESM) where respondents are prompted at random instants through the day and week, by a signal (usually a beep) from an electronic device, to describe their current activities and affective circumstances (Larson & Czikszentmihalyi, 1983). The London School of Economics (LSE) developed and released the free "mappiness" app for I-phones which uses ESM to measure affect (i.e. well-being) and its associations with various contextual factors (e.g. activities, location, company etc.). It enables users to chart their own happiness over time and in relation to these contextual factors and also provides a rich dataset for researchers to use for their own purposes (see http://www.mappiness.org.uk/).

ESM requires an immediate response from respondents meaning that recall problems can be avoided (i.e. reference period effects are entirely avoided because only current events are reported). This methodological approach also allows for direct measures (potentially including physiological measures too) of affective responses. Gershuny (2011) also points to fewer potential social desirability effects using ESM because honesty is the *easiest* policy for respondents (i.e. actively inventing more desirable responses would require considerable mental work). With reference to Bhat & Koppelman's (1999) criticism of diaries outlined above, ESM can also allow researchers to monitor experiential dimensions of human behavior at the time that this behavior occurs (Zuzanek, 2009). Therefore, this methodology may provide particularly useful insights into some of the underlying mechanisms of activity episode patterns.

Krueger & Schkade (2008, p.2) refer to ESM as the "gold standard" in measuring affective experiences of daily life as it minimizes the role of memory and interpretation. However, there are also many limitations of this approach, not least that it is expensive and very difficult to implement in large samples. Gershuny (2011) points out some additional limitations of ESM. Firstly, this methodology can induce significant respondent burden due to the amount that it intrudes into their normal lives. Furthermore, as ESM is only really acceptable with prompts relatively widely dispersed throughout the day (usually no more than 2-3 per day), this can lead to problems in estimating total elapsed time for activities due to discontinuity of observation. Moreover, as ESM provide no comprehensive coverage of a respondent's time, sequential events (e.g. what happened before or after the reported activity) that may have particular relation to affective responses about the activity, are altogether missing.

Clearly then, all methodological approaches to time use research have their associated benefits and limitations. In general, diaries have become established as the most reliable sources of time use data (Michelson, 2005). They are roughly comparable to ESM (yet much less expensive) and more valid than stylized questions (Bolger et al., 2003). They also provide researchers with the opportunity to collect data about the context of events, the activities that followed or preceded each and even, who was present with the respondent and the location for each event (Harvey & Royal, 2000; Baghal et al., 2014).

National time use research using diaries

In many countries, central statistical agencies conduct recurring time diary studies of their populations (e.g. Australia, Canada, Finland, Germany, Japan, Korea, New Zealand, the Netherlands, Norway, Sweden and many more - Harvey & Pentland, 1999). The American Heritage Time Use Study (AHTUS) compiled, major national samples of time-diary based time use studies in the USA conducted since the 1960's with the purpose of creating historically comparable time use statistics. Similarly, the Multinational Time Use Study (MTUS) is a collection of harmonized time use diary surveys based on nationally representative samples from over 20 countries from the 1960's to present day. Other organsiations have sought to the standardize classification activities (e.g. see UN guidance on International classification of activities in time use statistics – ICATUS – 2016).

Harmonized European Time Use Study (HETUS)

Since the year 2000 the MTUS diary surveys of many European countries have comprised part of the Harmonized European Time Use Study (HETUS). HETUS is a project coordinated by Eurostat with the aim of creating comparable and standardized time use statistics across European countries. The European Union began to support the harmonization of TUS and statistics in Europe in the early 1990's (Rydenstam, 1999). In 1996 and 1997, Eurostat launched a number of pilot studies that resulted in the guidelines on harmonized European TUS (HETUS). These included recommendations on the sample design, diary days, survey forms, activity coding lists, interviewers, data coding and estimators.

The 2000 HETUS guidelines have been the cornerstone of the European time use harmonization process. The Eurostat HETUS guidelines revised in 2008 further proposed a set of recommendations covering all steps and aspects from the design of surveys to the processing of results and recommend activity classification schema (Eurostat, 2009), with particular focus on the production of output-harmonised data. Together with national statistical offices, Eurostat developed guidelines for harmonized European time use surveys (HETUS, Eurostat, 2009). The 2009 documentation is an update subject to two principles: comparability with previous guidelines and simplification. For instance, EUROSTAT recently recommended lowering the age threshold of respondents and this has now been implemented in many European countries (in the UK, Netherland, Norway and Portugal children older than 12 may participate and in Finland the age threshold is 10 years). Zuzanek (2009) argues that inclusion of younger age groups provides information about the entry of teenagers to the labour force as well as the relationship between teens' study loads, sleeping habits, well-being and health - In the US, a National Adolescent Time Use and Risk Behaviour Study found that the time use patterns of 10th graders (i.e. 15-16 year olds) were highly predictive of what they would be doing one year after high school (Zuzanek, 2009).

Policy value of the HETUS surveys arises not just from the capacity to compare time use across European countries at the point of any given HETUS survey wave, but also to assess longer-term behavior change both within each participating country and across European countries (Pilot group 7). At the turn of the millennium around 20 European countries conducted TUS according to HETUS. In 2009 more than 40 international TUS had been conducted, as outlined by the authors. The author's note that the almost exponential increase of new time use studies since 2000 worldwide, emphasizes the internationally recognized importance of time use data for research and policy. A 2010 review by the Bureau of the conference of European statisticians highlighted the need to further improve comparability of TUS across countries and the utility of developing further practical guidelines for their implementation.

American Time Use Survey (ATUS)

The American Time Use Survey (ATUS) has collected a large number of time diaries each month from a sample designed to be broadly representative of the American population aged 15 or older. While the ATUS was first considered as a way to measure and place value on unremunerated work, there are many other potential analytical uses (Frazis & Stewart, 2004; 2007; Horrigan & Herz, 2004; Joyce & Stewart, 1999). Indeed, Abraham et al (2005) point out that the data from this ongoing national survey make it possible to examine the hours devoted to almost any activity a researcher may be interested in (e.g. childcare, household production, education, sleep, aerobic exercise, religious services, work, commuting etc.)

The ATUS methodology requires one person to be randomly selected from each household to participate. The selected person is assigned a specific day of the week about which to report. The ATUS is a computer assisted telephone interview (CATI) that takes on average, about 20 minutes to complete. During the interview, the preselected household member reports about his/her time use during the 24-hour period of the day before the interview. Interviewers use a set of scripted, open ended questions in conjunction with conversational interviewing techniques to walk respondents chronologically through the 24-hour day, collecting information about time spent on activities beginning at 4am the previous day up until 4am of the interview day. For each activity reported, the interviewer asks how long the activity took, recorded either as a duration or with start and stop times. Respondents are also asked questions about who was in the room with them (if at their own or another's home) or who accompanied them (if travelling or away from home) and where each activity took place, which are recorded using pre-coded categories (Phipps & Vernon, 2009). Subjective well-being has also been examined in previous ATUS (e.g. see Gimenez & Molina, 2015).

Other diary-based TUS

In 2000 the UK's first official TUS was carried out. This was a household survey where a household questionnaire was completed and then within each household, individuals aged 8 or over completed an individual questionnaire, two 24-hour diaries and a one-week work and full-time education sheet. Activities in this survey were collected within a slightly adapted version of the HETUS activity framework and the survey itself generally followed HETUS rules and guidelines (see UK 2000 TUS technical report, 2003). While creating a wealth of useful information, a survey of this form placed a heavy burden on respondents and was expensive to carry out. A pre-coded diary was also collected as part of the National Statistics Omnibus survey in 2005 (https://discover.ukdataservice.ac.uk/catalogue/?sn=5995). A 2014-15 UK TUS was also carried out, that was designed to be compatible with both the 2000 survey and other European time use studies carried out since 2008 (https://discover.ukdataservice.ac.uk/catalogue?sn=8128).

The UK Millennium Cohort Survey (MCS) collected time diaries and activity monitors (using accelerometers) from nearly 19,000 young people aged 14 from March 2015-16 (Chazitheochari et al., 2015). Like the HETUS, the MCS collects one weekday and one weekend diary. The survey is a 6-hour format diary format with a fixed range of categories for each of the diary domains. It used app, online and paper diary formats. It ran contemporaneously with the UK 2014-15 HETUS, offering the possibility of comparing the MCS diaries with the roughly 800 HETUS paper diaries for similarly aged people. Merz (2009) and Ricroch (2011) also provide good overviews of the German and French TUS respectively, which also primarily utilized diary-based methodological approaches.

As well as national statistical agencies, private firms also make use of time use information. The Nielsen consumer panel survey samples 300,000 households in 28 countries. It asks about consumption activities (scanned by the respondents using bar codes) and allows for the analysis of individual consumer behavior development over time for many years. In Germany too, the company GfK Global runs consumer scope collecting explicit time use information; with specific studies on gardening, media use that deepen the activity specific time use information collected (Merz, 2009).

What to collect in time use diaries?

Although a primary focus of this and many previous TUS has been activities, the purpose of the present research was to investigate the use of affect ratings as measures of subjective well-being in time diary studies. Therefore, for the purposes of this report, well-being is considered first, followed by consideration of activities and other important factors.

Well-being

The UK government has prioritized the collection of subjective well-being questions and the use of these questions in policy research. A growing recognition of the importance of subjective well-being also recently caused the government of Bhutan to switch from measuring GDP to measuring gross national happiness (Fisher et al., 2016). Understanding human well-being is a core task for both researchers and policy makers (McGillivray & Clarke, 2006). The majority of college students around the world consider happiness and life satisfaction to be extremely important and most respondents believe happiness is more important than money (Diener & Oishi, 1997). Happy people are also judged to have a more desirable life than unhappy people, to be better people and to be more likely to be admitted into heaven (King & Napa, 1998).

Krueger & Schkade (2008) report that from 2000-2006, 157 papers and numerous books were published in the economics literature alone using data on life satisfaction or subjective well-being. Growth in the field of subjective well-being reflects larger societal trends concerning the value of the individual; the importance of subjective views in evaluating life and the recognition that well-being includes positive elements that transcend economic prosperity (Deiner et al., 1999).

Defining well-being

Human well-being is an ambiguous concept that lacks a universally acceptable definition and has numerous and often competing interpretations. It cannot be directly observed so it cannot be directly measured (McGillivray & Clarke, 2006). Like other notions with close meanings (i.e. quality of life, human development) well-being has been the subject of numerous studies and intents to develop measures without really reaching a satisfactory one (Sauvain-Dugerdil, 2006). Terms like quality of life, welfare, well-living, living standards, utility, life satisfaction, prosperity, needs fulfillment, development, empowerment, capability expansion, human development, poverty and happiness are often used interchangeably with well-being without explicit discussion as to their distinctiveness (McGillivray & Clarke, 2006).

Determinants and correlates of well-being

In his broad review of the determinants of subjective well-being, Wilson (1967, p.294) concluded based on limited evidence that the happy person is a "young, healthy, well-educated, well-paid, extroverted, optimistic, worry-free, religious, married person with high self-esteem, job morale, modest aspirations, of either sex and a wide range of intelligence". More recent literature on the determinants of subjective well-being has studied the factors that make individuals happier, with some studies showing that self-esteem (see Deiner et al., 1999), and volunteering are positively related to individual subjective well-being (Binder & Freytag, 2013; Dolan et al., 2008; Meier & Stutzer, 2008). Indeed, using the well-being module of ATUS (2010), Gimenez & Molina (2015) found that those who devote any time to voluntary activities during the day report higher level of happiness than those who do not devote such time.

Harvey (2009) lists a number of things that matter in understanding well-being. These include work-life balance, intensity of social contacts, location of time use, subjective experiences of time (e.g. sense of time pressure) and enjoyment of selected daily activities. Perhaps the activity most generally associated with physical well-being is purposive exercise (Gershuny, 2011). According to this author, well-being is also promoted by, amongst other things, money, income, emotional/sexual satisfaction, an acceptable social/environmental context, and also by the way we make use of our time. Intrinsically satisfying work activities inside and outside the money nexus, healthy and enjoyable consumption and leisure, and an appropriate balance between work and leisure activities contribute to it, as do the economic and cultural resources that provide the prospect of continuing these into the future.

TUS also reveal that levels of subjective well-being correlate negatively with some activities like commuting, and positively with time spent in others (Krueger & Kahneman, 2006); that infrequent communication between spouses is an important predictor of possible family dissolution (Hill, 1988) and that social capital of volunteering and social networking contribute to happy and successful life careers (Zuzanek, 2009; Ravanera, Raiulton & Turcotte, 2003). However, there is not a simple answer as to the causes of subjective well-being. As Deiner et al (1999) explain, studies of religion, rumination, coping and attribution suggest that cognitive factors play an important role; studies of people with disabilities show that objective factors can matter (but people often adapt their goals to what is possible for them); studies of heritability show that personality plays a key role and cross cultural studies reveal that different factors correlate with subjective well-being in different cultures and different variables lead to subjective well-being for people with different values and goals.

According to Deiner et al (1999) the past 30 years have shown that the demographic correlates of subjective well-being do not account for much variance in subjective well-being. This is at least partially due to the fact that the effects of demographic variables are probably mediated by psychological processes such as goals and coping abilities. For example, gender and age may influence goals and objective life circumstances, but there is probably no direct path from these variables to happiness). Additionally, women report approximately the same level of global happiness and life satisfaction as men, but at the same time they are more likely to report greater levels of both positive and negative affect (Deiner et al., 1999).

Also there are many different resources that are related to people's goals and these resources are only modestly correlated with each other (Diener & Fujita, 1995). Therefore, any single resource is unlikely to have a strong effect when analysed across people. Judge (1990) cautions against interpreting correlations between domain satisfaction and global subjective well-being because

satisfaction with even neutral everyday objects may correlate substantially with life satisfaction, because of the influence of personality on these judgements.

Deiner et al (1999) suggest that future research needs to understand the interaction between personality and environmental factors. Although there are many studies examining correlations between personality factors and subjective well-being and many more on the environmental correlates of happiness, researchers are much less sure about the interactions of these two variable types in subjective well-being. More research is also needed on adaptation to understand when it does and does not occur, its limits and the processes underlying it. For example, people who are ill, physically unattractive (Diener, Wolsic & Fujita, 1995) and victims of crime often show average levels of subjective well-being probably reflecting the ability of these people to adapt to many conditions.

From their review of the (more recent) literature on subjective well-being and its correlates, Deiner et al., (1999, p.295) concluded that the happy person "is blessed with a positive temperament, tends to look on the bright side of things, does not ruminate excessively about bad events, is living in an economically developed society, has social confidants and possesses adequate resources for making progress toward valued goals". However, they also acknowledge that due to progress in the area of subjective well-being this description is likely to be over written in the coming decades.

Subjective well-being involves a multidimensional evaluation of life including cognitive judgments of life satisfaction and affective evaluations of emotions and moods (Diener, 1984; Argyle, 1987; Diener & Larsen, 1993; Eid & Diener, 2003). Sauvain-Dugerdil (2006) points to residual time as an indicator of well-being, where the issue is about what is done during this time left for oneself and how it is lived (i.e. time to be VS time to do) but that this is also not easy to measure. It is usually considered through free time as opposed to constraint time (i.e. work/studies and domestic tasks). They also propose that free time should be further disaggregated into leisure and semi-leisure. Semi-leisure includes social and spiritual involvements, whereas in leisure, priority is given to the self: self-rest, entertainment, self-autonomous achievement.

Well-being is clearly a multi-dimensional concept, yet until recently it was considered analogous to income and consumption levels. The authors suggest it is due to reasons of data availability and reliability that income "*continues to be regarded as the quintessential well-being indicator*" (Dasgupta, 2001, p.53). Indeed, Krueger et al. (2009) point out that the general well-being of a society is usually measured in terms of GDP, national income and consumption. This is not a new phenomenon either, it is one that sociologists (e.g. Durkheim, 1899) and economists (e.g. Easterlin, 1974) and the author of the Biblical Book of Proverbs have commented (cited in Gershuny, 2011), bears a complicated relationship to monetary measures, yet as this author argues, the wealth of nations does not necessarily equate to their happiness. In their report on the measurement of economic performance and social progress, the Stiglitz commission argued for a shift of emphasis from measuring economic production to measuring people's well-being (Stiglitz et al., 2009). The commission referred to the consensus that quality of life depends on people's health and education, their everyday activities, their participation in the political process, the social and natural environment in which they live, and the factors shaping their personal and economic security. It identified TUS as a key way to inform on well-being and progress, recognizing that indicators in this area remain deficient.

Measuring well-being

Therefore, to complement these financial measures researchers and policymakers have made some progress into incorporating subjective or affective measures (e.g. trust, happiness) into the discourse regarding societal well-being (Lee et al., 2016). Harvey (2009) identified four major types of time use relevant to the assessment of well-being. These were changing amounts of contracted time (paid work and education), committed time (unpaid work) necessary time (sleep, meals, personal care) and free time.

Although the task of creating a synthetic index of well-being and incorporating it into time use data is admirable, it is also tricky because unlike the relationship between GDP, unemployment rate and performance of the economy, the relationship between time use and well-being is not necessarily linear (Zuzanek, 2009). Too much free time does not make people happier, but neither does a shortage of free time. It is extremely difficult to quantify balanced use of time even though human well-being is predicated on it (Zuzanek, 2009). This author suggests that the key word for optimizing time use is balance, yet we do not seem to know how to measure how much of a given activity is "too much" for different gender, age, life-cycle or socio-occupational groups. From this perspective, Zuzanek argues that subjective measures of "time crunch" or "work-family balance" are important indicators of perceived life-style equilibrium which together with time use data can help us to capture important aspects of personal well-being. Considering their heavy workloads, studies of working families show that work-life imbalance leads to decreased psychological well-being, stemming from increased stress at home and at work (Offer & Schneider, 2011).

Quantifiable measures like real income or earnings have frequently been used to assess quality of life (under the assumption that a higher salary means a better quality of life). However, high salary jobs often imply long working hours and less leisure time and measures that ignore home production and leisure may be misleading. Time use data would allow consideration of an essentially broader set of resources in these measurements (Joyce & Steward, 1999). Recommendation 1 of the *Report by the commission on the measurement of economic performance and social progress (*the Stiglitz-Sen Fitoussi report) suggests that three aspects of subjective well-being (life evaluation, positive and negative emotions) should be collected as part of official statistics. It further recommends that positive and negative emotions (i.e. affect) need to be collected in real time and separately as *"the presence of positive affect does not imply the absence of negative affect"* (Stiglitz et al., 2009, p.146). Krueger et al (2009) also argue that ideally, one would like to know the proportion of daily time that an individual experiences positive/negative emotion, with an assumption being that the higher proportion of positive time experienced, the higher the well-being.

Ordinal well-being measurement scales

Fisher et al (2016) report that many HETUS surveys already ask a range of subjective well-being questions (e.g. how satisfied are you with your life overall, health, overall level of happiness) in individual questionnaires. Some HETUS surveys (e.g. the French, MCS) have used 5-point likert-type scales and 6-point scales, others (e.g. the UK) have used 7-point scales ranging from 1 (completely dissatisfied) to 7 (completely satisfied). Other surveys (e.g. Italy) have used an 11-point scale, ranging from 0 (not at all happy/anxious/satisfied) to 10 (completely happy/anxious/satisfied), and some surveys have used various scales for differing questions.

According to Fisher et al (2016), preliminary research comparing the UK, Italian and French HETUS formats suggest these surveys capture a similar range of emotional responses to diary episodes. No TUS has directly compared 5,7 and 11-point scales or scale order to determine which detail/direction of scale best captures variations in reaction to episodes in the time diary. The French experience also revealed that the horizontal scale allowing participants to see all the options at once was easier to complete and collected fewer errors.

Some HETUS surveys (e.g. the most recent French, Italian and the UK HETUS 2014-15 survey) have included one or two extra columns in the diary instrument. This approach collects emotion ratings throughout the whole day, revealing how emotions change in response to sequences of events and contexts, rather than restricting observations to snapshots.

Single or multiple questions to measure well-being?

Subjective well-being is most commonly measured by asking people a single question like "all things considered, how satisfied are you with your life as a whole these days?" or "taken all together, would you say that you are very happy, pretty happy or not too happy?". Such questions elicit a global evaluation of one's life. Surveys in many countries conducted longitudinally over decades have indicated that, on average, large increases in per capita national income have been found to have little effect on reported global judgments of life satisfaction or happiness over the past 40 years (Krueger & Schkade, 2008).

Since the 1980's time diary respondents have rated subjective well-being during specific activities. Lee et al (2016) report that TUS from Canada, USA and UK have all experimented with asking a single emotion question of all activities. In the 1985 study of American time use individuals rated enjoyment of each reported activity on a 10-point scale. The 1986 UK Unilever Household Research Project asked an enjoyment question for each activity reported in the diary (Gershuny, 2013). In the Canadian Family Time and Activity study, Michelson (1985) asked respondents about the degree of tension experienced in each episode on a 7-point scale ranging from 1 (very tense) to 7 (very relaxed). However, only a limited amount of published research has made any use of these well-being dimensions until recent years (Lee et al., 2016).

Krueger et al (2009) constructed an unpleasant index (i.e. the U-index) which was created by classifying an activity episode as unpleasant if the most intense feeling reported for that activity was negative. However, this scale had questionable validity as it is assumed that a relative ordering on a scale for assessing pain is comparable to the relative ranking on a happiness scale (without any underlying empirical research to establish equivalence – Lee et al., 2016). These authors went on to explain that the U-index was also dichotomous (i.e. pleasant or unpleasant), but that there are many activities which have some elements of both pleasantness and unpleasantness.

Single-item measures of subjective well-being have also been found to have only moderate test-retest reliabilities, usually between 0.4 and 0.66 even when asked twice in the same session one hour apart (Andrews & Whithey, 1976). Kahneman & Flett (1983) found that single-item well-being questions under instructions to consider the "past few weeks" or "these days" had reliabilities of 0.5-0.55 when asked within the same day. The authors note that there has been surprisingly little attention paid to reliability of subjective well-being measures, despite their widespread use.

One reason for the modest reliability of subjective well-being measures (e.g. compared to education and income which typically have reliabilities of around 0.9) could be the susceptibility of subjective

well-being questions to transient mood effects. For example, researchers have documented mood changes in response to subtle events like finding a coin before filling out a questionnaire, the current weather, or even question order, which in turn, influenced reported life satisfaction (e.g. Schwarz, 1987). Eid & Deiner (2004) used a structural model to separate situational variability from random error and basic stability. They found that anywhere from 4% to 25% of the variance in various well-being and satisfaction measures were accounted for by situation-specific factors. Ferring et al (1996) estimated the size of transient factors as between 12% and 34% of total variance.

Krueger & Schkade (2008) propose that reliabilities of subjective well-being measures can be used to determine the sample size needed to detect an expected difference between groups. Because these are modest reliabilities, the risk of incorrectly concluding that groups do not differ is of particular concern. They further suggest that an alternative design approach to larger samples would reduce error by sampling the same people at different points in time, or by selecting adjectives that are less situational and more person-specific.

Fisher et al (2016) outlined three broad methods of collecting well-being data from diaries that have been developed and used in previous TUS: Various regional US TUS included a number of items asking people how much they usually enjoy a range of activities, and at the end of the diary day ask respondents to identify events in their diary day they most and least enjoyed. Some surveys ask many questions after diary completion of randomly selected diary events or time periods of fixed duration. For instance, Kahneman & Krueger (2006) asked 4 negative questions (*how stressed, sad, tired, in pain were you during this time*?) and 2 positive questions (*how happy and meaningful did you find this time*?). These six well-being questions were also added to three ATUS episodes in 2010, 2012 and 2013 as part of the ATUS well-being module. In a recent assessment of the ATUS well-being module, Lee et al (2016) examined responses to these scales using factor analysis. It showed that the five items (without meaningfulness) fit well into one well-being scale explaining 51% of the variance, compared with 43% of the variance using the six-item scale. Robinson (2013) found that the measures in ATUS largely replicated enjoyment ratings in previous surveys, which demonstrates the validity of this well-being scale. Lee et al (2016) conclude that despite their limitations, the well-being items in the well-being module of the ATUS show sufficient reliability and validity and moderate variability, but that their suggestions and findings may not be applicable to well-being data in TUS utilizing different measures for subjective well-being.

Based on their research, Lee et al (2016) don't recommend using individual items. They explain that attempting to obtain separate individual effects while controlling for correlated items will not produce as high an association with the outcome as will combining related items into separate orthogonal scales. For example, when they included each of the emotion items individually in the model, the question about sadness was not as effective as other questions in predicting better health because it was highly correlated (0.6) with stress. As an alternative Krueger & Schkade (2008) highlight the increased reliability of the satisfaction with life scale (SWLS – Deiner et al., 1985) and suggest it is more reliable than single item questions as it is the sum of multiple items, so benefits from error reduction through aggregation.

Well-being outputs from time use research

There are three types of possible outputs from TUS on affect and subjective well-being. These are the mean level of subjective well-being for the population as a whole, the distribution of the mean levels of subjective well-being across people and the distribution of levels of subjective well-being across

activities (UNECE, 2013). The collection of subjective well-being measures from TUS is still in its early stages. Table 1 below shows some potential well-being outputs proposed by UNECE (2013).

Table 1: Showing potential well-being outputs in time use research

Measure	Outputs	Input requirement	Population
Affect/subjective well-being	Mean proportion of time spent in an unpleasant state	Calculated from diary	All persons
	Mean proportion of time spent in an unpleasant state by age group, sex, ethnicity, labour force status, highest qualification attained & household income quantile	Calculated from diary	All persons
	Mean net affect by primary activity group	Calculated from diary as the mean score of positive affect minus the mean score of negative affect states. For surveys measuring enjoyment on a single scale, net affect is the mean scale score for an activity. A normalization will be necessary if different scales are used	All persons
	Mean net affect by presence of others	Calculated from diary as the mean score of positive affect minus the mean score of negative affect states by "who with". For surveys measuring enjoyment on a single scale, net affect is the mean scale score for an activity by "who with". A normalization will be necessary if different scales are used	All persons
Sport and exercise	Average time spent in	Collection of secondary	All persons, and

Measure	Outputs	Input requirement	Population
	sport or exercise as primary or secondary activity	activity (minimum 24 hours diary)	separately for women and men & by age group
	Percentage participating in sport or exercise on any given day	Collection of secondary activity (minimum 24 hours diary)	All persons, and separately for women and men & by age group
Screen time	Average time spent watching TV, other visual electronic media, playing computer games, otherwise using a computer as primary or secondary activity – where the activity does not take place with sport or exercise or a physically active mode of transport at the same time	Collection of secondary activity, a yes/no flag column in the diary marking whether the respondent used the internet or a computer during the activity	All persons
Gender equality – Paid and unpaid work by gender	Average time spent in unpaid work by sex	24-hour diary with adequate sampling on weekends, and "who for" contextual variable, and gender variable	All persons
	Average time spent in total work (paid and unpaid)	24-hour diary with adequate sampling on weekends, and "who for" contextual variable, and gender variable	All persons
	Proportion of the population engaged in	24-hour diary with adequate sampling on	All persons

Measure	Outputs	Input requirement	Population
	paid work by sex	weekends, and "who for" contextual variable, and gender variable	
	Proportion of the population engaged in total work	24-hour diary with adequate sampling on weekends, and "who for" contextual variable, and gender variable	All persons
	Proportion of the population engaged in paid work by time of day and by sex	24-hour diary with adequate sampling on weekends, and "who for" contextual variable, and gender variable	All persons
	Average time spent on domestic activities by sex	24-hour diary with adequate sampling on weekends, and "who for" contextual variable, and gender variable	All persons
	Average time spent on caring activities by sex, for persons who provide childcare or adult care	24-hour diary with adequate sampling on weekends, and "who for" contextual variable, and gender variable	Persons who provided caring activities (child/adult care), persons who undertook (in)/formal voluntary work
	Average time spent on formal voluntary work activities by sex for persons who did voluntary work for an organisation	24-hour diary with adequate sampling on weekends, and "who for" contextual variable, and gender variable	Persons who provided caring activities (child/adult care), persons who undertook (in)/formal voluntary work

Measure	Outputs	Input requirement	Population
	Average time spent on informal voluntary work activities by sex for persons who did voluntary work for an organisation	24-hour diary with adequate sampling on weekends, and "who for" contextual variable, and gender variable	Persons who provided caring activities (child/adult care), persons who undertook (in)/formal voluntary work
	Average time spent on informal voluntary work activities by sex for persons who provided unpaid help/assistance to others	24-hour diary with adequate sampling on weekends, and "who for" contextual variable, and gender variable	Persons who provided caring activities (child/adult care), persons who undertook (in)/formal voluntary work
	Average time spent caring for children as a primary or secondary activity by sex for persons who provided childcare	Collection of primary and secondary activity, "who for" & "who with" contextual variables	Persons who provided childcare
	Average time spent providing care to persons with a disability as a primary or secondary activity by sex for persons who provided care to a person with a disability	Collection of primary and secondary activity, "who for" & "who with" contextual variables, questions collecting information on persons with a disability in household	Persons who provided care to someone with a disability

Measure	Outputs	Input requirement	Population
	Average time spent undertaking informal voluntary work for other outside the household/family by sex for persons who provided informal assistance to other	Collection of primary and secondary activity, "who for" & "who with" contextual variables,	Persons who provided informal assistance to others
	Proportion of total childcare time that is routine, physical, medical or supervisory (as a primary or secondary activity) that is performed by women	Collection of secondary activity, min. 24 hours diary	All persons

Source: UNECE, 2013

Policy implications of studying well-being in time use research

We know a great deal about the effect of policy changes on labour market behavior while relatively little is known about how government policies affect the way people spend their non-working time. Do they engage in more non-market production or do they consume more leisure? (Joyce & Steward, 1999). TUS provide information on the allocation of time to household production of substitutes for market output as well as on the allocation of time to leisure activities. This information is typically missing in other household datasets, but inevitable for showing how well-being depends on consumption and leisure (Apps, 2002).

Fisher et al (2016) highlight some important policy implications of studying well-being in time use research. Firstly, policy research using well-being seeks to promote greater well-being and reduce harm and suffering. Negative daily experiences are associated with negative overall outlook, but may have separate drivers and mitigators. Activity level well-being data inform understanding of what factors in the day make some experiences particularly unpleasant for certain groups and how policy might alter circumstances of daily experience to reduce negativity of these experiences. Secondly, policies which change people's behaviours can sometimes result in unintended consequences (e.g. convincing people to stop smoking may free up the day for other activities like eating too much). These policies may generate a range of outcomes. They may succeed in fostering behavior change, but also incentivise other behaviour change with negative consequences, making the overall effect of the policy worse than doing nothing. They may succeed in fostering behavior change and produce complimentary entailments, making the overall policy a greater success than anticipated. They may

have not effect whatsoever, or they may fail to achieve the desired behavior change, but incentivize other behavior change that has positive policy value. Finally, Fisher et al (2016) point out that emotional responses can represent a significant part of the judgments of the success or failure of a policy.

Future well-being time use research

Deiner et al (1999) argue that future well-being research within the time use paradigm should consider focusing on examining the causal direction or correlates of happiness using sophisticated methodologies, the interaction between internal factors (e.g. personality traits) and external circumstances, and the processes underlying adaptation (considerable adaptation to both good and bad circumstances often occurs, yet the processes responsible for this adaptation are poorly understood). They also recommend that theories be refined to make specific predictions about how input variables differentially influence the components of subjective well-being.

Activity duration

Generally, the minimum unit or "grain" of time in a daily diary is between five and fifteen minutes, which corresponds to one line on a printed form. In weekly diaries, the minimum unit is a quarter of an hour or a half an hour (Chenu & Lesnard, 2006). HETUS guidelines recommend 10-minute time slots (Eurostat, 2001). In time diaries, time length measurements are sensitive to how detailed the descriptions are. Extremely brief diaries filled out in haste by poorly paid interviewers and/or rushed respondents, leave out short episodes or those that depend on more strongly structuring activities. Chenu & Lesnard (2006) report that the breadth of this phenomenon may be measured by comparing the French surveys completed in 1986 and 1998. The level of definition expected of interviewers and respondents decreased with the doubling of the diary "grain" (each line in 1986 represented five minutes, but in 1998 it was 10 minutes, so respondents were asked to describe their day in 144 lines rather than 288). Interviewer task load was lightened. In 1986, interviewers were instructed to revise diaries during their second visit to respondents after they had identified material that might have been left out, such as getting from one place to another, work breaks and other low-visibility episodes. In 1998 the interviewers no longer received this instruction. In 1986 the daily diaries included an average of 27 episodes whereas in 1998 the figure was 20.

Activity data

Self-complete time diaries offer a detailed description of the patterns of individuals' daily activities. Multi-national data of this kind also provides the opportunity to compare the way in which time is divided between different people in different societies (Robinson & Godbey, 1999; Gershuny, 2000). Several authors (e.g. Kitamura et al., 1996; Kraan, 1996; Bhat & Misra, 1998 and Lawson, 1996) have focused on individual time allocation with specific emphasis on the tradeoffs and relationship between in-home and out of home activities. For instance, Kitamura et al. (1996) studied individual participation in, and time allocation to, in-home and out of home discretionary activities. Their results indicated a strong negative effect of being employed and having a long work commute on out of home discretionary activity time. Many researchers are also interested in comparing time use between different groups, such as time spent in childcare between men and women. Time use data collected on an ongoing basis also allows for across time comparisons. Fisher et al (2016) highlight the importance of measuring physical activities too, particularly in relation to health and obesity risk research and also for other policy research (e.g. efficiency of physically active transport or energy use associated with such behaviours).

Activity episode analysis emphasizes the context (why, where, when, with whom, duration, sequence) of activity participation. In fact, many of the recommended outputs from TUS require the collection of contextual information about the activity. This may include information about who was with the respondent during the activity, why the respondent undertook the activity (for whom), the location of the activity or the mode of transport during the activity. The collection of such secondary measures of how people are using their time therefore, becomes an important additional variable.

A common approach to collecting information on a respondent's subjective attitude to different activities in TUS has been to ask questions on the most/least preferred activity at the end of the questionnaire. While these questions may be useful for some specific purposes, they capture a different sort of information to questions on positive/negative emotions in that they focus only on the activities provoking the most extreme responses, they lack information on intensity of feeling and they require a consciously processed judgment by the respondent which may result in different responses to that collected by questions focusing more specifically on emotional state (UNECE-2013).

Classification of diary activities

Though within a time use diary the respondent is characterizing his/her activity in a time spell in his/her own words, only coded activities are available for transcribers and researchers in the data analysis and interpretation phases. Thus an appropriate coding for all thinkable research interests is a challenging task. However, there are international harmonizing approaches to classification like the HETUS project (Eurostat, 2009), the United Nations (Bediako & Vanek, 1999) as well as alternative approaches (Hoffman & Mata, 1999). For example, Bianchi & Wight (2010) produced a taxonomy to classify activities into 7 broad categories: work (income producing activities in the labour market); housework (domestic duties that constitute unpaid work, such as food preparation, house cleaning and maintenance, shopping, laundry, gardening and car repair); childcare (maintenance, supervisory and interactive activities that revolve around the needs of children such as feeding, bathing, putting to sleep, assisting with homework, playing, cuddling and soothing) personal care (self-care like eating and grooming); communication (interactive activities like talking on the phone, conversing and emailing); transportation (commuting, school run etc) and mental labour (various thoughts related to work and family matters such as coordination of schedules and time constraints).

Despite the existence of classification taxonomies and standardized coding schemes, coding diary entries can nevertheless be a complicated task that is often "coloured" by the particular research domain of interest. Chenu & Lesnard (2006, p.2) ask for example whether an episode noted with the phrase "read the bible" be classified as "religion" or "reading"? or is the restaurant bill paid after dinner with friends a food or leisure expenditure? These examples illustrate that the choice of one or another coding attests that classification schemes used are viewing the world through certain "glasses", some are particularly interested in religion, others in literacy etc. Double description coding is also prohibitively unwieldy (Chenu & Lesnard, 2006), due for instance, to the levels of complexity it can add to data analysis. Keeping a record of the designations used by respondents does make it possible to recode later using different classification schemes, and opens up the door to ethnographic analysis, but as these authors point out, it can also lead to identifying respondents by name, which goes against the promise of anonymity usually made to them.

Further complications may also arise in the classification of subjective well-being within diaries. Previous authors have stressed that the extent to which a population engages in non-market

production and leisure needs to be collected in time use research. Sauvain-Dugerdil (2006) suggests that leisure time should be further sub-divided into leisure and semi leisure. By studying the emotional associations with, and implications of, activities, it may be possible to further validate increasingly narrowing classifications of leisure and other activities.

Single or multiple activities?

A lot can be learnt about society by knowing how people spend their time during the typical day. However, inconsistency in the recording of time use, specifically, in how we record details of people's participation in more than one activity at a time (i.e. multi-tasking) may be preventing the full understanding of how people use their time in their everyday lives (Kenyon, 2010). Termed in different ways (e.g. co-occurring activities, simultaneous activities, overlapping activities, concurrent activities, parallel activities, secondary activities and polychromic time use), multi-tasking has long been recognized as important (Ironmonger, 2003). In his 1960's study, Szalai (1972) recognized that accounting for multi-tasking gave a fuller picture of the reality of everyday behavior. Indeed, as outlined by Kenyon (2010), multi-tasking data reveal that activity participation is not a zero-sum game, in which the addition of one activity requires subtraction of another: activities can be added without taking any away. However, it is only recently that some, but by no means all, time use researchers have started to record such data.

Defining multi-tasking

Despite the recognized importance of multi-tasking data in time use studies, there is an absence of a universally agreed upon definition of multi-tasking. Mostly, definitions are not even offered in the literature, but when they are, they differ, meaning that survey respondents are left to define multi-tasking themselves (Kenyon, 2010). This author points to this as a serious methodological flaw in time use research as when different studies do define multi-tasking but in different ways, this reduces the extent to which we can make informed comparisons between them. Further, when different respondents define multi-tasking in different ways, this reduces the extent we can compare different peoples' time use, therefore reducing the reliability of any findings. How we record multi-tasking influences conclusions about its prevalence and importance, and how multi-tasking is defined influences how it is recorded. Therefore, questioning what we mean by multi-tasking is of vital importance to our understanding of time use.

Drago (2011) points out that it is not obvious how much time spent in secondary activities should be counted because reported secondary activities can be either true simultaneous activities or short duration sequential activities that are reported as secondary. For example, truly simultaneous activities are "preparing dinner" while "listening to the radio" whereas short duration sequential activity might be "getting the mail" while "watching TV".

In her survey of UK internet users, Kenyon (2010) uncovered five conceptualisations of multi-tasking raised by participants that included consecutive multi-tasking (two distinct activities which may occur in the same block of time but which clearly occur at different times); simultaneous multi-tasking (the simultaneous conduct of two distinct activities each undertaken with a separate purpose); active multi-tasking (active participation in multiple acts) and on-call multi-tasking (combines two or more activities when presence is required yet participation cannot be considered to be truly active, such as childcare) and absent multi-tasking (doing two activities but only focused on one).

Robinson & Godbey (1997) conceptualise multi-tasking as the deepening of time, although Kenyon (2010) suggests that it may be more useful to consider it the broadening of time. Questions have been raised as to the extent to which humans are able to multi-task, it has instead been suggested that simultaneous multi-tasking is actually rapid consecutive multi-tasking at a timescale too small to be recorded in surveys, or perhaps even to be observed (Kenyon, 2010).

Lindquist & Kaufman-Scarborough (2007, p.269) provided the following definition of multi-tasking (which they termed polychronicity) "*a form of behavior wherein a person engages in two or more activities during the same block of time, while monochronicity occurs when a person engages in one activity at a time*". They provided examples of polychronicity that included "telecommuting while the clothes dryer is going and dinner is in the oven" and "the grocery shopper who is having photos processed, sushi made, and is calling mom on the cell phone concerning a greeting card process, all during the same block of time".

Multi-tasking requires complex mental activities like planning, prioritizing and error monitoring and is crucial for performance of daily activities (Burgess, 2000; Burgess et al., 2000; Burgess et al., 2008). Without the ability to multi-task, "*one would have to always finish one task (e.g. cooking the vegetables for a meal) before starting another (e.g. cooking the other parts of the meal)*" (Burgess et al., 2008; p.243).

Neuropsychological experiments show that multi-tasking is however, often an inefficient way to accomplish tasks (Rogers & Monsell, 1995) and that the amount of time lost and the likelihood of committing errors while switching repeatedly between two or more tasks increases as a function of the tasks' complexity (Rubenstein, Meyer & Evans, 2011). Overall, people can easily perform automatic and routine tasks together with more complicated tasks such as eating and listening to music, but when tasks require conscious attention and planning, efficiency likely decreases significantly. These findings raise potential concerns about the consequences of multi-tasking for individuals' well-being and functioning (Healy, 2004; Javid & Varney, 2007; Wallis, 2006).

As outlined above, much of the literature uses the term multi-tasking to refer to situations in which people perform secondary activities alongside primary activities. However, in this report, the term "co-occurring activities" will be used in its place (except in those cases where the term multi-tasking was specifically referred to in the literature) because the diary study utilised specific-length time slots to record activity data, within which, it is was impossible during the analysis phase to determine whether primary and secondary activities were undertaken at the same time or sequentially yet within the 10-minute time slot.

Why collect multi-tasking data?

Drago & Stewart (2010) list a number of reasons that researchers might be interested in multi-tasking. There are research communities interested in studying subjective well-being, work-life balance and the extent to which people, especially women, multi-task to get more out of their day; other researchers want to examine secondary activities to gain a more complete picture of childcare (as much of this is done as a secondary activity). Similarly, some researchers wish to measure household production and would want to include household work that is done as a secondary activity. Moreover, for a whole range of important research questions, it is necessary to capture all episodes of a particular activity.

Multi-tasking in childcare and housework

These issues relate to an important reason that time use data is collected in the first place (i.e. to learn more about the time people spend doing unpaid work). Unpaid work refers to activities one could hire someone else to do rather than doing oneself and can be measured directly as a primary activity in any time use data set. Unpaid work is typically captured in two time use categories: childcare and household activities. Childcare tends to account for most of the time devoted to unpaid work as a secondary activity. By allowing parents to simultaneously attend to multiple obligations, multi-tasking can create a sense of greater time, that is, it can deepen the intensity of time, as well as maximize efficiency (Bianchi et al., 2006; Sayer, 2007). Research shows that in dual-earner families, mothers combine housework (Lee & Waite, 2005; Sayer, 2007; Sayer et al., 2009) and childcare (Bianchi et al., 2006; Craig, 2006, 2007; Craig et al., 2010; Ironmonger, 2004; Sayer, 2007; Zick & Bryant, 1996) with other activities more frequently than fathers, and that these trends continue to exacerbate the unequal division of labour between mothers and fathers. In the 2000 National Survey of Parents, 67% of married mothers, but only 42% of married fathers indicated that they multi-task "most of the time" (Bianchi et al., 2006). Percentages are substantially larger among dual earner families where both parents work 50 hours a week or more (85.9% of women and 59.1% of men report frequently multi-tasking (Bianchi & Wight, 2010). As an appropriate balance between work and leisure activities contribute to well-being it is important to understand the implications of multitasking on this complex phenomenon. The heavy workloads of working parents and stresses at home (that frequently housework and childcare) that are often associated with multitasking have previously been shown to have a detrimental effect of subjective well-being (Offer & Schneider, 2011).

Previous empirical research also shows that failure to ask about concurrent activities tends to results in an over estimation of certain activities like housework (Kan, 2007; Kitterod, 2001; Kitterod & Lyngstad, 2005; Robinson, 2005). Drago (2011) reported that most secondary activities occur in the home, and women are more likely to report them, which together reflect the messiness of home life and the prominent role that women continue to play in the home.

Should we even collect multi-tasking data?

Traditionally time use diaries have asked respondents "what they were doing?" and if they "were doing anything else?" during various portions of a 24-hour day (e.g. Bianchi et al., 2006). In practice, secondary activity data in a lot of time use research are usually not used in time use analysis (e.g. see Bianchi et al., 2006). When they are analysed, researchers typically take the naïve approach of dividing the time spent in the episode between the two activities. So, a person who is eating and watching television would be classified as having spent half an hour watching TV and half an hour eating (Drago, 2011).

A small number of authors however, have also expressed caution in collecting multi-tasking data. Drago (2011) warns that a significant danger in changing the ATUS to systematically collect secondary activities is that it may change the way primary activities are reported. There is research too, supporting this notion. Kitterod (2001) analysed two days of Norweigian time use diaries where secondary activity reports were explicitly requested of respondents only on the second day. She found differences in the pattern of primary activity reports across the two days and suggests the data are not comparable. Drago & Stewart (2010) examined secondary activities in ATUS data (i.e. data that was transcribed but not previously coded). They found that secondary activity reports in ATUS were not comparable to other TUS that explicitly ask for information on secondary activities.

However, arguing in favour of collecting multi-tasking data in 2010, Kitterod suggested that surveys collecting limited or no secondary activities also collect fewer episodes, as well as more limited detail associated with "indistinct and fuzzy activities" such as watching TV, using social media, social activities and childcare. The research here is based on the older style of diary collection, which may not be as relevant as collection for time use moves to app and online.

Multi-tasking prevalence

Other studies too, suggest that multi-tasking is highly prevalent. Around 95% of the population report multi-tasking each day (Hungerford, 2001). In the UK 2014-2015 survey (see Gershuny & Sullivan, 2017), 92.8% of diaries contained at least one secondary activity in at least one episode (Fisher et al., 2016). People participate in more than one activity concurrently for approximately one third of the day (Bittman & Wajcman, 2009; Floro & Miles, 2001; Hungerford, 2001; Ruuskanene, 2004) such that multi-tasking can add up to seven hours to the average working day (Kenyon, 2008). Multi-tasked activities are not trivial activities either, but those that impact upon quality of life and life chances, such as childcare and other caring activities (Budig & Folbre, 2004; Ironmonger, 2003); domestic work (Bittman & Wajcman, 2000; Sullivan, 1997); passive leisure (Baron, 2008); communications activities (Baron, 2008); social network activities (Kenyon, 2008); and online activities (Kenyon, 2008). Furthermore, evidence suggests that these activities are more likely to be recorded as secondary than as primary activities. They are therefore, under reported when only primary activities are considered. Thus, it can be considered that the failure to recognize multi-tasking has distorted the picture of popular time use devoted to these activities, leading to an inaccurate account of the amount of time that people spend on these activities.

There is also evidence that highly educated respondents are more likely to report secondary activities. When secondary activities are reported the most common activities are socializing, relaxing and leisure, eating and drinking, and household activities. However, respondents who work longer hours are less likely to report secondary activities, while those over the age of 70 are more likely (Drago, 2011). Furthermore, time use studies indicate only a small gender gap in multi-tasking. For example, Bianchi et al (2006) showed that the number of multi-tasking hours per week is almost identical among married mothers (80 hours) and fathers (78 hours). However, studies focusing on specific activities (e.g. housework and childcare) point to qualitative differences in experiences of multi-tasking by gender, and accounting for secondary activities widens the gender gap in time spent on unpaid work (Craig, 2006, 2007; Ironmonger, 2004; Lee & Waite, 2005; Sayer, 2007).

Accounting for multi-tasking has implications for the understanding of well-being, inequality and disadvantage in society. Inclination to multi-task has been linked to demographic factors including age, culture, educational attainment, employment status, gender, presence of children, and income (Floro & Miles, 2003) and some studies also indicate that multi-tasking may be becoming more prevalent in response to social changes (see Kenyon, 2010) and therefore, more important to our understanding of time use.

Based on their work, Drago & Stewart (2010) recommend that regardless of whether secondary activities are systematically collected, it is important to ask respondents to break apart activities. Such requests make respondents more likely to report short-duration activities as primary rather than secondary. Furthermore, if secondary activities are requested, it is important to collect information on their duration. For instance, results from the 2006 & 2007 ATUS "eating and health" module clearly showed that there is a potential to grossly over estimate time spent in secondary activities unless data on the duration of the secondary activity are collected separately.

What other information to collect?

Other participants in activity

The HETUS surveys collect limited grids with simplified categories of who else was present during activities to minimize respondent burden while ensuring that diaries capture time respondents are with their parents, household children too young to complete their own diaries, other household members, with other people, or alone. This is the most varied element across HETUS diary instruments. However, (Fisher et al., 2016) point out an important limitation with this approach in that it is effectively saying to people not in couples and do not live with children/parents that the survey team does not care about them or who they spend their time with, as contact with significant others who influence daily life (e.g. close friends, carers, non-resident parents) is effectively obscured.

The ATUS collects a more detailed range of who else was present (some participants register up to 20 people present during activities). Other studies using app-based or online data collection methods (e.g. Happiness Research Organisation, MOTUS) have also suggested that people will add larger numbers of people with specified detail without impacting on response rates or data quality, but these surveys aren't limited to paper only surveys (Fisher et al., 2016). There are many reasons for understanding who else was present during diary-reported activities. For a start, rich contextual detail about activities, like this, could have significant implications in understanding the associations between activities, the contexts in which they take place and subjective well-being. Previous research has also outlined the importance of others in general (Krueger & Kahneman, 2006) and of significant others (Hill, 1988) for instance, as correlates and predictors respectively, of subjective well-being. Fisher et al (2016) proposed further reasons why who else was present should be considered in all future time use research: Firstly, understanding care relationships (e.g. parenting, looking after people with disabilities and aging populations) requires more detailed information on who does what and for whom. Interactions with others are also gendered and additional research on who else was present enhances gender research and gender equality promotion. They also highlight other surveys like ATUS which capture a wide range of who else was present information (including household members as well as a wide range of other categories of people living in other households). Finally, they propose such contextual information could be extended by also considering time people spend with animals and other people online. The authors conclude that capturing this data would result in more policy research using this particular element of the data.

Other devices

Watching TV is the most frequently cited activity modified by online or web-enabled devices. This is not particularly surprising as so many programmes offer supplementary content online and people often share their viewing experiences with others using social media (Fisher et al., 2016) Such findings will be of particular interest to researchers interested in media use.

Since the first round, HETUS survey classification codes have distinguished some activities that take place on the internet and smart devices from off-line activities. In the second round of HETUS, statistics delegates discussed adding a separate tick box column for activities taking place at least part on-line or on smart web-enabled devices. The 2014-2015 UK HETUS survey (see Gershuny & Sullivan, 2017), implemented this idea in which participants were asked "did you use a smartphone,

tablet or computer?" Adding this column appears to have had no effect on response burden or response rates. Pretesting of the instrument also identified no comprehension issues, concerns noted in focus groups or clear impact on completion time (Fisher et al., 2016).

However, there are also ambiguities with this approach on the UK version as it is not clear whether ticking the device-use box during an online activity (like updating a website while listening to a podcast) means using the same device or multiple devices at the same time (Fisher et al., 2016). These authors refer to an unreferenced comment Klas Rydenstam of Statistics Sweden in which he noted a further potential complication with marking internet-based and computer-based activities: people new to the internet may be more likely to recognize and report online activities as taking place online than those more "old hand" at this for whom "check bank balance" may mean checking on an app, but who would consider this detail so obvious that they would not report it.

Methodological issues in time use research

Response rate in TUS

Abraham, Maitland & Bianchi (2005) found a number of similar factors were related to the probability of participating in the ATUS survey, such as age, race, education, marital status, hours worked by the respondent and his/her spouse, ages of children in the household, presence of other adults in the household, home ownership and urbanicity. However, when Abraham et al., constructed new weights based on response propensities and compared the time use estimates to those produced using ATUS standard weights, there were only minimal differences. These findings suggested that despite nonrandom non-response, the estimates showed little evidence of non-response bias.

Phipps & Vernon (2009) explain that features of the sampling design contribute to non-response. The ATUS randomly select one person from a household and preselects a day of the week for which that person is to report. This design ensures the sample is representative and all days of the week are equally represented. However, if that person is unavailable then an alternative person in the household cannot complete the survey as a proxy respondent. Also, when the designated person is busy on the assigned reporting day, they might be more difficult to contact or may refuse to participate.

Relatively low response rates of 56%-58% have raised concerns there might be bias in time use estimates. O'Neill & Dixon (2005) found that several demographic characteristics including age, race, marital status and presence of relatives in the household were related to the probability of being unable to contact a sample member, or for the sample member to directly refuse to participate, but their findings showed little evidence of bias across time use categories. The category of personal care (sleeping, showering, grooming and personal/private activities) showed the biggest evidence of bias, but the difference was still relatively small (i.e. 12 minutes out of a total of 12 hours).

Speaking from general experience in this field, Fisher et al (2016) explain that the more complex an instrument becomes, the more likely people are to refuse to participate. Even when people supply a detailed account of their activities, the level of non-response in additional fields can be high. Glorieux & Minnen (2009) examined the effects of using longer diaries and concluded that longer periods of observation clearly lead to better data and more accurate estimates. However, they also concluded that the issue of non-response may be a significant issue in determining whether to use longer diaries.

Abraham et al (2005) approached the issue of survey non-response by discussing two alternative hypotheses about the nature of non-response to be expected in a time diary study like ATUS. The hypotheses were:

1. That busy people are less likely to be among the respondents because they are less frequently at home and, if contacted, less willing to take the time to respond to requests for participation in surveys VS

2. That people with weaker social ties are less likely to be interviewed – i.e. a persons' response propensity reflects the degree to which he/she is socially integrated, or, put differently, the degree of attachment to the broader community

If hypothesis 1 were true this would be a particular problem for a time diary study like ATUS, since it is precisely the use of time that such studies are designed to measure, and the under-representation of busy people could seriously distort the estimates produced (Abraham & Mackie, 2005). Regarding hypothesis 2, difficulties in contacting people who are less socially integrated may arise because they move away or do not have valid phone numbers etc, in addition to their perhaps being less likely to be at home. If people with strong social connections spend their time differently than people with weak social connections, differences in response propensities between the two groups could lead to bias in aggregate time use estimates.

Their analyses indicated that difficulties in contacting sample members accounted for a higher percentage of non-response than outright refusals to participate. Noncontact explained about 60% of non-response to the ATUS. They note that noncontact in the ATUS is noteworthy not only because it is so high, but also because the personal and household characteristics of those who are contacted differ systematically from the characteristics of those who are not. Even after controlling for demographic characteristics such as age, sex, race, education that are taken into account in standard weighting adjustments, they found that people whose characteristics suggest weaker ties to the broader community are more difficult to contact than those whose characteristics suggest a greater degree of social integration. They concluded that although even random non-response reduces effective sample sizes and raises standard errors, nonrandom non-response is much more worrisome because it creates the potential for bias in survey estimates.

Response rate in well-being related TUS

Some studies have examined the effects of well-being questions on response rates. In 2009 the ATUS the well-being module was not included and the response rate was 56.6%, it slightly increased to 56.9% in the first year the well-being module was included but in subsequent years that the well-being module was included, the ATUS response rate dropped to 52.2% in 2012 and 49.9% in 2013, and raised slightly to 51%in 2014 when the well-being module was not included (research cited in Fisher et al., 2016, p. 9).

However, Fisher et al (2016) reported that the US Bureau of Labour Statistics (BLS) found no effect of well-being supplement on response rates. The French National Institute of Statistics and Economic Studies (INSEE) included the enjoyment column in a subsample of their survey and were therefore, able to compare response rates across the same survey, and they found it was not associated with a drop in response (OECD, 2013). Furthermore, anecdotal evidence from the 2015 UK Millennium Cohort surveys suggests that including an enjoyment column actually helped response rates, with young people expressing particular interest in it and reporting finding it particularly meaningful (research reported in Fisher et al (2016). Similarly, early analysis of the French HETUS suggests that

respondents are at least as likely to return completed enjoyment columns as they are to answer other context columns (with response in enjoyment fields sometimes being higher than in other context fields). Moreover, the UK 2014-2015 HETUS surveys had to extend their fieldwork collection period and take other measures to increase overall response. Also, the UK followed the French diary example and added an enjoyment field in a subsample of diaries in the first few months of data collection, and interviewers reported they found sampled household members showed far more interest in the survey when they completed those with enjoyment fields.

In the second round of HETUS, statistics delegates discussed adding a separate tick box column for activities taking place at least part on-line or on smart web-enabled devices. The 2014-2015 UK HETUS survey implemented this idea in which participants were asked "*did you use a smartphone, tablet or computer?*" Adding this column appears to have had no effect on response burden or response rates (Fisher et al., 2016).

The future of the diary method in TUS

The use of GPS tracking devices alongside the collection of time use (and related travel) surveys is now well established (Kracht, 2006; Krenn et al., 2011). Initial work using time diaries completed by people who also wore accelerometers also suggests that these independent measures compliment as well as expand the range of details that time diaries collect (Kelly et al., 2015; Fisher et al., 2015).

Very little published methodological work has analysed the effectiveness of app and online diaries, and many of the formal national sample surveys are recent with some still in field. However, the most recent HETUS survey in Denmark (2008-2009) offered respondents the option to complete some survey instruments online. While relatively few took the online option, automated checks and prompts in the online instruments produced higher quality data with fewer errors (Bonke & Fallesen, 2010). The Modular online time use survey (MOTUS) collects an online diary from a sample of Belgian adults which has a format suitable for small and large screens and covers the same domains currently in HETUS (Minnen et al., 2014). 2014 MOTUS data are close in time to the 2013 Belgian HETUS survey, allowing direct comparison of outcomes of collecting data via both modes. The project is ongoing and offers modules readily adapted for future HETUS pilot tests. Also, the UK millennium cohort survey (MCS) is collecting time diaries and activity monitors (using accelerometers) from nearly 19,000 young people aged 14 from March 2015-16 (Chazitheochari et al., 2015). Like the HETUS the MCS collects one weekday and one weekend diary. The survey is a 6-hour format diary format with a fixed range of categories for each of the diary domains. It used app, online and paper diary formats. It runs contemporaneously with the UK 2014-15 HETUS, offering the possibility of comparing the MCS diaries with the roughly 800 HETUS paper diaries for similarly aged people.

Fisher et al (2016) identified several important issues concerning the eventual transformation of diaries to a more online format. One such issue was that the small screen of a tablet, smartphone, smart watch and other smart devices will require reformatting of the diary, meaning that respondents will see the unfolding of their account in a different way than the paper or laptop or larger size diary screen will offer.

Fisher et al (2016) suggest that for all sections of the survey, rigorous comparability and quality checks will need to consider the time required to complete each instrument, response rates by section and the basic profile of respondents who complete the whole survey and each major

section/instrument (age group, sex, education, employment status, region of country, live in household with young children). These and other issues are currently receiving detailed consideration by the Task force for the Modernisation of Time Use Statistics at Eurostat.

Conclusion

This literature review has illustrated the utility of time use research in the collection of valuable information about a population's working activities, other activities, health behaviours, transportation patterns and a wealth of other information. The implications of time use research for social policy as well as for understanding individual and societal well-being have also been discussed. It has outlined a variety of methods for collecting time use data including surveys with stylized questions, time diary studies, and experience sampling methods and has shown that most authors and researchers consider the time diary to be the most appropriate method of collecting this data. This review went on to introduce a range of national and international time use studies conducted by various statistical agencies and private companies as these frame the context within which most time use research studies have been conducted. Following this, the literature review considered the range of information that must be collected to permit the comprehensive analysis of time use data including activity data, activity duration, other participants in activities, devices used in activities and multi-task or co-occurring activity information. The review went on to discuss the importance of these factors in understanding well-being, and in the collection of well-being information itself. It showed that well-being can be measured using ordinal scales and it would be most appropriate to measure this phenomenon using multiple questions.

Furthermore, this review has illustrated the implications of studying well-being for informing social policy and that its measurement has minimal negative (and even some positive) implications for respondent burden and response rate. This literature review also highlights the importance of collecting well-being information (in light of various approaches like HETUS to harmonise time use data collection methods) for making meaningful comparisons over time within and between countries, populations and different types of people in terms of their activity participation and the well-being derived from this. Finally, the review considers the potential future of the diary method in time use research, the various forms this might take and the research questions that will need to be addressed in validating and using such methods. This review implies a fruitful future for this methodology in terms of the information it can provide to inform social policy as well as in drawing conclusions for, and making comparisons between, the well-being of societies, populations and individuals throughout the world.

Expert Panel

To ensure the study was well-informed, a range of opinions were gathered from various experts around the world, in the fields of time use, survey design and well-being. Conversations with these experts concerned well-being measurement issues (e.g. wording, scale characteristics, response characteristics) and the range of issues associated with capturing activities in sufficient detail to enable useful and meaningful analysis within a well-being context. At each stage of the project the experts were asked for their opinion on the proposals (e.g. reviewing the proposed scale and the diary design).

Focus group study

Introduction

The present research was commissioned to investigate the implications of measuring well-being scales as part of a time use diary study. Although other previous research has also considered time diaries and well-being, including those directly relevant to the HETUS instrument, it was considered important in the design of this research to begin by discussing the construction of diary study instruments that incorporate well-being measurement, with experts in the relevant fields.

It was decided that for this initial study, a focus group would be the most appropriate methodological approach. A focus group would ensure wide-ranging coverage of the relevant issues while also facilitating rich and constructive discussion from a diverse range of perspectives associated with the diary design. The semi-structured nature of this approach allowed for the research team to be exposed to varying (potentially novel) viewpoints, while also (through an experienced facilitator) ensuring sufficient coverage of the pre-defined discussion areas.

The pre-defined discussion areas were determined through examination of the relevant academic literature, proposals and recommendations from time use subject matter experts (e.g. Fisher et al., 2016) and preliminary conversations with other experts in the fields of time use, survey design and well-being (see expert panel). They primarily concerned well-being measurement issues (e.g. wording, scale characteristics, response characteristics) and the range of issues associated with capturing activities in sufficient detail to enable useful and meaningful analysis within a well-being context.

Sample

The participants of the focus group comprised 16 individuals; 4 male and 12 female. Sampling was purposeful, so that researchers had expertise in the area of research being conducted as they were drawn from the professional fields of time use research, well-being, survey design and methodology, and data analysis. These participants had particular expertise in time use measurement, well-being scales, the development of respondent materials and survey question wording.

Data collection instruments

The focus group was presented with a range of materials that included the different measurement scales used to capture the quality of time spent on an activity, the various options for the specific wording of particular questions and a table contained in the previous UK HETUS Time Use diary that demonstrated the different codes for the activity groups.

The measurement scales consisted of two uni-polar scales (i.e. those with either solely positive or solely negative scale points – with the exception of zero). These were a 5-point scale and a 7-point scale, ranging from 'not at all' to 'very much', where the mid-points were 3 and 4 respectively (see Appendix A). The participants were then presented with two bi-polar scales (i.e. those containing negative and positive scale points). These were a 5-point scale and a 7-point scale, with 0 being the mid-point for both scales (i.e. the first scale started on -3 and the second started on -2) (see Appendix B).

The focus group participants were presented with 3 variations of a question, used to measure well-being within the context of time perception. The discussion surrounded the words 'pleasant/unpleasant', 'enjoy' and 'value' (e.g. 'How much did you enjoy your time?' - see Appendix C). Finally, the participants were presented with a table containing the codes for specific activity groups, along with examples demonstrating activities that would fall within that group (see Appendix D).

Procedure

A focus group was selected to ensure that each participant had the opportunity to deliver their thoughts on the materials that would be presented to respondents, to explore individual attitudes on the way in which the quality of time is captured, and thus provide the research team with a detailed consultation surrounding these issues. The focus group lasted for 90 minutes and was held in a convenient location and time for the researchers. The discussions were recorded and fully transcribed.

The focus group was facilitated by a moderator trained in cognitive interviewing and assisted by a member of the research team. The moderator used a semi-structured interview approach and encouraged discussion between respondents. The moderator used a series of open-ended questions about scale lengths, well-being activity question wording, scale types, and scale appearance. Participants were encouraged to explore and offer opinions on all the presented scales and what they felt would bring the most value to the study.

To start the session the moderator outlined the purpose of the research to the respondents and discussed its aim to inform the Harmonised European Time Use Survey (HETUS). Each respondent was provided with an excerpt from a diary to fill in to familiarise themselves with the context within which, the discussion could follow.

The moderator then asked respondents a series of questions on each of the following aspects of the diaries:
- The length of the well-being question scales
- The use of bi-polar versus uni-polar scales
- The wording of the well-being rating question
- The way the respondent was required to respond to the well-being question i.e. writing numbers in or circling numbers printed in the diary
- Types of activities coded and whether these could be captured by the various question wording options

For each aspect discussed, the moderator presented the respondents with examples to aid discussion (see appendix A - D).

Participants of the group were encouraged to give their thoughts on proposed scales and different variations of question wordings, before making a decision on what to implement in the diary. The group also identified whether the activities presented in the UK Time Use Survey were current and thus could be captured appropriately by the well-being questions in the diary.

Participants also reviewed the process for selecting answers, and thus considered the level of difficulty in responding, as well as considering potential interviewer effects and societal expectations. Each discussion centered on gaining an understanding of individual questions and

terminology, and deliberating on alternative terminology for questions that were harder to interpret and understand.

Results

Well-being question wording

All members of the group agreed that 'enjoyable' was the most effective way of measuring a person's quality of time and is easier to rate than the word 'pleasant'. Participants expressed their concerns surrounding the impact events earlier in the day would have on the ratings given to activities that occur later that day. They felt that the same event may be scored differently on different days depending on the activities that had occurred. The group also felt that the word 'value' related to a task that one 'had to do'. They felt that the term was ambiguous and potentially had many interpretations, but was mostly associated with financial and moral values. It was decided that the scales of 'enjoyment' and 'value' were not equivalent, but both could be used to get an accurate depiction of time, although cannot be rated in the same way. The group suggested using 'satisfaction' instead of 'value', which would relate to the standard subjective well-being question, and thus concluded that these words would best capture quality of time. Contrary to this, a well-being expert felt that having both 'value' and 'enjoyment' scales would be a better way of identifying attitudes and thoughts towards a period of time.

Rating scale appearance

Focus group respondents asserted that bi-polar scales should not be used, as visually they add confusion. In particular, the negative scale numbers could result in respondents unconsciously altering the way they interpret the scale, where respondents assume that the negative numbers are associated with negative activities undertaken rather than simply the lower scale points are lower than others. When comparing the scales for each question a small number of participants preferred the 5-point scale over the 7 and 10-point scales, as it gave a more accurate indication of attitudes towards the time slot of an activity, despite it being harder to choose a rating. Participants expressed that they were more likely to score in the middle of a 7-point scale, but not as likely in a 5-point scale. Members of the group seemed to understand the individual connotations of the end points of the 5-point scale, explaining that 3 would be the mid-point thus indicating indifference, 5 reflected large levels of enjoyment and 1 meant no enjoyment. However, respondents felt that longer scales (i.e. 1-7 or 0-10) provided more flexibility in responses. The majority of the group preferred the 7-point scale compared to the 5-point scale, as it showed greater differentiation, but not too much of a burden to reflect the different answers.

Many participants preferred a scale that offered them the choice of circling a number, rather than writing it themselves. After discussing their experiences of filling in the scale, participants agreed that if respondents had to write their rating in a box they would be more likely to select midpoints or the highest numbers on the scale, for example on a 7-point scale 3s and 7s. It was therefore asserted that there would be greater variation in responses if people were to circle their answers on a scale, as it is easier to visualize a rating and compare it to previous activity rating. Additionally, this would be more accessible for respondents with varying levels of numeracy.

Recall and mood effects

Participants experienced some difficulty recalling and rating their mood during an activity on that particular day, when compared to rating their mood at the point of recording their daily activities, especially when the two time points were different. Participants expressed that recalled mood was almost always affected by one's current mood. It was found that using 'enjoyment' as an indicator of well-being was likely to be influenced by events that had occurred during that day, and that activities assigned a particular level of enjoyment on one day may be different on a different day. Participants were also tempted to change their rating of a certain activity after considering their rating of another activity; it was evident that previous scores were used as an anchor for each activity rating.

Recommendations for testing in the diaries

The focus group participants concluded that two well-being wording options could be useful to compare in the diary data collection. Both options consisted of a visual 1 to 7 scale that would require participants to circle the correct number to the question. For option one the diary would examine: 'How much did you enjoy your time?', with option 2 examining 'How satisfied were you with your time?'

Diary Study

Introduction

The diary study comprised the principle component of this research project. The recommendations proposed in the literature and the focus groups led to refinement of the 2015 UK HETUS diary design for the present research project. Specifically, the diaries would include two types of well-being questions (aiming to measure enjoyment and life satisfaction) and the analysis would attempt to determine the efficacy and utility of using these different scales as measures of subjective well-being. It was decided that each of these questions use the same 7-point measurement scales to control for the impact of other factors unrelated to the question wording or chance.

Sampling frame

To maximize the study's ability to reach a diverse range of respondents, the diary was distributed to respondents of the Opinion and Lifestyles Survey (OPN). The sample only included private households and the household members living in these private households. Institutions and business were excluded from the survey. The present study consisted of respondents from across England, this was supplemented by respondents in Scotland to ensure circa 500 respondents could be sampled in total.

Respondents who completed the OPN were asked if they were willing to undertake future research, if they were, they were offered the opportunity to complete the 6-hour format time use diary. The respondent who completed the diary was the same respondent who completed the OPN survey. Distributing the diary as a follow up to another survey has given us the chance to link data collected in the survey such as well-being or labour market information given by the respondents in the time diaries. The data for the present study has not been imputed or weighted.

For the OPN, a sample of households is selected at random from the Royal Mail's Postcode Address File (PAF) of 'small users' as the sampling frame. The PAF contains approximately 26 million addresses in Great Britain. The sample for the UK Time Diary Study was drawn in two stages. At the first stage, a random sample of primary sampling units (PSUs), based on postcode sectors, was selected. Within each selected PSU, a random sample of postal addresses was then drawn. In England and Scotland 64 postal sectors were selected, with probability proportionate to size. Within each sector, 30 addresses were chosen randomly giving an initial sample of 1920 addresses each month. One person per household was randomly selected as the respondent. The interviewer determined the household composition and the respondent was selected from amongst all the over-16s using a Kish grid.

Sample Size

A total of 899 respondents who completed the OPN survey were offered the diary, with 378 respondents consenting to complete a diary. This resulted in 159 returned diaries[6]. The further survey participation was optional for OPN participants.

---

[6] This includes diaries that were received, including those deemed to be poor quality, see the "diary quality check" section in the method for details on diary quality.

Field procedures

Interviewer training

Interviewers were briefed on the administration of the diary via interviewer instructions sent out in July, in advance of the August OPN survey field period. The instructions covered a summary of the study, how to fill the diary out and administration. Interviewers were given copies of the diary to familiarize themselves with prior to commencing fieldwork, which included a pre-completed example page (see Appendix E). A full set of written instructions covering diary procedures was also provided for interviewers (see Appendix F for a copy of the interviewer instructions).

Fieldwork period

The interviewing period for OPN started in the first week of August and continued for the duration of the month. Interviewers were instructed to make up to eight calls at an address at different times and on different days of the week.

Contacting respondents

Respondent letters

Participants eligible for the Time Diary study received a letter provided to them by the interviewer after completion of the OPN survey. The letter explained the purpose of the study and stated we would like them to complete the diary within the next 7 days. The name of the study was changed for fieldwork purposes to a simplified 'Time Use study'. This avoided highlighting that the study was testing out scales, which can seem off-putting and less important to potential participants. The letter also provided reassurances about confidentiality as well as contact details and a link to the participant website for any queries respondents had.

The letters were handed out to respondents who said they would be interested in future research after completing the OPN survey. Interviewers gave the respondents' time to read the letter after they had explained the time use diary they would be required to fill out. Interviewers left the letter, a copy of the diary and a pre-paid return envelope with the respondent. A copy of respondent letter is available at Appendix G.

Confidentiality

Participants were informed in the respondent letter that that any information they provided would remain confidential. Participants were provided with a telephone number for ONS's free enquiry line that they could telephone if they had any queries. Any substantive queries or complaints were subsequently passed on to research team to deal with. Interviewers would remind participants that the diary was voluntary before asking whether the respondent consented to complete the diary.

The diary

Placing the diary

All participants who had completed the OPN interview, and who consented to partaking in future research were eligible to complete a diary. By this stage interviewers had built up a rapport with the participant and introduced the diary as something that participants could complete in their own time.

A short summary of the diary and what was required was included in the OPN survey to make it easier for interviewers to discuss the diary with the respondent. Interviewers would explain the diary briefly, explaining that it was voluntary before asking whether respondents consented to complete the diary. Once consent had been given by the respondent, the interviewer would explain the protocols. The questionnaire also included checks to ensure protocols were being followed correctly – such as handing the respondent the materials and filling out the respondent serial number on the front of the diary.

Interviewers completed the details on the front of the diaries, specifically the serial number. They then explained to participants the procedure for completing the diary, showing them the instruction page, example page, checklist questions at the end of the diary and the section on whether the respondent consented to a follow-up interview about their experience of filling in the diary.

Diary completion

Diary types

Interviewers were provided with two versions of the diary ("satisfied" and "enjoyed"), each containing instructions, examples, a checklist and questions about whether it was a usual or unusual diary day for any reason. This was to avoid interviewer bias (where certain interviewers obtain respondent consent than others), ensuring a more even distribution of each diary type. Interviewers were asked to alternate the diaries they gave out after each interview. The diary types were indicated on the front either "A" or "B" in large font so the interviewer could monitor which diaries had been handed out and would minimize the risk of the interviewer handing out the incorrect version. Approximately half of the respondents received the "satisfied" diary and the other half received the "enjoyed" diary. Each address was randomly allocated one diary day; one-week day or one weekend to capture a variety of life styles. An example diary page can be found in Appendix H.

Completion by participants

Interviewers left the diaries and a return envelope with the participants. They showed participants the instructions and example pages at the front of the diary. Interviewers would indicate the number to ring on the back of the diary for participants to call if they had any queries during the diary completion period.

The diary pages required participants to record what they were doing in specified ten-minute slots. The selected time periods in the diary covered early morning, mid and late morning, through to afternoon, late and early evening. These time periods were selected to give a variety of activities to allow for capturing varying lifestyles to test the appropriateness of the well-being scales proposed by

the focus group. The diaries selected 6 time slots from the HETUS 24-hour diary to reduce respondent burden[7], whilst retaining the ability to assess the effectiveness of the well-being scales.

The diary was split into a series of columns, which requested the following information for each time slot:

- main activity undertaken during time slot (and secondary activities should two or more activities be carried out at the same time)
- whether a smartphone, tablet or computer was used
- the location or the mode of transport
- if the respondent was with another person[8]
- "enjoyment" or "satisfaction" scores (depending on diary type)

Participants were encouraged by the interviewer and in the instructions to complete the diary from time to time during the course of the day, not just at the end of the day.

The interviewers informed the respondents how they should be recording the information in the diary:

- Write down in as much detail as possible what activity was undertaken in every 10-minute slot across the 6 hours
    - Write only one main activity on each line
    - Mark the duration of activities clearly
- Record a secondary activity whenever something is done simultaneously with the main activity.
- Do not forget to separate the travel from the activity and also record the location and mode of transport in the appropriate columns
- Don't forget to indicate the use of a computer, smartphone or tablet.
- Record whether the time is spent alone or together with people. You do not need to be doing the same thing as other people you are spending time with.
- The 'full' diaries have a column to record enjoyment/satisfaction of activities using a scale from 1-7, with 1 being that they didn't enjoy/were not satisfied with the activity at all and 7 being that they enjoyed/were satisfied with it very much.
- To ensure linkage of data from the OPN to the diary the OPN respondent must be the diary respondent, not another member of their household.

Checklist questions

At the end of each diary, a series of contextual questions about the day was asked. It covered when participants completed the diary, the type of day and whether they were on a trip on the diary recording day. It also contained a reminder to check that they had completed the diary adequately and a question on whether they consented to a follow up telephone call to discuss their experiences of filling out the diary. These questions can be found in Appendix I.

---

[7] A number of the diary entries squeezed most of the day's activities into the slots provided, and it appeared that a full day account would have been simpler for some respondents than the abbreviated day.

[8] The "were you alone" column was edited for this study from various checkboxes to one checkbox for "alone" and a write in column for the relationship to other people the respondent spent time with. The reason for this was two-fold: it was felt that the grid-like design was quite off-putting for respondents to fill in and it excluded some relationships that respondents may feel are important to them.

Diary return

Interviewers left a return postage paid envelope with the respondent to ensure the diary was returned as quickly after completion as possible. There was no follow-up to ensure respondents remembered to complete the diary.

Diary Response

In total, 378 diaries were distributed to potential respondents from the Opinion and Lifestyles (OPN) survey who indicated they would be willing to participate. 159 completed diaries were received by ONS. However, as outlined in the section on diary quality below, three diaries were eliminated from subsequent analyses because they failed basic quality checks. Therefore, the response rate was 42.1% and when the poor-quality diaries were denoted as non-response, response was 41.3%. Whilst this sampling approach is not directly applicable to HETUS sample surveying, it is worth noting that the response rate for this study is within 10 percentage points of the diary response rate within the UK HETUS survey collected by the University of Oxford at 32.8%.

To maximize diary response the date to accept diary returns was extended from the end of September to the end of October 2017, with the aim being to ensure the maximum number of diaries were returned and coded before fieldwork closed. Both UK HETUS surveys, collected by Office for national statistics in 2000/01 and the University of Oxford Centre for Time Use Research in 2014 also had to extend the period of data collection to improve response rates.

Diary coding, quality check and editing

Coder training

In advance of their first coding assignment, all coders were provided with a half-day project-specific training session for the Time Diary Study led by a time use expert, who had previously worked at the Centre for Time Use Research. The training session covered the navigation of the diary, exploration of and familiarisation with the code frames, missing data codes, rules and best practice for coding main and secondary activities and standardisation of raising queries. Any diaries that coders were unsure about were checked by time use expert. An update was then provided to the coders with the knowledge of any potential issues to avoid inconsistencies in coding.

Each coder completed the coding for a subset of diaries initially. These were reviewed and discussed to ensure the coders had satisfactorily fully grasped the rules. Throughout the coding period a series of spot checks were carried out by a second coder to ensure the codes entered were accurate.

Diary coding

To ensure the process of coding was efficient; coders directly input the data into a diary spreadsheet. Coders were provided with a set of diary coding instructions and a Word version of the full code frame for reference. The process was iterative, with additional advice passed on to the coding team by a time use expert during the coding period. The code frame was static and was used to code the following items:
- Main activity

- Secondary activity (or activities)

The coding was primarily undertaken by one coder, then the quality was checked by a time use expert.

Diary quality

As with all ONS projects, the diaries were checked for quality of the information that respondents had provided. The following quality checks were undertaken:

1) a check that the diary episodes and starting time slots were in the correct sequence with no duplications or gaps
2) a check that all diaries had 360 minutes coded
3) a check for peculiar patterns or recurrence of any typos spotted
4) a check that diaries contained logical episode patterns and reflected basic domains of activity most people do almost every day (sleep or rest or downtime; eating or drinking; personal care; movement, travel or exercise)
5) a check that the gender of the diarist (based from the name on the cover page of diaries) matched the actual gender recorded of the respondent in the OPN file
6) a check the day of the week and version of the diary recorded on the diary matched the information assigned prior to interview.

To have sufficient detail for analysis, there were several criteria the diaries were required to meet:

- Diaries should not be missing significant amounts of time (for this study ONS allowed for no more than 20 minutes missing; in the literature for 24-hour diaries, National Statistics Institutes allow between 30 and 90 missing minutes)
- Diaries should also have a minimum of 7 episodes (ONS allowed 6 due to the short period observed). Additionally, diaries should include at least 3 of the four daily activity domains if they are relatively low episode diaries; for the present analysis, diaries with at least 2 daily domains were allowed
- Diaries should have the day of the week recorded and the gender and couple status of the diarist should be known.

The quality checks revealed that there were only three low quality diaries. These had limited episodes, with minimal information on activities and ratings. These diaries were eliminated from the analysis and between them contained 14 episodes. A total of 2418 activity episodes from 156 diaries were included in subsequent analyses (1175 from enjoyed diaries and 1243 from satisfied diaries).

Diary analysis

Well-being activity ratings

Overview of ratings

Figure 1 below shows that the most frequently selected rating was 7 ("enjoyed very much" and "very satisfied"). The frequency of ratings for both satisfied and enjoyed diaries increased with each scale rating, indicating that rating selection was more likely to be towards the upper end of the scale regardless of diary type.

The largest difference in the average scale rating, between satisfied and enjoyed diaries, was for rating 7, with a difference of 4 percentage points. Scale points 2, 4, 5 and 6 were more frequently rated in satisfied diaries whereas scale points 1, 3 and 7 were more frequently rated in enjoyed diaries.

Figure 1: Average diary rating for enjoyed and satisfied diary respondents

England and Scotland, 2017
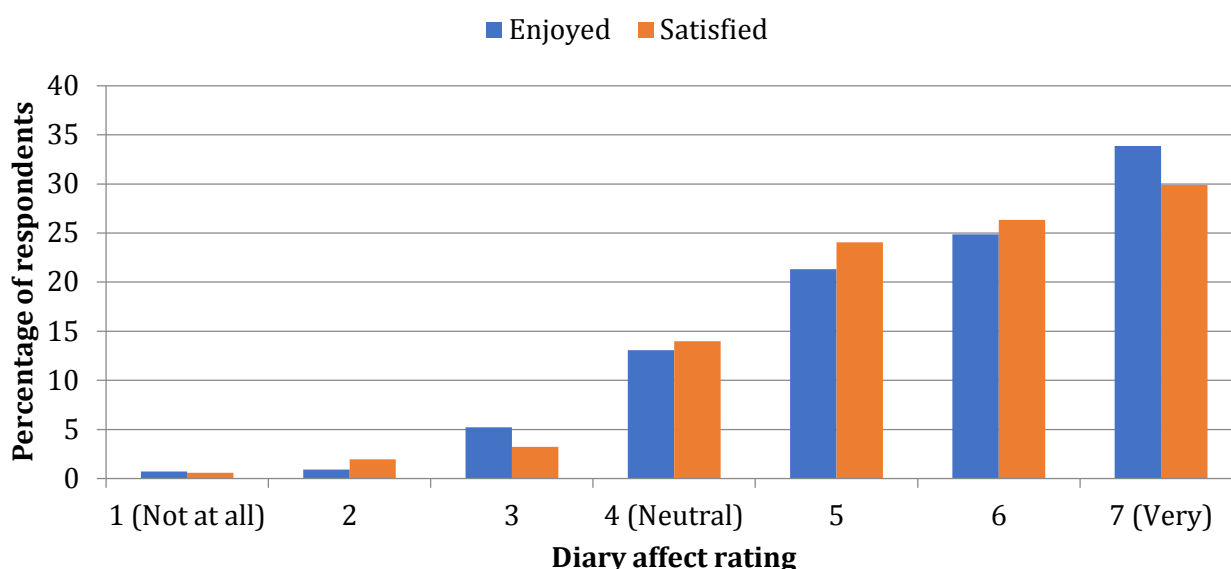
Source: Time Use feasibility study, Office for National Statistics

For the purposes of visual comparison, figure 2 shows how diary respondents rated their well-being in the Opinions and Lifestyle (OPN) survey. This shows a similar trend towards upper ends of the scale as seen in figure 1, but in figure 2 the frequency of ratings appears to tail-off at the highest points of the scale, whereas in the diary well-being scales the highest frequency of enjoyed and satisfied responses were scored at the most extreme end of the scale[9]. This response trend will have implications for the interpretation of any findings related to the diary well-being scales within this section, as responses are also much more likely to favour the upper more positive, points on the well-being scales.

---

[9] It should also be noted that figure 1 shows the percentage of activity episodes in which different well-being scale responses were selected but figure 2 shows the percentage of diary study respondents who had also on a previous occasion provided well-being ratings on the OPN survey.

Figure 2: Diary respondents' well-being ratings

England and Scotland, 2017

Legend:
- Overall, how satisfied are you with your life nowadays?
- Overall, to what extent do you feel that the things you do in your life are worthwhile?
- Overall, how happy did you feel yesterday?
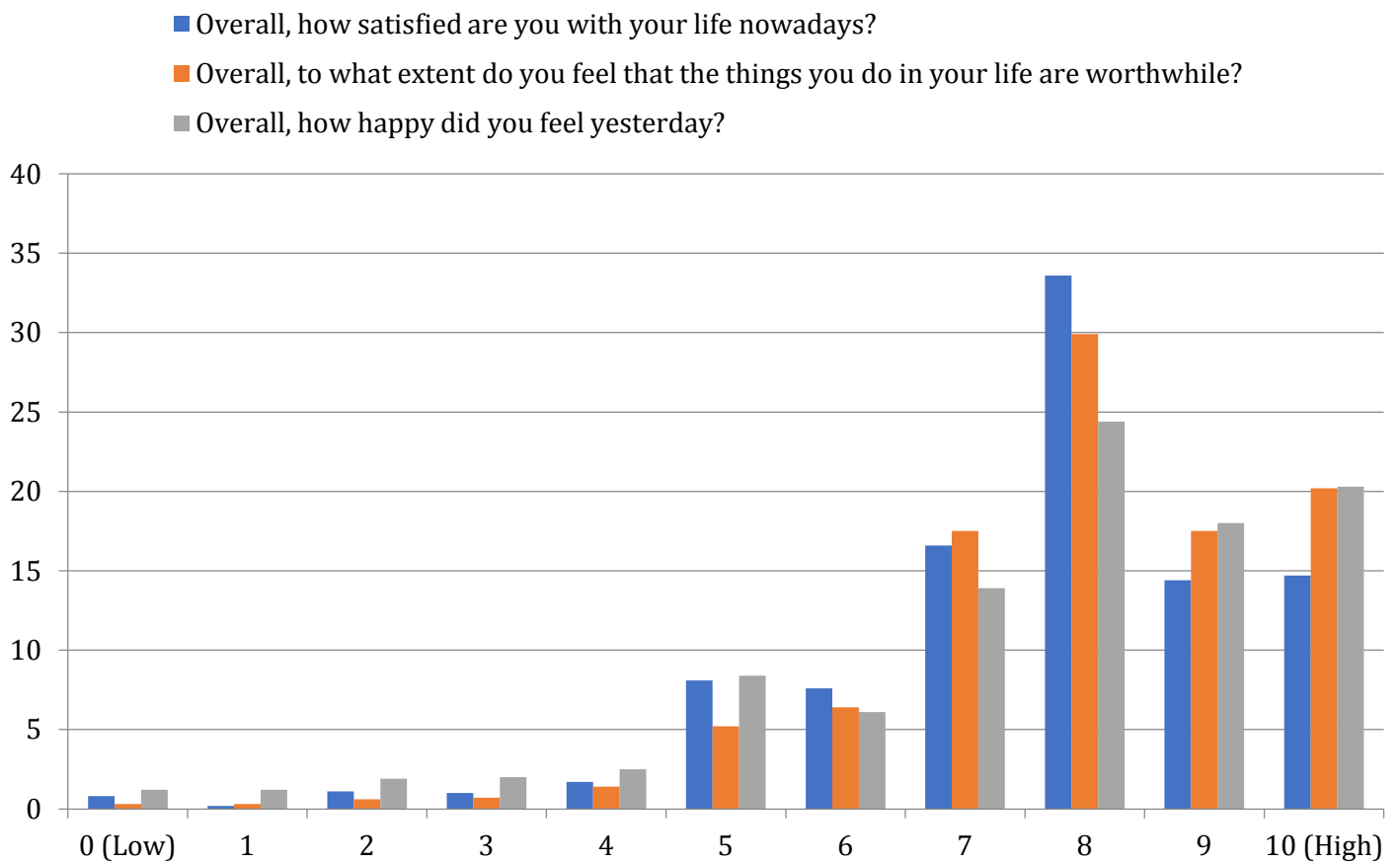
Source: Opinion and Lifestyle Survey, Office for National Statistics

Day of the week diary was completed

In total, 1708 activity episodes occurred on weekdays and 550 activity episodes occurred on weekends. Figure 3 suggests that most activities occurred on Wednesdays (18.5%), followed by Saturdays (16%). For specified weekdays, the fewest number of activities occurred on Sunday (7.3%), followed by Friday (8.4%). 14 participants did not note which day they completed the diary, which resulted in 8.5% of episodes without a specified day of diary completion.

Figure 3: Percentage of activity episodes by day of diary completion

England and Scotland, 2017
Percent (%)
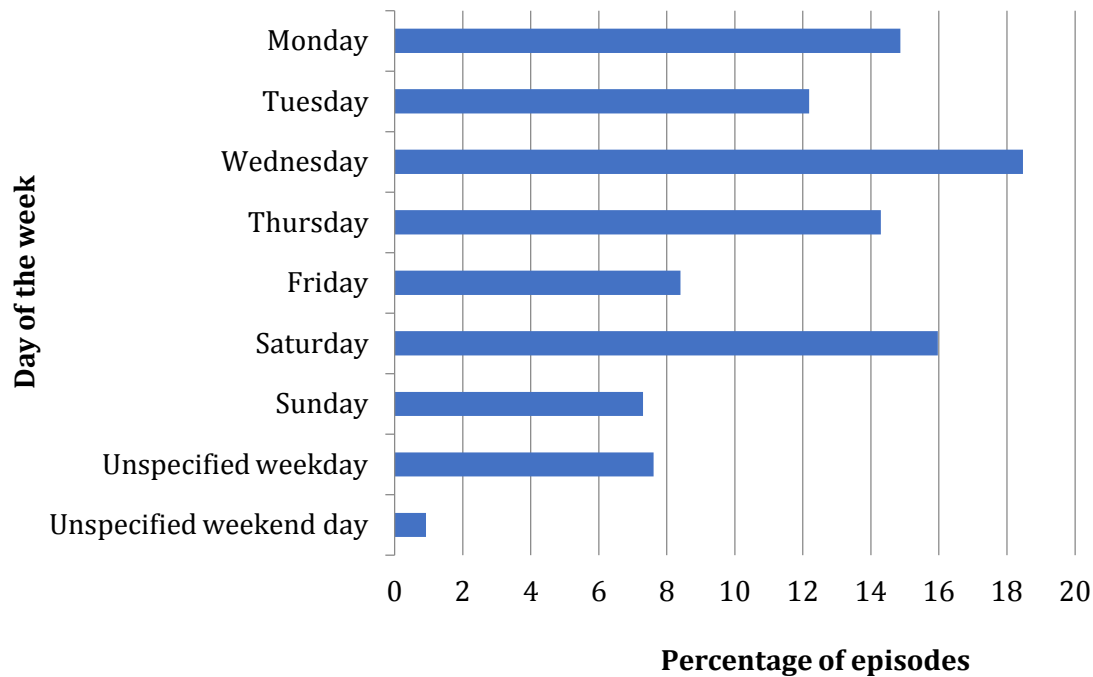
Source: Time Use feasibility study, Office for National Statistics

Figure 4: Activity locations reported on weekdays and weekends

England and Scotland, 2017
Percent (%)
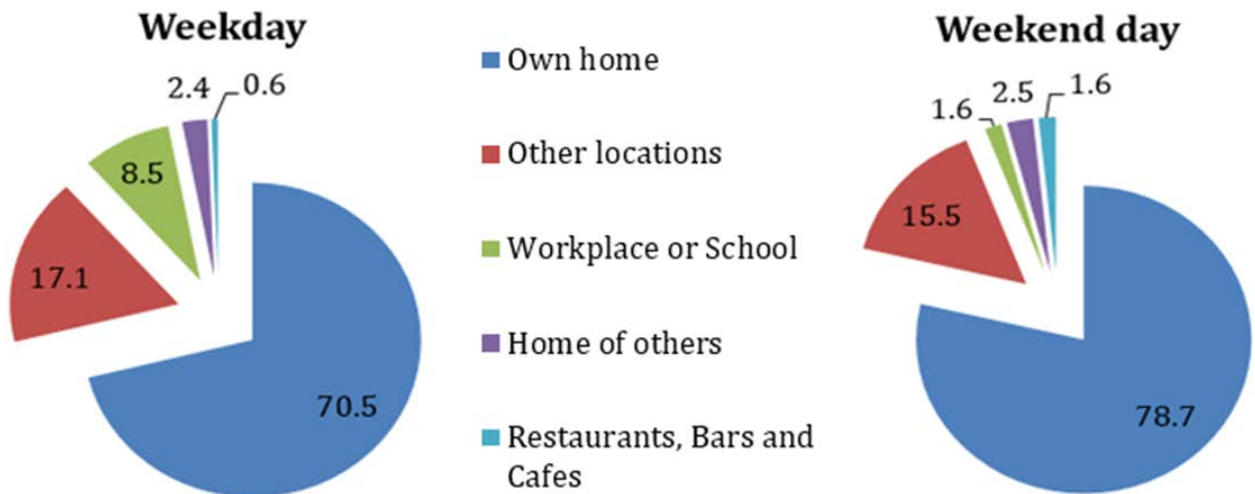
Source: Time Use feasibility study, Office for National Statistics

Figure 4 shows that across all locations, activities occurred most frequently at home, which was true for both weekdays and weekends (70.5% and 78.7%, respectively). Restaurants, Bars and Cafes etc were reported as the least frequent location for both weekdays and weekends at 0.6% and 1.6% of activities occuring there. Workplace or school and other locations were more likely to have activities completed in those locations during the week (8.5% and 17.5%, respectively) than on the weekend (1.6% and 15.5%, respectively).

Main activities

Activities

Smart devices used in main activities

Figure 5 below shows the main activities in which smart devices were most and least frequently used (see appendix K for a list of all main activities and whether smart devices were used). It is unsurprising that they were infrequently used in sleep/naps but is more surprising they were used so infrequently in gardening and all other travel. It is also interesting how frequently they were used in employment and employment related activities.

Figure 5: Main activities in which smart devices were most and least frequently used

England and Scotland, 2017

Percent (%)
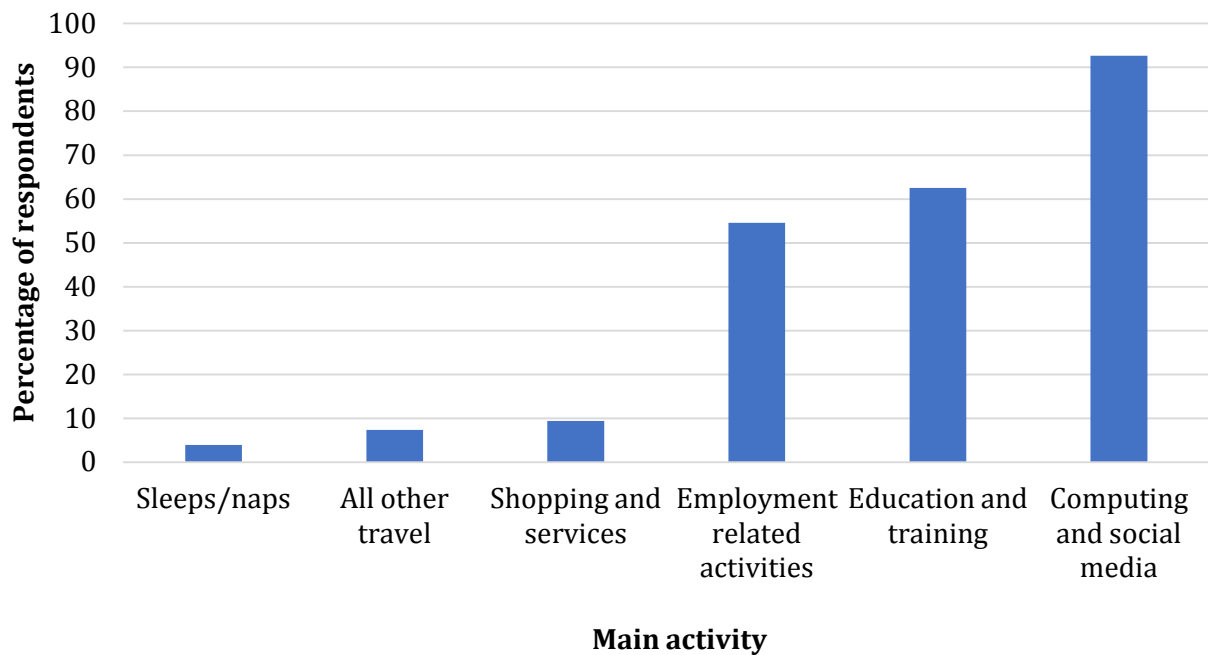
Main activity

Source: Time Use feasibility study, Office for National Statistics

Main activities and secondary activities

Figure 6 below shows that the most frequently reported main activities were "Sleeps/naps" at 16.7% of all episodes, followed by "Other domestic work and household management" at 10.5%. It also shows that the least frequently reported main activities were games (0.6%), Childcare – read/teach/play (0.8%), and Voluntary/religion (1%).

Figure 6: Most and least frequent main activities

England and Scotland, 2017
Percent (%)
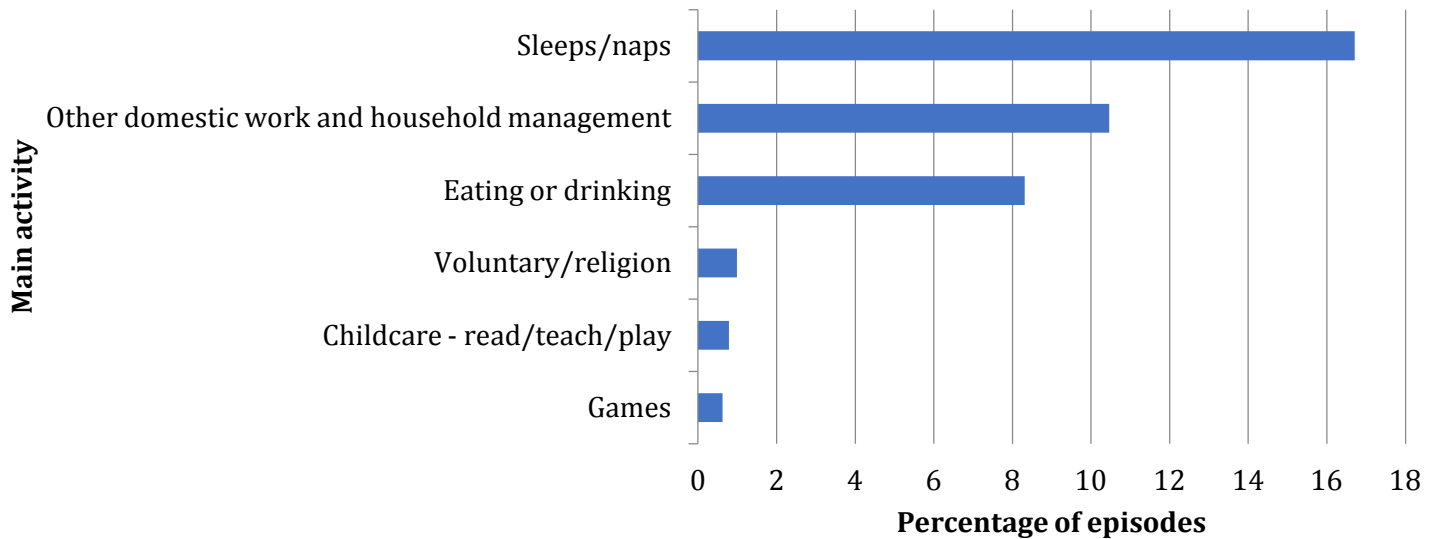
Source: Time Use feasibility study, Office for National Statistics
1. Figures will not sum to 100% owing to inclusion of only some episodes

Figure 7 below shows the most frequently reported secondary activities were "Audio/visual media use" at 24.9%%, followed by "Social leisure" at 13.9%%. The least frequently reported secondary activities were "Employment related activities" (0.6%) and "Games" (0.8%).

Figure 7: Most and least frequent secondary activities
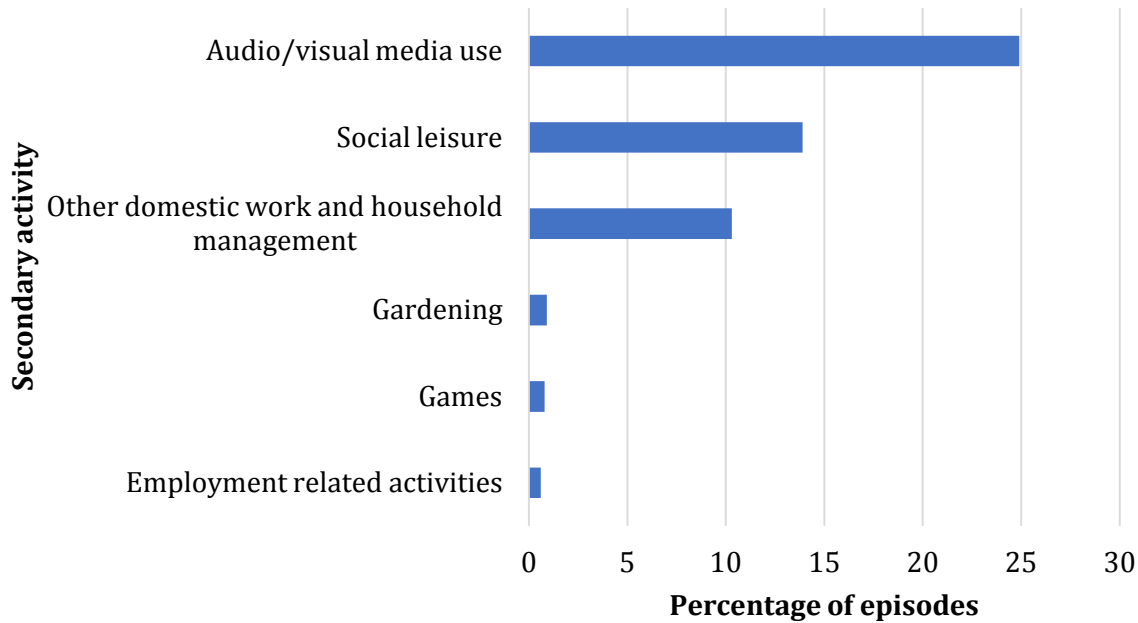
England and Scotland, 2017
Percent (%)

Source: Time Use feasibility study, Office for National Statistics
1.   Figures will not sum to 100% owing to inclusion of only some episodes

Figure 8 shows that the most popular main activities with which, secondary activities were also undertaken were "other domestic work and household management" and "eating/drinking" (155 and 130 episodes, respectively).

Figure 8: Main activities most frequently associated with secondary activities (data labels show number of secondary activities)

England and Scotland, 2017
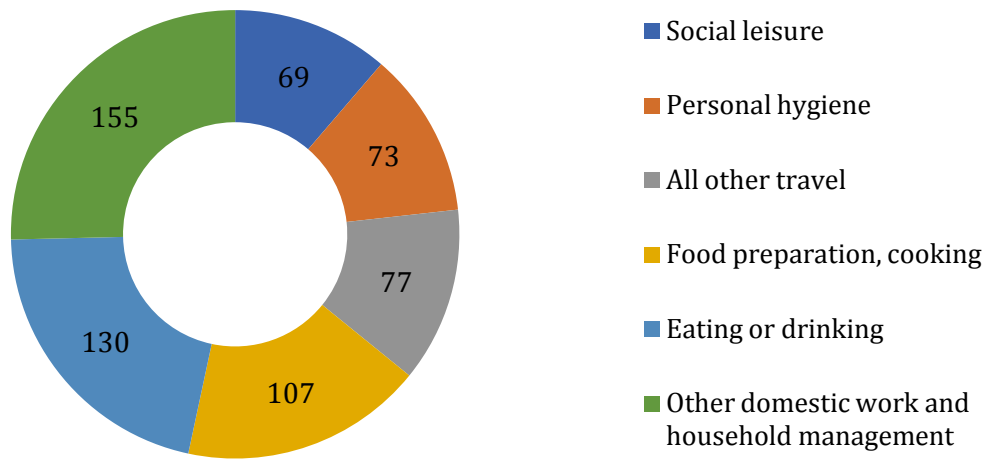
Source: Time Use feasibility study, Office for National Statistics

Figure 9 shows that the main activities least associated with secondary activities were "games" (7 episodes) and "Childcare – read/teach/play" (9 episodes) and "Commuting" (9 episodes). As would be expected, "Sleeps/naps" also feature in this group.

Taken together, both figures 8 and 9 could indicate that tasks which are less likely to co-occur with secondary tasks completed (figure 9) are more mentally demanding than the main activities more likely to co-occur with secondary tasks (figure 8).

Figure 9: Main activities least frequently associated with secondary activities (data labels show number of secondary activities)

England and Scotland, 2017
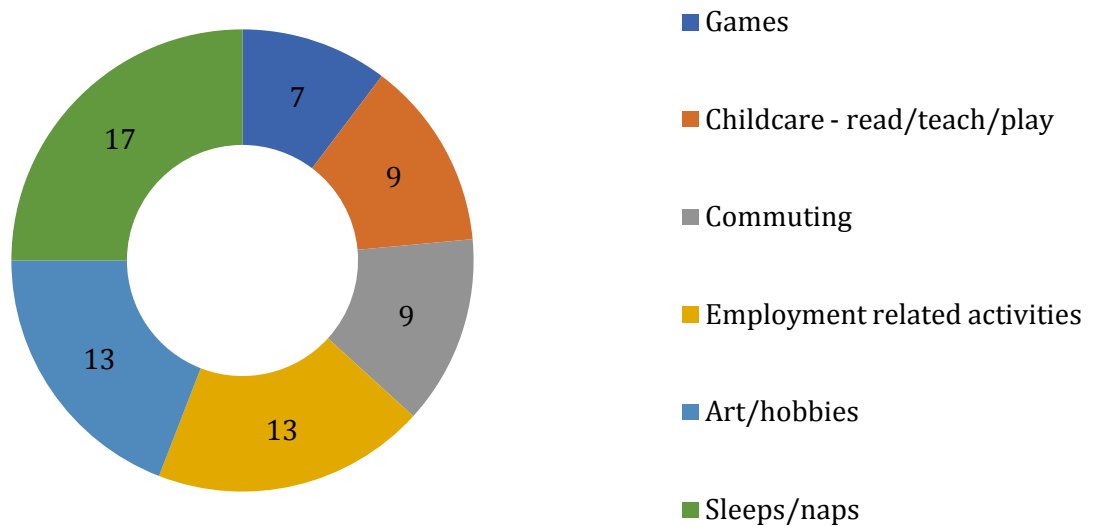
Source: Time Use feasibility study, Office for National Statistics

The presence of others when undertaking co-occurring activities

Table 2 below shows that when undertaking both co-occurring and singularly-occurring activities respondents were with their spouse or partner more frequently than they were with anyone else. This was most frequently the case for non-co-occurring activities (46% of episodes compared to 33.8%). Respondents note being accompanied by a work or school colleague more than three times more frequently when undertaking singularly-occurring compared to co-occurring activities (16.7% of episodes compared to 5.1%).

Table 2: Other people present during co-occurring and singularly-occurring activities

England and Scotland, 2017

Accompanying persons	Singularly-occurring activities		Co-occurring activities	
	Number of episodes	**%**	**Number of episodes**	**%**
Spouse or partner	417	46	263	33.8
Own child who lives with you or not specified	146	16.1	145	18.7
Own child who lives elsewhere	15	1.7	28	3.6
Grandchild or other child relative	22	2.4	43	5.5
Parent	18	2	33	4.2
Other adult relative	27	3	41	5.3
Sibling	15	1.7	16	2.1
Friend	57	6.3	80	10.3
Neighbour	n/a	n/a	8	1
Work or school colleague, boss, teacher, client	151	16.7	40	5.1
Other people - including relationship unclear	9	1	29	3.7
Animal	28	3.1	51	6.6

Source: Time Use feasibility study, Office for National Statistics

Activities and well-being

Together figures 10 and 11 below show that "Art/hobbies" and "Sport/exercise" were rated as among the most enjoyable and satisfying main activities. However, there was little consistency between the least enjoyable and satisfying main activities. For enjoyment "Childcare – read/teach/play" was rated as one of the more enjoyable activities, whereas for satisfied "Games" rated as one of the more satisfying activities.

Figure 10: Least and most enjoyed main activities

England and Scotland, 2017
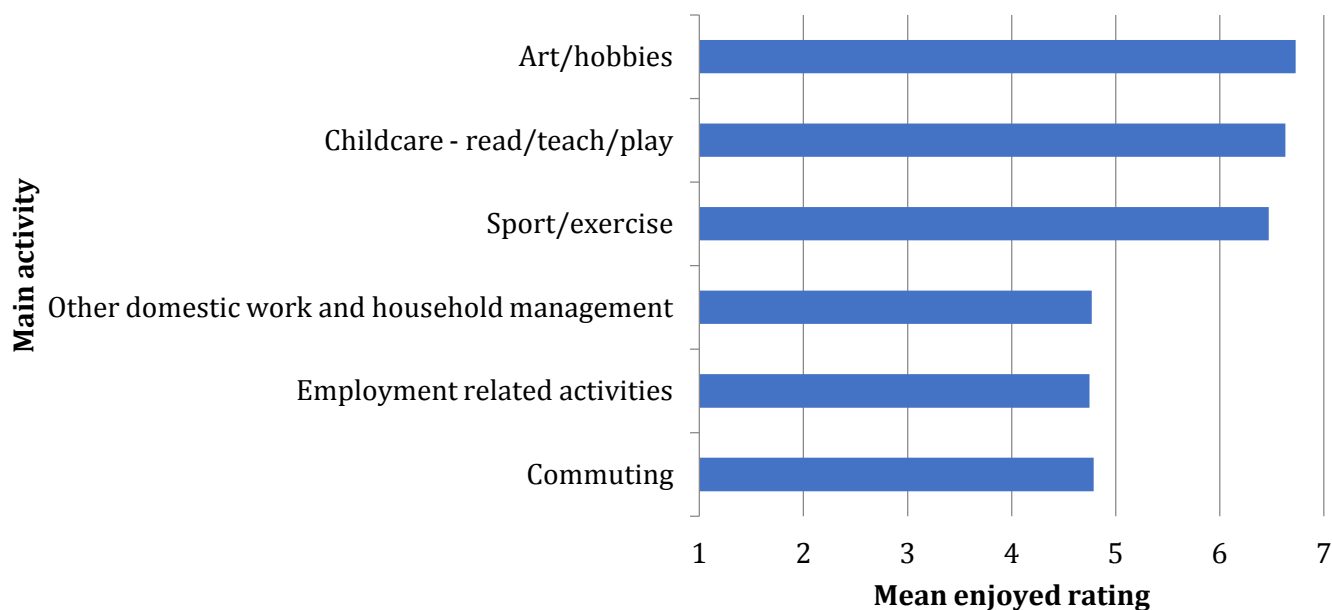
Source: Time Use feasibility study, Office for National Statistics
1. Episodes in which no well-being rating was provided by participants were excluded from this analysis

Figure 11: Least and most satisfying main activities

England and Scotland, 2017
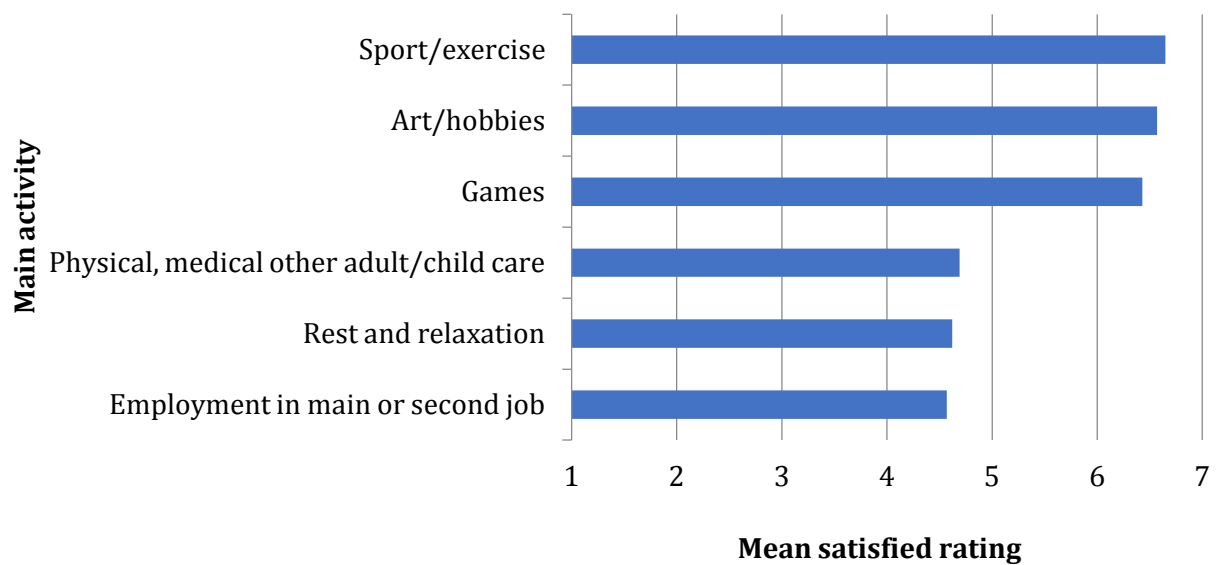
Mean satisfied rating

Source: Time Use feasibility study, Office for National Statistics

1. Episodes in which no well-being rating was provided by participants were excluded from this analysis

Activity ratings and differences between well-being groups

Figure 12 illustrates that the biggest difference between the satisfied and enjoyed diaries was for the main activity "Rest and relaxation[10]" at 1.3 points on the well-being scales. The second biggest difference was for the main activity "Physical, medical, other adult/child care" at 1.2 scale points.

It also shows that in most cases enjoyed ratings were rated higher than satisfied ratings. The only activity that didn't follow this trend was "Other domestic work and household management", which might well be expected to be more satisfying than enjoyable.

Figure 12: Largest differences for main activities between satisfied and enjoyed diaries

England and Scotland, 2017

---

[10] Several activity descriptions have been shortened for simplicity in this section. For example, Rest, relax, think, do nothing has been shortened to "Rest and relaxation". See appendix M for a glossary showing all activity descriptions and their shortened versions.
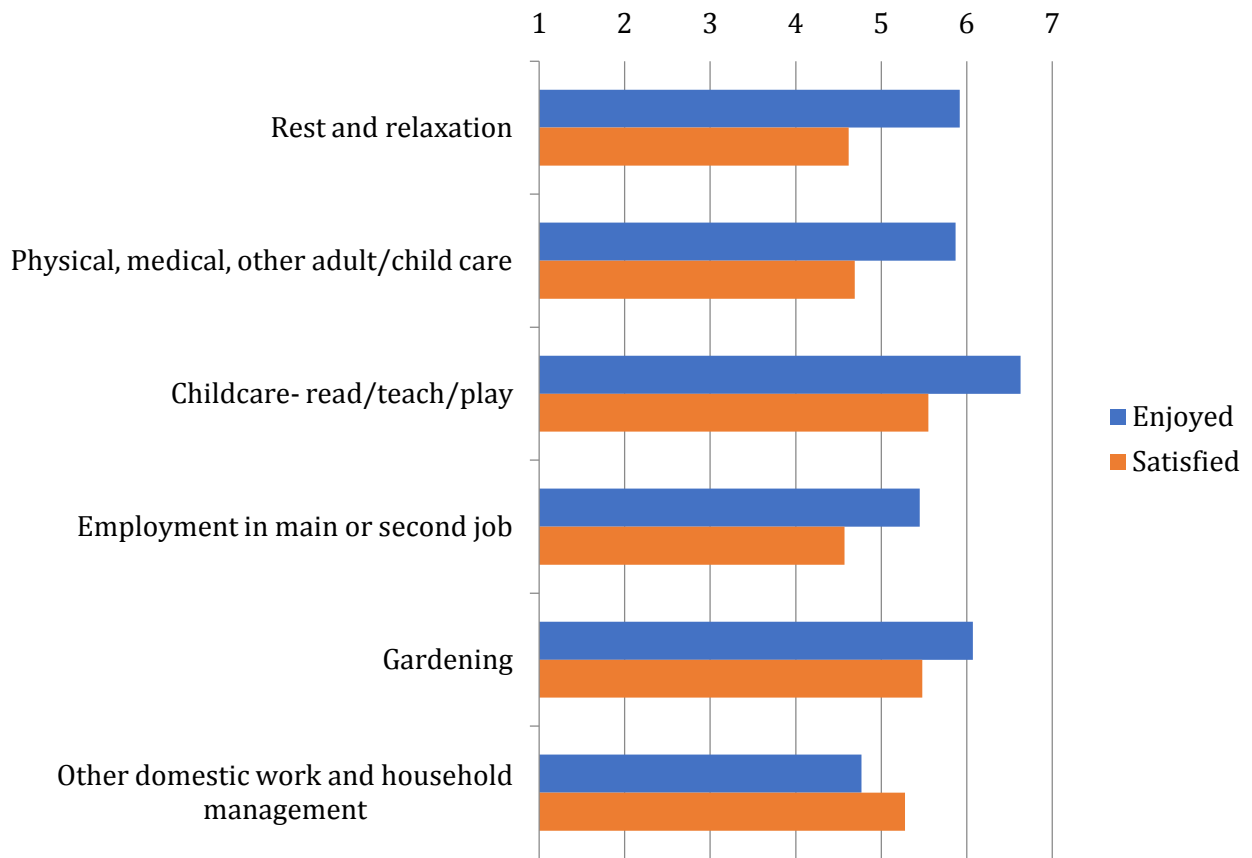
Source: Time Use feasibility study, Office for National Statistics
1. Episodes in which no well-being rating was provided by participants were excluded from this analysis

Literature-driven analysis of specific main and secondary activities

Housework, childcare and well-being

As covered in the co-occurring activities section of the literature review one aim of time use research has been to gain a greater understanding of domestic work, which is frequently under-reported in other surveys. Therefore, diary study data were examined to examine potential associations between these particular activities and well-being. Table 3 below shows the mean well-being ratings and (number of analysed episodes) for childcare and housework activities when these activities were performed as primary and secondary activities. Table 3 shows that when childcare was undertaken as a main activity, the mean enjoyed rating (6.3) was markedly higher than the mean satisfied rating (4.9) but that when housework was undertaken as the main activity, the mean satisfied rating (5.3) was higher than the mean enjoyed rating (4.9).

Table 3 also shows that when childcare was undertaken as a secondary activity, the mean enjoyed and satisfied ratings were both high but there was very little difference between them and when housework was undertaken as a secondary activity there was no difference between enjoyed and satisfied ratings.

Table 3: Showing the mean enjoyed and satisfied ratings and number of affected episodes, when childcare and housework were reported as the main activity

England and Scotland, 2017

| | Main activity | | | | Secondary activity | | | |
	Childcare		Housework		Childcare		Housework	
	No. episodes	mean rating	No. episodes	mean rating	No. episodes	mean rating	No. episodes	mean rating
Enjoyed	27	6.3	209	4.9	37	6.1	67	5.3
Satisfied	34	4.9	202	5.3	28	6.2	70	5.3

Source: Time Use feasibility study, Office for National Statistics
1. Episodes in which no well-being rating was provided by participants were excluded from this analysis

Contexts and well-being

Figure 13 below shows that respondents who felt rushed, and travelled by walking or using other physical transport (e.g. bicycle) had higher ratings for satisfaction than for enjoyment. Conversely, respondents who undertook their activity in the workplace or school or who travelled using unspecified transport or taxi during the main activity reported higher enjoyment.

Figure 13: Largest differences between mean enjoyed and satisfied ratings by main activity and context

England and Scotland, 2017
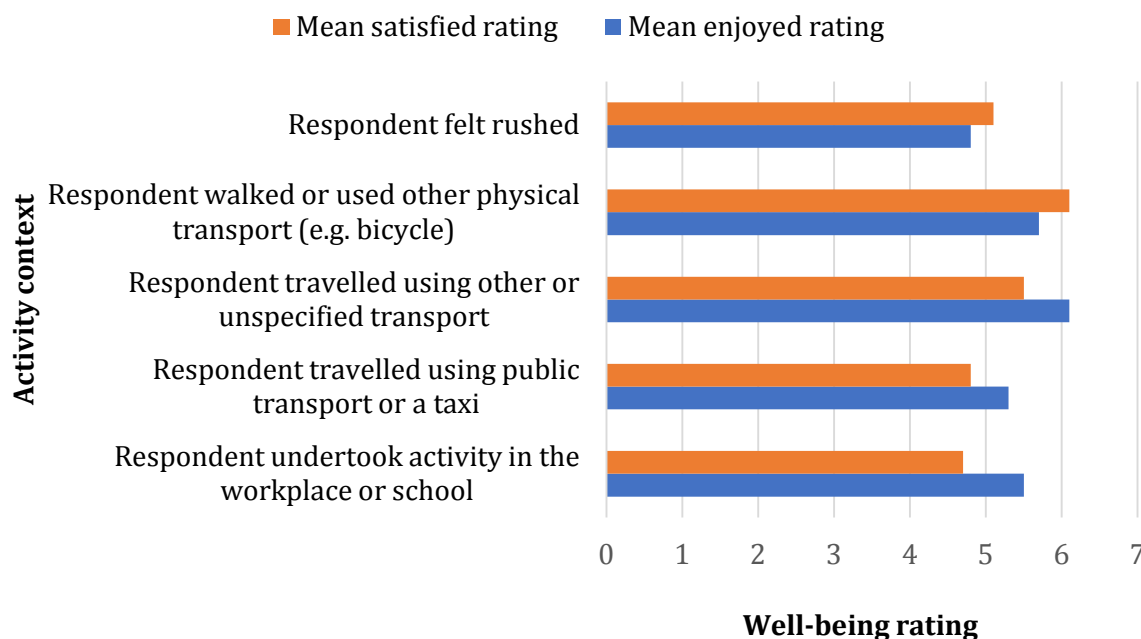
Well-being rating

Source: Time Use feasibility study, Office for National Statistics

1. Episodes in which no well-being rating was provided by participants were excluded from this analysis

Figure 13 reveals that "Respondents undertook main activity in workplace/school" had the largest difference between satisfaction and enjoyments rating (0.8 rating points). The second largest difference between ratings was for "Respondents travelled using public transport or taxi" or "other/unspecific transport" with a 0.6 scale point difference (see appendix J for all mean ratings).

Well-being, co-occurring activities and location

Table 4 below shows that there was little variation in diary-study respondents' reported enjoyment and satisfaction of activities when undertaken in combination and in isolation in various locations. It does however, show that respondents enjoyed singularly-occurring activities at others' homes much more than co-occurring activities at these places. It also shows that those who completed satisfied diaries found singularly-occurring activities in the workplace to be much less satisfying than those who completed enjoyed diaries, enjoyed singularly-occurring in this location.

Table 4: co-occurring and singularly-occurring activities in different locations: mean emotion rating for enjoyed and satisfied diaries

England and Scotland, 2017

Location	Enjoyed		Satisfied	
	Singularly-occurring activities	Co-occurring activities	Singularly-occurring activities	Co-occurring activities
Home	5.7	5.5	5.7	5.6
Other home	6.7	5.5	5.5	6.4
Other place	5.7	5.7	5.6	5.7
Restaurant	6.3	6.8	n/a	6.2
Workplace	5.5	5.4	4.6	5

Source: Time Use feasibility study, Office for National Statistics
1. Episodes in which no well-being rating was provided by participants were excluded from this analysis

Validation of diary well-being scales and well-being scales from the OPN survey

Four well-being questions were asked as part of the ONS Opinions and Lifestyle survey (OPN). To determine the degree of association between diary well-being scales and these well-being scales (i.e. to determine if they are measuring similar well-being constructs), both scales were first standardized as this diary study and OPN used scales with differing scales of measurement: The OPN well-being scales ranged from 0 to 10 but the diary well-being scales ranged from 1 to 7. Standardised versions of each of the scales were then correlated. The table in appendix L shows that there were small, yet significant correlations between standardized versions of the diary enjoyed scale and a standardized version of the enjoyment-related OPN scale: "Overall, how happy did you feel yesterday?" And between standardized versions of the diary satisfied scale and the following OPN well-being scales:

1. "Overall, to what extent do you feel that the things you do in your life are worthwhile?"
2. "Overall, how happy did you feel yesterday?"
3. "Overall, how anxious did you feel yesterday?"

The table in appendix L shows that there was a slightly stronger correlation between the diary satisfied scale and the OPN scale happiness question (i.e. OPN scale number 2 above) than between the diary enjoyed scale and this OPN scale question. The strongest correlation was between the diary satisfied scale and the final OPN well-being scale which asked "Overall, how anxious did you feel

yesterday?" This correlation was (and should have been) negative because this particular OPN scale was scored in the opposite direction.

Cognitive testing

Introduction

The aim of the cognitive testing was to conduct between 20 – 25 cognitive interviews to gain insight into the participants' experiences of completing the Time Use diary. Cognitive testing is a diagnostic technique that explores the processes employed by people when they complete survey materials and answer questions, such as comprehension, recognition, recall and decision-making/response (e.g. how do they respond to the wording of the question/or how suitable is the measurement scale for capturing all ranges of well-being).

By exploring in a qualitative way, the processes by which people interpret and respond to questions, we can identify potential sources of measurement error. This allows us to address them via appropriate revisions to the well-being scales, ensuring that they measure what we want them to measure as accurately as possible.

Methodology

Cognitive Interview structure

Two rounds were undertaken: The first round was completed by the ONS telephone operations team and the second round was completed by ONS researchers. For the first round, two sessions were run in September: one during the day, and one in the evening to ensure people with different work/life patterns could be captured. Interviewers in these sessions were given a list of participant names and the diary type the respondent had filled out. Interviewers rang the respondents, with the aim of interviewing respondents at that time. If that was not possible, interviewers were asked to make an appointment with respondents for a more convenient time.

For the second round, interviews were conducted on an appointment basis. These appointments ran throughout October 2017, with varying times throughout the day, ranging from 9am – 7.30pm. The content of the cognitive interviews was the same as in round one.

Sample

In total 22 cognitive interviews were completed. In round one, 11 interviews were completed and in round two 11 interviews were completed. Initially, 40 participants were selected to be contacted for a cognitive interview. These were selected from the 82 who consented to complete a follow-up telephone interview after completing a diary. Participants were selected to ensure a variety of people were interviewed, accounting for an even split of enjoyed/satisfied, a range of ages, as well as a variety of diary qualities e.g. full vs. sparse, 1 score for everything.

After the first session in round one, it became clear than non-contact would be quite high, therefore the cognitive interview sample was supplemented with an additional 20 respondents. The same selection criterion was applied for the reserve sample.

Key sample demographics

The key demographics of the cognitive interview sample are shown in table 7 below.

Table 3: Showing key demographics of the cognitive interview sample

Demographic characteristic	Levels of demographic characteristic	Number of participants
Gender	Male	8
	Female	14
Age	16 to 54	8
	55 to 64	4
	65 and over	10
	Mean age = 59.3 years	
	Higher professional occupations or Intermediate professions	8
	Lower professions and semi-routine	6
	Not classifiable for other reasons (inc student)	8
Diary type	Enjoyed diary	8
	Satisfied diary	14

Source: Time Use feasibility study, Office for National Statistics

Interviewer training

For this study, we utilized ONS interviewers who have completed a full three-day course in cognitive interviewing led by National Centre for Social Research. Additionally, interviewers were provided with two topic guides on the cognitive interviews they were going to undertake (one for participants who had completed the satisfied diary and one for participants who had completed the enjoyed diary[11] – see appendix N). These guides included the main questions to ask; along with probes for the interviewers to ask, should the respondent struggle to answer the initial question. Interviewers were also given an empty diary to familiarise themselves with the content. Interviewers were briefed by either a member of the research or field team and encouraged to ask questions before they started interviewing on the content if they felt in any way unsure or unclear about anything. Interviewers were reminded not to provide leading comments in response to respondent answers.

---

[11] The topic guides were identical except that one referenced enjoyment of activities whereas the other referenced satisfaction with activities in line with the differing diary types participants had completed.

Interviews

After obtaining consent for the interview to take place (including audio recording the interview), the interviewer asked questions surrounding 4 main areas (see appendix N for an example topic guide):

> How the respondent found completing the diary
> Questions about the respondents' day yesterday (the day before the cognitive interview)
> What ratings the respondent would give their activities from the day before
> How the respondent understood the wording of the well-being scale

There was a time delay between respondents completing diaries and the cognitive interviews. As a result, participants may have experienced poor recall, which could have impacted upon the quality of the information provided in the interviews. To limit the impact of recall issues, participants were asked to recall yesterday's activities, and many of the questions they were subsequently asked in the interview concerned these activities.

The interviews lasted between 15 – 45 minutes, and were conducted over the telephone. The discussions were recorded, fully transcribed and entered on to a transcription form that mirrored the topic guide. Responses were then collated and coded as part of an iterative process.

Analysis approach

A theoretical thematic analysis (TA) approach was used to analyse the data (as engagement with time use literature occurred prior to the cognitive interviews). Thematic analysis is a method for identifying, analysing, and reporting patterns (themes) within data. TA is a more flexible method, allowing for a more explorative and fluid analysis approach.

The researcher examined the reality of participants filling out the diary, thereby using a realist or "essentialist" approach, as this best reflected what the participants believed their experiences to be. Themes were identified by the researcher[12] . The researchers focused on the 'keyness' of a theme rather than its frequency – i.e. in terms of whether it captures something important in relation to the overall research question. Therefore, themes have been considered as such if they occurred in more than one interview or occurred various times within one interview as that theme had particular prevalence for that individual. The researchers felt this was important not to overlook. The themes identified in the research do not go beyond the semantic content of the data, focusing only on the surface meanings of what participants have said rather than the underlying implicit meaning.

---

[12] As with all qualitative research, it is important to note, that researchers are not free from their own perceptions of human knowledge, as such data is not coded free from any pre-conceptions. However, researchers sought to minimise pre-conceptions by the following approaches: 1) having multiple coders, 2) having other researchers review the analysis and 3) seeking alternative explanations.

The analysis was carried out with the following iterative structure:

Stage 1: Coding:
    1. familiarization with the data
    2. generating initial codes
    3. grouping into themes
    4. Reviewing themes
Stage 2: Descriptive Analysis: defining themes
Stage 3: Explanatory analysis

Whilst themes were identified in the data, the researchers have presented the analysis by each aspect of the diary aspects as this is expected to be more useful for users drawing conclusions on the diary effectiveness.

Results

General impressions of the diary instrument

Prior to more detailed probing about specific aspects and functioning of the diary instrument, participants were first asked about their general experiences in completing the diaries. More than 3 in 4 responses found the diaries straightforward or had no problems to complete:

"Part of the time I did it just after I'd done the task or if I went I kept an eye on the time. It did make me very aware of the time and I just jotted things down to fill in when I got home later so I wouldn't forget what I'd been doing".

This ease the participant felt, on reflection, about their experience completing the diary might be indicated through the use of the word "just".

Only 1 participant reported that they found diary completion to be only ok and only 1 participants reported difficulties at this initial interview stage (concerning the "who with?" response option).

"Made a few mistakes on being alone or being with people. Do I put a cross or not in the box was confusing whether to say I'm on my own or a tick to say I'm with someone? Such a little box forgot about it sometimes. Would have been better to have to write it."

The use of the phrase "little box" demonstrates the space issues the respondent had with the diary, and how such space issues added to their confusion. The respondent explicitly notes potential for diary instrument changes: having write in boxes for the whole of the who with section, could allow more space in future diaries.

Two participants reported no difficulties with the questions they were asked and the same number also reported that they found the diary completion experience "*repetitive*", "*bizarre/surprising*" and "*interesting*".

Some participants also expressed minor frustration over an inability to record sufficient detail about their activities using the diary instrument:

"Felt it maybe didn't give one enough opportunity to say what you actually done"

"There was not always maybe enough room to write everything that you want to if your handwriting is a bit on the large side like mine is, so that was the only negative thing but apart from that it was quite a fun task to do"

Electronic diaries should be able to allow more space to alleviate such frustrations; however, space will still be an issue to consider in paper diaries. The researchers will have to choose whether

collecting all activities respondents want to record is more, or less crucial than other information collected in the diary.

Despite the low numbers of participants initially reporting overall difficulty with the diary instrument they were asked in a follow-up question about any particular aspects they struggled with. More than half the participants who answered this question reported no difficulties and that they fully understood the task they were set:

"It was just a log of what I've done."

The word "just" is again used by a respondent which may indicate their perceived ease of filling in the diary.

Of those who reported difficulties, a third of reports concerned difficulties with the concept of co-occurring tasks, with the same proportions reporting difficulties determining the level of detail to provide in recording their daily activities as well as potential concerns about operational definitions related to the "who with?" diary response option:

"I think you need to define what "being alone" and what being "with someone" is – [because] you could be in the same building technically someone's there with you, but you're doing a task on your own does that count?"

This respondent could be recounting their thought process at the time explicitly for the "who with?" section. Alternatively, the respondent's use of the word "technically" could suggest that when they were looking for issues with the diary they argued that semantics within the diary needed greater clarification as some respondents could find that confusing.

Determining the main activity and secondary activities

Three themes came out when discussing the meaning of the main activity from the respondents: "attentional task focus", "temporal tasks" and "perceived task priority". A third of participants asked how they decided what to write as the main activity, reported that this was easy to do. Just under a third of these participants understood main activities as those upon which attention is most focused, a fifth of participants understood main activities as those they spent most time doing and just under a third of participants understood main activities as the tasks they were doing at that time only. Just over 10% of participants indicated that they used the instructions to make this determination and a similar proportion indicated they experienced little or no task variation yesterday.

Responses indicating that the participant understood main activities as those upon which attention was most focused:

"Quite often I work on computers and design things so there was always something that took up the most of my attention, so whatever I was focusing on the most was the main activity"

"When I was picking the blackberries, I was just doing that and trying not to fall off the ladder, [pause] when I am sewing I am also watching television so I suppose there was a secondary activity in that. When I was doing the jigsaw, I was listening to the radio so."

Responses indicating that the participants understood main activities as those they spent most time doing:

"Housework was my main activity. [Because] I spent most of my time doing that"

"I guess it would be the amount of time I spent on it"

Responses indicating that the participants understood main activities as the tasks they were doing at that time only:

"I'll have one main thing I'll have to get done in the house. I think I was washing curtains or blinds or something that day"

"I think I just thought what was the most important"

Several quotes have prefixes in such as "I think"/ "I guess"/ "I suppose", which could suggest of the respondents were struggling to recall how they designated these tasks as main or secondary. Alternatively, these phrases could indicate that these respondents may not have consciously processed how they designated a task as secondary or main at the time, suggesting their categorisation was more instinctual.

Collectively, this indicates that respondents have different interpretations when reporting different activity information. If researchers felt comparability between respondent diaries was important diaries may need to contain more explanation on how to choose different types of tasks (main versus secondary tasks), which would add clarity if respondents were unsure, or processed the categories unconsciously.

Over half the participants asked how easy or difficult they found it to decide what the main activity and secondary activities were, reported that they found this easy. Just over 1 in 4 of responses indicated that they understood co-occurring in relation to focus of attention. Just under a fifth of participants reported that they understood co-occurring in a temporal manner and just over 10% of responses indicated that they understood the primary task as what they were doing, with no consideration of the secondary task. A similar proportion of responses indicated a need for further clarification. For instance:

Participant understands co-occurring as focus of attention

"in that case it would depend on whether the radio was on in the background or washing up was the purpose or if I was listening to a radio program and happened to be washing up at the same time"

"usually when I'm at home I've got the radio on or the television on for company so obviously if I was doing something else then listening to the radio or watching something would probably be the secondary activity"

Participant understands co-occurring in a temporal manner

"Secondary activity I think was my laptop, [because] I did it after"

"Quite hard as we multi-task all the time and flip from one to another."

Participant understands co-occurring as primary task with no consideration of secondary task

"The main activity was what I was actually doing at the time"

"I haven't really got any main activity, its all sorts of things, it just depends on the day. It was fairly easy deciding what I was doing"

Further clarification needed

"secondary activities … I didn't quite understand what that meant"

Activity well-being ratings

Easiest activities to rate

Participants were asked about the activities from yesterday they felt would be easiest to rate. Those given satisfied diaries considered house repair/DIY/household tasks, relaxation/leisure and gardening among the easiest activities to rate. These participants explained that these activities were easiest to rate for many reasons, including due to task factors and novelty factors:

"It was a simple confined task"

"I got it done and I was pleased with what I'd done."

"[Because] it was a defined task"

"It's something I don't do very often"

Those participants who completed enjoyed diaries considered family time and novel tasks to be easiest to rate. Some of these participants reported that family time was easiest to rate because it is unambiguous and because they considered pleasurable time to be particularly easy to rate:

"Because it's so clear cut and a nice thing to do"

"Probably being on the boat and the general overview of the day. It's when you start to think of the negative side, that's hard, so the pleasurable things are much easier to rate"

Among both groups, family time and paperwork/correspondence were considered easiest to rate. One participant given an enjoyed diary, considered the level of satisfaction an activity provided as important in determining how easy it is to rate the enjoyment of that activity:

"Found it easier to rate the more satisfying activities"

*(*Interviewer: Why would that be?*)*

"You're feeling happier and things are going better"

Most difficult activities to rate

Participants were also asked about the activities from yesterday they felt would be most difficult to rate. Over a third of those given satisfied diaries reported that ordinary activities (e.g. personal hygiene, walking, cooking, commuting) were most difficult to rate because they are so routine and mundane:

"its not that memorable and it's something you do every day"

"Doing to the vacuuming. I mean how do you rate doing the vacuuming?"

"I think the mundane things are the most difficult to rate"

"Probably quite difficult to rate going for a walk I suppose."

Over 15% of these participants considered no activities difficult to rate and one of these participants misunderstood the question:

"I don't think any really. I don't any were dissatisfactory at all. No no no I wouldn't have found anything difficult to rate."

About a quarter of those participants who completed enjoyed diaries also reported that ordinary activities were most difficult to rate because they were mundane and just something they "had to do":

> *"They have to be done and not particularly pleasurable...A bit mundane"*

> *"Probably driving to work, I don't not enjoy it.... It's something I have to do. So, it's a, it's a, necessary evil"*

> *"Yeah because you don't think about what you're cooking, you just do it. It's just something I do"*

Other activities those given enjoyed diaries found difficult to rate included frustrating and time constrained activities:

> *"Filling out and completing forms against the time scale".*

Classification of "middle of the road" activities

Those participants who were given satisfied diaries suggested a long list (with little agreement) of activities which were not their most or least satisfying, but housework or no activities featured more prominently than others (e.g. commuting, leisure/exercise, socialization, entertainment/TV).

Most of those given enjoyed diaries suggested housework and entertainment related activities could be considered not most or least enjoyable:

> "*watching telly in the evening.*"

> *"playing games on my laptop"*

Further probing revealed that for many participants it was the novelty value of the tasks that made it particularly difficult to answer this question:

> *"Quite difficult actually, cause it's just things that you do every day – you don't normally rate it – you just sort of do it"*

> *"I think difficult for that particular answer.... Because there's nothing of particular note"*

> *"Quite hard to rate that experience as quite an ordinary event watching a film"*

Participants' understanding of well-being scale points

At various points throughout the cognitive interview, participants who had completed satisfied and enjoyment diaries were reminded of the well-being scale in their diary and asked to provide ratings of various aspects of yesterday's activities. Later in the interview, they were also asked to provide ratings on the scale piloted with the other group, with which participants were unfamiliar. Furthermore, they were also directly asked specific questions about particular points on the scales that ranged from 1 (not at all) to 7 (very much), such as *"what does a 3 mean to you?"* or "*can you tell me the difference between a 5 and a 6?".*

Responses to all these questions as well as any further instances, in which participants related a description to a number on satisfied and enjoyed scales, were collated and coded into short descriptor terms.

As part of an iterative analysis process, these descriptor terms were categorized as representing different potential factors or "elements/components" of enjoyment and satisfaction for this small sample of diary study participants, who had, at the time the cognitive interviews were conducted, actively engaged in the process of attaching ratings to these emotional states, in two separate and temporally-spaced research studies (i.e. the diary study and the cognitive interview).

Figures 16 and 17 below show the different factors that were associated with the numerical labels on the satisfied and enjoyed scales respectively. They show that there were four satisfied factors and five enjoyed factors that represented five or more continuous points on the respective scales. Other factors represented scale points on a more intermittent basis and only the satisfied scale has sufficient factor coverage to represent every point from 1 to 7 on the scale. The tables in appendices O and P show these factors as well as specific example descriptor terms suggested by participants as representing the various points on the satisfied and enjoyed diary scales.

Table 4: Showing a summary of factors associated with different points on the satisfied scale

Factors associated with different scale points

Satisfied scale points						
1			Satisfaction factors			
2		Individual factors		Situational factors	Enjoyment factors	
3	Task factors	Individual factors	Satisfaction factors	Situational factors	Enjoyment factors	Ordinal factors
4	Task factors	Individual factors	Satisfaction factors			Ordinal factors
5	Task factors	Individual factors		Situational factors	Enjoyment factors	Ordinal factors
6	Task factors	Individual factors	Satisfaction factors	Situational factors	Enjoyment factors	
7	Task factors		Satisfaction factors	Situational factors	Enjoyment factors	

Source: Time Use feasibility study, Office for National Statistics

Table 5: Showing a summary of factors associated with different points on the enjoyed scale

Factors associated with different scale points

Enjoyed scale points						
1						
2	Enjoyment factors	Situational factors				Individual factors
3	Enjoyment factors	Situational factors	Task factors	Ordinal factors	Satisfaction factors	Individual factors
4	Enjoyment factors	Situational factors	Task factors	Ordinal factors	Satisfaction factors	
5	Enjoyment factors	Situational factors	Task factors	Ordinal factors	Satisfaction factors	Individual factors
6	Enjoyment factors	Situational factors	Task factors	Ordinal factors		
7	Enjoyment factors	Situational factors	Task factors		Satisfaction factors	

Source: Time Use feasibility study, Office for National Statistics

Some of these (or similar) descriptor terms represented identical points on both scales. Table 6 below shows that similar satisfaction-related terms were suggested by participants to represent points low and high points on both scales, similar enjoyment-related terms and task factors were suggested to represent low, mid and high points on both scales and similar ordinal terms were used to represent mid-points on both scales.

Table 6 shows that task non-completion/achievement was used to represent the mid-point on both scales rather than scale points below this. It is also interesting that routine task-level factors denoted 1 scale point above the mid-point (i.e. scale point 4) on both scales, and descriptor terms suggesting "a necessary task" were used to represent 1 scale point below the mid-point on both scales.

Although extremely preliminary, and based on a small, non-representative sample, the similarities outlined in table 6 suggest that these terms for the points they aligned with on both scales, may be indicative of more global well-being scale point labels. However, further research would be required to explore this in more depth with a more robust, methodological approach. It would certainly be interesting if future research determined whether such terms were aligned with the same scale points for further measures of well-being, beyond enjoyment and satisfaction.

Table 6: Showing the same or similar descriptor terms used to describe identical points on both enjoyed and satisfied scales

General factor	Specific factor	Scale points represented	Enjoyed term	Satisfied term
Satisfaction	Satisfaction	3	*"Dissatsified"*	*"Dissatisfying"*
		7	*"Satisfying"*	*"Always satisfied"*
				"Continually satisfied"
Enjoyment	Enjoyment	3	*"Not enjoyed"*	*"Not enjoyed"*
			"Not happy"	*"Unhappy"*
				"Not particularly happy"
		5	*"Quite enjoyed"*	*"Quite happy"*
				"Chuffed"
		7	*"Really enjoyed"*	*"Enjoy task"*
			"Enjoyed that"	
Ordinal	Ordinal	4	*"Average"*	*"Average"*
			"Neither"	*"Either"*
				"Normal"
				"Neutral"
Task	Outcome	4	Task non-completion	Task non-completion
			Task non-achievement	Task non-achievement
		7	Task completion	Task completion
			Task achievement	
	Timing & Novelty	3	*"Necessary task"*	*"Something has to be done"*
			"Routine"	*"Routine"*

Source: Time Use feasibility study, Office for National Statistics

1. Quotes in italics in the examples column denote direct quotes from participants on the meaning of scale points. Those examples not in quotes were derived by the research team from participants' responses.

In a range of comments, participants also expressed frustration and confusion in assigning values to psychological attributes using rating scales like these:

"I can't believe you're asking me this! There is no difference. You can't possible take a complex human reaction and decimate it between individual numbers. It's not possible. I can tell you the difference between 1 and 7, but not between 5 and 6. Nobody can tell you tell, and anyone who is is just pretending.

"I don't understand why someone thought this was a good question to ask?!"

"But if you rated something 5 and it was really a 6 it doesn't really matter. In all these surveys asking to rate 1 to 5, 1 to 10. But you can't dissect to the nth degree, you've just got to say "on that scale, on that particular occasion, it was a 6".

"I think of 5 as being medium, but when I say that I'm thinking of it being between 1 and 10, really truly medium should be 3.5 shouldn't it, so therefore I have miscalculated it. I believe that is I was to see this written down my mind would say 3.5 but if I don't see anything I might automatically go for 5 being medium"

Other miscellaneous comments from participants suggested further preferences and individual differences in approaching rating scales:

"Nothing I do was worth giving a 1"

"I don't know if there's any activity I would absolutely rate a 7 so I think 6 for me is probably as good as it gets and 5 is further down the scale"

Such frustrations demonstrated by the participants highlight how arbitrary they feel assigning meaning to a number on a scale. It suggests a need for further research that engages with respondents, as well as careful consideration of the methods used in time use research to elicit emotional reactions to daily activities.

Rating changes depending on context rather than task

Participants who completed satisfied diaries reported a number of factors that could cause their satisfaction ratings to change (although one participant suggested this did not happen because they don't experience good/bad days). The most frequent reasons cited for this included the type of activity, individual characteristics and situation/context:

"I just found the sort of mundane things it probably depends on how I was feeling at the time"

"I'd probably just state it on the feelings at the time"

"That might have made me feel better if I'd been wide awake. I'd probably feel better if I'd had a cup of tea"

"Your well-being on the day affects the task. You can be doing the same task and personal circumstances around you will affect it. Some days you can do a task and enjoy it, other days it will be a chore. It is what it is on the day".

Most participants who completed enjoyed diaries also indicated their enjoyment ratings could change. The most popular reasons cited included environmental conditions and individual characteristics (e.g. mood). A minority of those given enjoyed diaries suggested that their enjoyment of certain activities would remain the same irrespective of the type of day. For example, one participant cited the novelty value of activities as related to this:

"I think washing up is always [going to] be a 5, even with music on or having somebody help you, it's still gonna be a chore"

Comprehension of well-being scale wording

Those participants who completed satisfied diaries were asked what they thought satisfaction meant to them. A range of definitions were provided but the most popular definitions included task factors (i.e. achievement/completion), emotional characteristics related to enjoyment and outcome factors:

"Having achieved a little bit"

"Satisfactory is that you're happy for other people to look at it and say yes you've done an ok job, if someone did something I would look at it and think its met the persons abilities to do that"

"You're happy with what you've done"

"Well happiness really, something which has gone well and given me pleasure so happiness is the nearest thing really"

"You don't think, I should've done this. I should've done that..."

"Day went well and no dramas"

Those participants who completed enjoyed diaries were asked what they thought enjoyment meant to them. A range of definitions were provided but the most popular definitions included task factors and emotional characteristics related to enjoyment:

"Not necessarily a difficult task or even if it is its one that you've got the ability to achieve without too much difficulty so even if its painting a painting or DIY if you've got the skills to do that, if it comes to you easily then you would derive enjoyment from it, so something that, even if it was challenging you could still achieve a big result and would bring enjoyment"

In a range of comments, participants also expressed puzzlement, confusion and bewilderment in their comprehension of the diary well-being scales:

"This is the trouble with surveys like this, you're trying to put a number to something that's very complex"

"I mean, satisfied, does that mean the housework's been done properly and you've cooked a meal properly, or does it mean you'd really rather do something else?"

"I'm getting too philosophical. You can't put philosophy into numbers, can you?"

Similarities and differences between enjoyment and satisfaction

In their responses to multiple questions, some participants indicated the ways in which the concepts of satisfaction and enjoyment were similar or different to one another.

For instance, some participants indicated that satisfaction was a precursor to enjoyment and that enjoyment was a precursor to satisfaction:

"If you're experiencing enjoyment then you're satisfied"

"I am satisfied then there's enjoyment there"

However, other participant responses suggested no such association:

"You can enjoy a drink, but a drink doesn't satisfy you"

"You go get a suit made you're satisfied with the suit, you don't enjoy the suit"

"Because I was satisfied having completed the tasks yesterday, but in no way, can you call that enjoyment. So satisfied was getting routine and the tasks of the day sorted out, but in no way, can that be classified as enjoyment."

Additionally, in some cases, participants who completed satisfied diaries explicitly referred to aspects of enjoyment when providing their responses.

"I'm a very calm and even person, I don't get wildly excited about things or get too depressed at anything."

"I'd probably go for a middle rating but that wouldn't be exactly right because neither happy or unhappy is not it doesn't probably cover it"

"felt happy"

Similar and opposite well-being terms

Figures 14 to 15 below show synonyms of well-being terms suggested by participants[13]. They show that the most frequently cited synonyms for satisfied were nice/good and adequate/reasonable, and the most frequently cited opposite term for satisfied was insufficient/problem. The figures also show that a range of synonyms for enjoyment were cited with equal frequency.

Figure 14: Showing participant-suggested synonyms for satisfied

England and Scotland, 2017
Percent (%)
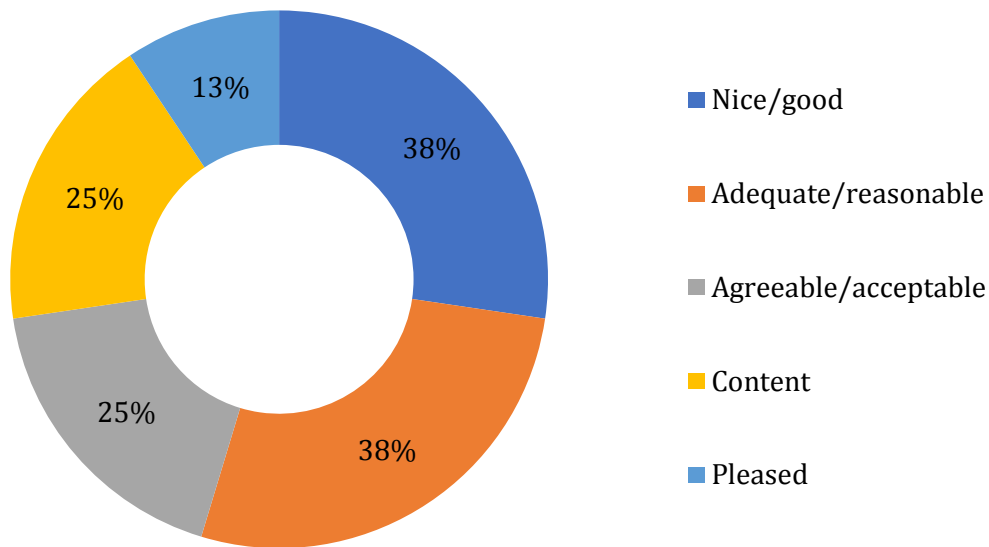
Source: Time Use feasibility study, Office for National Statistics

---

[13] Percentages may add up to greater than 100% because some participants suggested several terms.

Figure 15: Showing participant-suggested synonyms for enjoyed

England and Scotland, 2017
Percent (%)
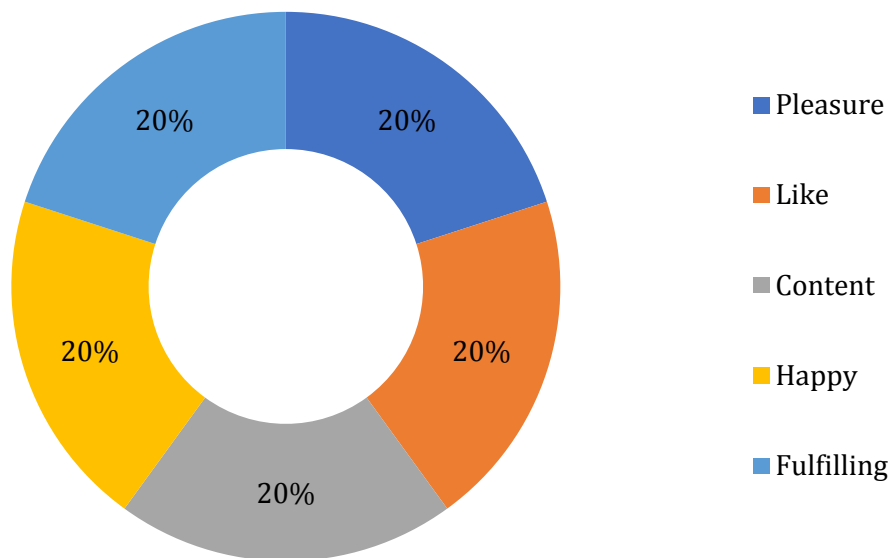
Source: Time Use feasibility study, Office for National Statistics

There is less variation in the terms respondents gave to represent enjoyment compared to satisfied. This indicates that there may have been less agreement in understanding of the word "satisfied" compared to "enjoyed".

Rating the alternative well-being scale

Those participants who completed enjoyment diaries were asked what they thought satisfied meant to them. A range of definitions were provided but the most popular definitions included task factors (for example achievement/completion/process) and output factors (for example done enough/average).

> *"I wouldn't particularly say I enjoyed it but I am satisfied that I have done it and completed it".*

> *"Satisfied to me is what I would call a normal day".*

Those participants who completed satisfied diaries were asked what they thought enjoyment meant to them. A range of definitions were also provided here but the most popular definitions included emotional connotations and indications that enjoyment meant "*something more*" than satisfaction:

> *"Enjoyed is a bit above that, you get that little extra smile and it's beyond what you expected"*

> *"Got far more range if you had used enjoyment, boring and fun"*

> *"Satisfied really didn't do it for me as a scale, but I enjoyed doing the diary"*

About half the participants who completed a satisfied diary indicated that they would find it easier to rate their enjoyment than their satisfaction:

"I'd say satisfaction is slightly harder not by a lot just requires more thought"

1 in 3 participants who completed the satisfied diaries suggested there would be no difference in rating ease/difficulty or that they were such different concepts so none was harder or easier than the other:

"They're too different…. You can't substitute one for the other. You would have to ask two questions. They're not the same".

"And use the word enjoyment instead of satisfied…would depend on what the activities were that I had done"

"Um, not more difficult just different"

Just under half of those participants who completed an enjoyed diary considered that satisfaction would be harder to rate. Just under a quarter of these participants thought both concepts would be equally difficult to rate and the same proportion considered the concepts too different to be able to determine this:

"Harder to rate, because enjoyment is such a positive feeling it is easier to recognize and differentiate. Whereas satisfactory could be quite average"

"I think it would because satisfied could be a bit confusing"

"Rating satisfaction would be much harder than enjoyment. Enjoyment is straightforward in my view. Satisfaction has to be achieved"

Yesterday's activities

These accounts were classified by raters as exhibiting good, moderate or poor recall of yesterday's activities. Over 3 in 4 of participants provided a good level of detail concerning yesterday's activities, just over a fifth or yesterday's activities were recalled in moderate detail and just under a fifth were recalled at a poor level of detail.

An example of good recall of yesterday's activities

Most of the day, well I got in from work just after 6am, 6.20 then I had breakfast. [I] went to bed about 7am in the morning. Then I sleep from about 7am till 1pm in the afternoon, then getting washed and dressed for – that's about half an hour then lunch half past 1 till 2. Then watching television for an hour and then, household chores for about 15 minutes, that takes me up to about quarter past 2. Then chatted with another member of the household for about half an hour, then from about quarter to 3 to about half past 6 I was doing some exterior decorating on the house. Then from half past 6 – 7 I had a cup of tea then biscuits then I watched a bit of [television] for about half an hour to about half past 7. Then washing up for about quarter of an hour, then from 8.15 until. Until quarter to, till about half past 9 I watched [television] and quarter to 10 or 9.45 I drove to work. And then from 10pm till midnight – working.

An example of moderate recall of yesterday's activities

"I was out quilting from 10 o'clock in the morning to 3 o'clock in the afternoon, so that was my main activity, nearly all day…..You set your sewing machine, then get coffee or dinner and carry on".

An example of poor recall of yesterday's activities

"I would include that yesterday was spent doing emails, paperwork and forms."

The level of detail in recalled activities does have potential implications for the need for future research concerning the selected method for conducting time use research. As detailed in the literature review, a methodological difference between the HETUS and ATUS instruments concerns the manner in which time use data is collected. The time diaries in this study followed the HETUS approach, but the ATUS approach also uses a telephone interview to ask about yesterday's activities. While ATUS is undeniably robustly designed, with methodological strengths to address any potential limitations (e.g. interviewers specifically experienced in eliciting required information), our very limited experience in this study may indicate that while this is a potentially useful methodological approach for most respondents, some may fail to provide a sufficient level of detail in their responses. It would certainly be interesting for future research to examine this issue in greater detail, comprising for a start, an experimental approach, larger sample sizes and varying levels of interviewer experience.

Recall of well-being during yesterday's activities

To familiarize participants with the task of recalling the emotional associations with, and implications of, yesterday's activities, they were first asked to describe how they felt yesterday, more generally.

More than 3 in 4 participants reported feeling positive yesterday with only about a 1 in 4 reporting that they felt negative, neutral or both positive and negative.

Examples of positive emotional recall of yesterday's activities:

*"Quite happy actually as I was doing something at my desk that I was quite enjoying. The journey home was quiet, not too bad for a change and it's always nice to come home to my husband and a bit of time talking to my Mum. So yeah yesterday was a happy day". (*enjoyed diary*)*

*"Well I got quite a bit done, so I was pleased" (*enjoyed diary*)*

"*Particularly satisfied because it was beautiful paint and it did a wonderful job that I wasn't expecting" (*satisfied diary*)*

"*Satisfied*" – (enjoyed diary)

Example of positive and negative emotional recall of yesterday's activities:

"Getting ready to go in the morning was hard work and not so pleasant as you try to organize children but the actual day once we got there and got sorted, was extremely pleasant so two ends of a spectrum"

Satisfaction and enjoyment of yesterday's activities

As outlined in the literature review, the utility of asking questions about extreme responses has previously been challenged by some authors in the emotion and well-being research community. They argue that such questions are more likely to "tap" in to cognitive skills (e.g. judgment and decision making), rather than truly measuring emotional state. However, for these participants, several months might have elapsed between the date at which they completed their diary and the date of their cognitive interview. Therefore, despite reasonable objections concerning the validity of these types of questions, participants in this study were first asked to consider yesterday's most and

least satisfying/enjoyable activities to further familiarize participants with the process of thinking about activities within an emotional context.

Most satisfying activities yesterday

Participants were asked about their most satisfying activities yesterday. They mostly referred to specific activities, with the most popular including house/garden repair and housework:

"Definitely defrosting the freezer."

Other popular satisfying activities included food preparation/consumption, office work/paper work and hospitality/socializing. However, without prompting, many of those participants who completed satisfied diaries listed various task-relevant factors (e.g. achievement, completion, objects/materials, skill learning/maintenance) in providing their answers to this. For example:

Office/paperwork: *"Completing the paper work for the solicitors. Getting that out the way"*

Gardening: *"Getting the bits and pieces in the garden sorted".*

House repair: *"Well I achieved what I set out to do on the decorating I was pleased with it"*

The participant above also expressed enjoyment-related adjectives in relating satisfaction to interviewers.

Participants were asked what aspects of their activities made them particularly satisfying. 70% of these participants explicitly referred to task factors (e.g. task completion, achievement) as components of satisfying experiences, suggesting these played a key role in activity satisfaction for these participants:

"I made some breakthroughs in learning a new piece of software so that was quite exciting so learning new skills, didn't really learn any new skills just tried to polish up some really bad ones"

A fifth of these participants referred to the importance of contextual factors such as environment/setting and other people, in contributing to their satisfaction with yesterday's activities:

"Probably all having lunch together..... I suppose because I got a grandchild that's 12 weeks old and a grandson that's two years old and they're just lovely"

A fifth of these participants also referred to individual-level factors (e.g. motivation, time for self) in contributing to their satisfaction with these activities. For example:

"Just haven't had the time to do it. And yesterday was one of those days yes well we'll get stuck in and get sorted"

Least satisfying activities yesterday

Participants were asked about the activities they felt least satisfied with yesterday. Participants mostly referred to specific activities, with the most popular including nothing, work/schoolwork, food preparation/household tasks, crises and watching TV.

The reasons participants felt least satisfied included the time of day, task factors (i.e. achievement/unpleasantness, frequency) and individual factors (i.e. attention, motivation, frustration):

"It's frustration of not knowing what's going to happen really and not knowing what you can do about it because you've just got to wait."

Activity enjoyment yesterday

Participants were asked about the activities they enjoyed the most yesterday. They mostly referred to specific activities, with the most popular including work and necessary tasks:

"I was also at work today and I do enjoy my work, yes work was good"

Other activities participants reported enjoying the most included family time, hobbies/ voluntary work and seeing friends:

"Using my sewing machine, which I haven't been able to use for a long time"

"Seeing my friend and having a chat"

Some participants also referred to factors related to the frequency with which they engage in such activities, suggesting that for these participants novel, infrequent, outside-of-the-norm activities were particularly enjoyable:

"Designing a form template, its just one of those mundane tasks that I don't have to think about and quite enjoy just getting on with quietly – it makes a change from having to deal with people and going to lots of meetings and being pestered for information, it was nice to have a few hours peace and quiet"

"An experience you wouldn't normally have otherwise"

Participants were asked about the activities they enjoyed the least yesterday. Responses included housework/DIY, task factors (i.e. activity preparation, fixing errors) travelling, paper work, using technology and getting up:

"Having to correct some mistakes that I'd made"

The reasons for not enjoying these activities included individual factors (for example frustration, motivation) and task factors (for example task skill, task difficulty and task completion).

Location

Just over a third of participants reported that recording their location was an easy/straightforward task and nearly 3 out of 4 participants reported that this was because their responses simply reflected where they were at that time, or that they were at home. However, a third of comments suggested participants required further instructions, that they were confused with or in some way misunderstood this diary section. One participant expressed a degree of confusion or frustration over repeatedly being required to specify location, despite staying in the same place throughout the day:

"It didn't seem overly relevant as on the day I was staying in one spot, it was just repeating myself throughout the day"

However, the main issues reported, for this aspect of the diary instrument concerned the level of specificity required for this diary section:

"Just put at home [because] I live in not a big house but a substantial house and I don't always sit in the same room"

"I don't know, should I have specified a room, rather than just at home?"

"When I was in bed obviously I thought do I just put at home or do I put in bed if I was in bed but I wasn't quite sure on that how much personal detail I should go into"

"Would just say "pub", for pub quiz, I can't remember if I said the "[insert specific pub name and town name]" for example. So it might need to tell you whether you need to be more vague or a bit more specific."

Conclusion

Few respondents commented on difficulties filling in the diary before follow-up questions about difficulties faced. Respondents had different interpretations of how to assign a main or secondary task, there was no consensus or majority approach for how tasks were assigned. Respondents felt that more space for activities in the diary would have helped them record more information. Co-occurring (main vs secondary) tasks was found to be the area of most confusion for respondents.

Enjoyment and satisfied were perceived to be different terms, with differing connotations for how respondents rated their activities. Participants proposed a range of potential labels for numerical points on both scales that broadly fit within 6 different factors. Several of the same or similar descriptor terms were also proposed to represent identical points on both scales, particularly for low, mid and high scale points. These may have utility as scale point descriptors for more general measures of well-being (e.g. those with satisfaction and enjoyment dimensions).

The results should be used as indicators for improvements to be made to the 2020 HETUS, rather than definitive as this analysis is explorative. Nevertheless, the results provide some evidence for enjoyment as the measure of choice for well-being in HETUS. Further analysis of words similar to enjoyment to represent well-being is recommended.

Recommendations

HETUS 2020

1. The literature suggests that the well-being element of time use surveys produces results policy-makers find more directly relevant. The literature also indicates that respondents find this element of the survey particularly relevant in recent time use survey implementations. This dimension would be used widely if collected, adding urgency for adding this field.
2. An increasing number of national time use surveys incorporate a well-being element. HETUS 2020 remains cutting edge by including this component in diary design recommendations. Missing this element out misses a major advantage of the effort and cost of conducting a time use survey.
3. The concept of enjoyment should be used to measure well-being in HETUS 2020 as it was more meaningful to respondents than the concept "satisfied". The enjoyment column produced a wider range of responses across both activities and context variables (mode of transport, location, who else was present), produced more variation across days of the week and times of day, and produced more policy-relevant activity distinctions than asking about satisfaction. Alternative terms reviewed by this project were deemed less appropriate to a multi-purpose survey capturing a range of national statistics.
4. Respondents in the pilot diary and the cognitive interviews reported finding the satisfaction as opposed to enjoyment more relevant for reporting how they experienced routine housework, commuting and daily living tasks. People may not enjoy doing the ironing, brushing their teeth, or cleaning the bathroom, but they appreciate that such a task was performed well (or feel less happy if the task was not done well or they did not have time to properly complete the task). While this point is worth noting in terms of the limitations of choosing a single well-being measure, and this knowledge is relevant to understanding the experience of housework and other routine activities, the more blunt the information captured about the range of other activities, context of activities, and times of day make this a less useful option for measuring well-being. Capturing satisfaction is not relevant to measuring the volume and assigning economic value to the range of unpaid production activities more often performed by women, and contributes only modestly to this central function of time use surveys.
5. 1-7 uni-polar scales provide respondents with flexibility of various ratings whilst remaining generally comparable to other means of well-being collection (e.g. survey questions on well-being).
6. Diaries containing full scales with numbers that respondents can circle or otherwise mark allow respondents to visualize the scale to compare ratings. Enabling visualisation of scales will be easier once diaries are collected online or via apps as space restrictions would be less applicable.
7. For the period where paper diaries remain a mode of diary collection, a more flexible who else was present column with a tick box for alone and a box to fill in others present reduces the complexity of the diary grid and could create more space for an enjoyment section. The simplified, open-ended self-completion box captures a wider range of persons present than the existing HETUS design, and may prove easier to complete as respondents record these categories of people who are relevant to their daily activity account, rather than trying to make sense of the HETUS categories if these are not applicable.

8. Our effort to pilot a reduced day diary was not successful. Multiple respondents tried to fit in activities which clearly took place over the whole day into the 6 one-hour blocks. A limited consecutive observation window which allows respondents to fill in a complete story about a part of their day might prove easier to complete than a selection of hours over different times of the day.

9. Further testing of understandings and alternative definitions, conceptualizations, synonyms and opposites of enjoyment to:
    - Examine their relative suitability/appropriateness as alternative measures of enjoyment or as complementary measures as part of a multi-dimensional enjoyment scale
    - Examine their association with particular activities or activity types with a view to explain the using different terminology to explain enjoyment in different activity contexts
    - Particular consideration will need to be made to the concept that best translates across European languages for the 2020 HETUS survey.

References

Abraham, K., & Mackie, C. (2005). Beyond the market. *Washington, DC: National Academy of Sciences.*

Abraham, K., Maitland, A., & Bianchi, S. (2005). Nonresponse in the American Time Use Survey: Who is missing from the data and how much does it matter? University of Maryland, ATUS Early Results Conference.

Abraham, K. G., Maitland, A., & Bianchi, S. M. (2006). Non-response in the American time use survey: who is missing from the data and how much does it matter?

Al Baghal, T., Belli, R. F., Phillips, A. L., & Ruther, N. (2014). What are you doing now? activity-level responses and recall failures in the american time use survey. *Journal of Survey Statistics and Methodology*, *2*(4), 519-537.

Andrews, F.M. and Withey, S.B. (1976). Social Indicators of Well-Being , Plenum Press, New York.

Apps P 2002. Gender, Time Use, and Models of the Household. Faculty of Law, University of Sydney

Argyle, M., & Crossland, J. (1987). The dimensions of positive emotions. *British Journal of Social Psychology*, *26*(2), 127-137.

Aslaksen, Iulie (Statistics Norway) and Koren, Charlotte (Institute of Applied Social Research), "Unpaid Household Work and the Distribution of Extended Income: The Norwegian Experience," Feminist Economics 2 (3), 1996, pp. 65-79.

Baron, N. (2008), Adjusting the volume – Technology and multitasking in discourse control, in: Katz, J.E. (eds.), Handbook of mobile communication studies, Cambridge, MA, MIT, 117-194.

Bediako, G., & Vanek, J. (1999). Trial international classification of activities for time use statistics. In *Time use-research, data and policy: Contributions from the international conference on time use (ICTU), FFB-Schriftenreihe* (Vol. 10, pp. 151-161).

Bevans, G. E. (1913). *How workingmen spend their time*(Doctoral dissertation, Columbia University).

Bhat, C.R. and R. Misra (1998) Discretionary activity time allocation of individuals between inhome and out-of-home and between weekdays and weekends, submitted to Transportation for possible publication in a special issue on Time Use.

Bhat, C. R., & Koppelman, F. S. (1999). A retrospective and prospective survey of time-use research. *Transportation*, *26*(2), 119-139.

Bianchi, S. M., Robinson, J. P., & Milke, M. A. (2006). *The changing rhythms of American family life*. Russell Sage Foundation.

Bianchi, Suzanne M. and Vanessa R. Wight. (2010). "The Long Reach of the Job: Employment and Time for Family Life." Pp. 14–42 in Workplace Flexibility: Realigning 20th-Century Jobs for a 21st-Century Workforce, edited by K. Christensen and B. Schneider. Ithaca, NY: Cornell University Press

Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to survey quality* (Vol. 335). John Wiley & Sons.

Binder, M., & Freytag, A. (2013). Volunteering, subjective well-being and public policy. *Journal of Economic Psychology*, *34*, 97-119.

Bittman, M. and Wajcman, J. (2000). "The Rush Hour: The Character of Leisure Time and Gender Equity." Social Forces 79:165–89.

Boarini, R., Comola, M., Smith, C., Manchin, R., & De Keulenaer, F. (2012). What makes for a better life?: The determinants of subjective well-being in OECD countries-Evidence from the Gallup World Poll. *OECD Statistics Working Papers*, *2012*(3), 0_1.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual review of psychology*, *54*(1), 579-616.

Bonke, J. & Fallesen, P. (2010) "The Impact of Incentives and Interview Methods on Response Quantity and Quality in Diary - and Booklet-Based Surveys" *Survey Research Methods* 4(2): 91-101.

Bonke, J. (2005). Paid work and unpaid work: Diary information versus questionnaire information. *Social Indicators Research*, *70*(3), 349-368.

Brancato, G., Macchia, S., Murgia, M., Signore, M., Simeoni, G., Blanke, K., & Hoffmeyer-Zlotnik, J. (2006). Handbook of recommended practices for questionnaire development and testing in the European statistical system. *European Statistical System*.

Budig, M., & Folbre, N. (2004). Child Care vs. Child-Minding: Measuring Activities, Responsibilities, and Time. *Family Time: The Social Organization of Care*, 51-68.

Budlender, Debbie, "The Policy Implications of Time Use Surveys: Lessons from South Africa," 2002, unpublished manuscript.

Burgess, P. W. (2000). Strategy application disorder: the role of the frontal lobes in human multitasking. *Psychological research*, *63*(3), 279-288.

Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, *38*(6), 848-863.

Burgess, Paul W., Iroise Dumontheil, Sam J. Gilbert, Jiro Okuda, Marieke L. Schölvinck, and Jon S. Simons. (2008). "On the Role of Rostal Prefrontal Cortex (Area 10) in Prospective Memory." Pp. 235–60 in Prospective Memory: Cognitive, Neuroscience, Developmental and Applied Perspectives, edited by M. Kliegel, M. McDaniel, and G. Einstein. Mahwah, NJ: Erlbaum "Canadian Women – Making an Impact," Women's History Month Newsletter, Status of Women Canada, 1998.

Chatzitheochari, S., Fisher, K., Gilbert, E., Calderwood, L., Huskinson, T., Cleary, A., & Gershuny, J. (2015). *Measuring young people's time use in the UK millennium cohort study: A mixed-mode time diary approach*. CLS Working Paper 2015/05.

Chatzitheochari, S., Fisher, K., Gilbert, E., Calderwood, C., Huskinson, T., Cleary, A. & Gershuny, J. (2015) Measuring Young People's Time-Use in the UK Millennium Cohort Study: A Mixed-Mode Time Diary Approach CLS Working Paper 2015/05, London, UK: Centre for Longitudinal Studies, Institute of Education.

Chenu, A., & Lesnard, L. (2006). Time use surveys: a review of their aims, methods, and results. *European Journal of Sociology/Archives Européennes de Sociologie*, *47*(3), 335-359.

Cloud, Kathleen and Garrett, Nancy, "A Modest Proposal for Inclusion of Women's Household Human Capital Production in Analysis of Structural Transformation, "Feminist Economics 2(3) 1996, pp. 93-119.

Craig, L. (2006). Does father care mean fathers share? A comparison of how mothers and fathers in intact families spend time with children. *Gender & society*, *20*(2), 259-281.

Craig, L. (2007). Is there really a second shift, and if so, who does it? A time-diary investigation. *Feminist review*, *86*(1), 149-170.

Craig, L., & Mullan, K. (2010). Parenthood, gender and work-family time in the United States, Australia, Italy, France, and Denmark. *Journal of Marriage and Family*, *72*(5), 1344-1361.

Craig, L., Mullan, K., & Blaxland, M. (2010). Parenthood, policy and work-family time in Australia 1992—2006. *Work, Employment and Society*, *24*(1), 27-45.

Dasgupta, P. (2001). *Human well-being and the natural environment*. Oxford University Press.

DeMaio, T. J., & Rothgeb, J. M. (1996). Cognitive interviewing techniques: In the lab and in the field.

Diener, E., & Larsen, R. J. (1993). The experience of emotional well-being.

Diener, E. (1994). Assessing subjective well-being: Progress and opportunities. *Social indicators research*, *31*(2), 103-157.

Diener, E., & Fujita, F. (1995). Resources, personal strivings, and subjective well-being: a nomothetic and idiographic approach. *Journal of personality and social psychology*, *68*(5), 926.

Diener, E., Wolsic, B., & Fujita, F. (1995). Physical attractiveness and subjective well-being. *Journal of Personality and Social Psychology*, *69*(1), 120.

Diener, E., Suh, E., & Oishi, S. (1997). Recent findings on subjective well-being. *Indian journal of clinical psychology*, *24*, 25-41.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological bulletin*, *125*(2), 276.

Dolan, P., Peasgood, T., & White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of economic psychology*, *29*(1), 94-122.

Drago, R. (2011). *Secondary Activities in the 2006 American Time Use Survey* (No. 446).

Drago, R. W., & Stewart, J. C. (2010). Time-use surveys: issues in data collection on multitasking. *Monthly Labor Review*, *133*(8), 17-31.

Durkheim, É. (1899). Une révision de l'idée socialiste. *Revue philosophique*, *48*, 163-172.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. *Nations and households in economic growth*, *89*, 89-125.

Elson, D. (2002). Gender justice, human rights and neo-liberal economic policies. *Gender justice, development and rights*, 78-114.

Eurostat. (2001), Guidelines on harmonised European time use surveys, https://www.testh2.scb.se/tus/tus/doc/Circa%20Principal%20Document.doc, (accessed 16 June, 2009).

EUROSTAT (2009), Harmonised European Time Use Surveys – 2008 Guidelines, EUROSTAT Methodologies and Workingpapers, KS-RA-08-014-EN.pdf, http://epp.EUROSTAT.ec.europa.eu/portal/page?_pageid=1073,-46587259&_dad=portal&_schema= PORTAL&p_product_code=KS-RA-08-014

Ferring, D., & Filipp, S. H. (1996). Measurement of self-esteem: findings on reliability, validity, and stability of the Rosenberg Scale. *Diagnostica*, *42*(3), 284-292.

Fisher, K., & Layte, R. (2004). Measuring work-life balance using time diary data. *electronic International Journal of Time Use Research*, *1*(1), 1-13.

Fisher, K., & Gershuny, J. (2013). *Time use and time diary research*. Oxford University Press.

Fisher, K., Gershuny, J., Glorieux, I., Minnen, J., Paäkkönen, H., Pantelic, V. & Wennekers, A. (2016). Future of time-use research in Europe - Pilot Group 7: Designing experiments for calibration: cross-mode, cross-country, and cross-time comparability. 12-01-2016, first working report.

Fisher, K., Cappadozzi, T., Castell, L., Papic, D.D., Gershuny, J., Kabadayi, A., Paäkkönen, H. & Payne, C. (2016). Future of time-use research in Europe - Pilot Group 6: New dimensions in time diaries. 10-01-2016 - first working report.

Fisher, K., Chatzitheochari, S., Gilbert, E., Calderwood, L., Fitzsimons, E., Cleary, D., Huskinson, T. & Gershuny, J. (2015) "A Mixed-Mode Approach to Measuring Young Peoples' Time Use in the UK Millennium Cohort Study" *Electronic International Journal of Time Use Research* 12(1): 174-180.

Floro, M. S., & Miles, M. (2001). *Time use and overlapping activities: evidence from Australia*. Social Policy Research Centre, University of New South Wales.

Foasberg, N. M. (2014). Student reading practices in print and electronic media. *College & Research Libraries*, *75*(5), 705-723.

Forsyth, B. H., & Lessler, J. T. (1991). Cognitive laboratory methods: A taxonomy. Measurement errors in surveys, 393-418.

Frazis, H., & Stewart, J. (2004). What can time-use data tell us about hours of work. *Monthly Lab. Rev.*, *127*, 3.

Frazis, H., & Stewart, J. (2007). Where does the time go? Concepts and measurement in the American Time Use Survey. In *Hard-to-measure goods and services: Essays in honor of Zvi Griliches* (pp. 73-97). University of Chicago Press.

Gershuny, J. (2000) Changing Times: Work and Leisure in Postindustrial Society, Oxford University Press, New York.

Gershuny, J. (2011). Time-use surveys and the measurement of national well-being. *Centre for Time-use Research, Department of Sociology, University of Oxford*.

Gimenez-Nadal, J. I., & Molina, J. A. (2015). Voluntary activities and daily happiness in the United States. *Economic Inquiry*, *53*(4), 1735-1750.

Glorieux, I., & Minnen, J. (2009). How many days? A comparison of the quality of time-use data from 2-day and 7-day diaries. *electronic International Journal of Time Use Research*, *6*(2), 314-327.

Grimm, P. (2010). Social desirability bias. *Wiley International Encyclopedia of Marketing*.

Gross, J., & Swirski, B. (2002). Time use surveys and gender equality. *Adva Centre, Tel Aviv.*

Harvey, A. S. (1993). Guidelines for time use data collection. *Social Indicators Research*, *30*(2-3), 197-228.

Harvey, A.S. (1997). From activities to activity settings: behavior in context. In D. Ettema and H. Timmermans (eds.), Activity-Based Approaches to Travel Analysis, 209-228, Elsevier Science Ltd., New York.

Harvey, A., & Royal, M. (2000). Use of context in time-use research. In *expert group meeting on methods for conducting time-use surveys* (pp. 23-27).

Harvey, A.S. and Singleton, J. (2009), Time-use and balance, in: Life balance multidisciplinary theories and research, Matuska, K. and C.H. Christiansen (eds.), Slack and AOTA, Bethesda, MD.

Healy, M. (2004). We're All Multitasking, but What's the Cost?. *Los Angeles Times*, *19*.

Hill, M. S. (1988). Marital stability and spouses' shared time: A multidisciplinary hypothesis. *Journal of Family Issues*, *9*(4), 427-451.

Hoffamn, E. and A. Mata (1999), Measuring working time – An alternative approach to classifying time use, in: Merz, J. and M. Ehling (eds.), Time use - research, data and policy – Contributions from the International Conference on Time Use (ICTU), FFB-Schriftenreihe Band 10, Nomos Verlag, Baden-Baden.

Hook, J. L., & Chalasani, S. (2008). Gendered Expectations? Reconsidering Single Fathers' Child-Care Time. *Journal of Marriage and Family*, *70*(4), 978-990.

Horrigan, M., & Herz, D. (2004). Planning, designing, and executing the BLS American time-use survey. *Monthly Lab. Rev.*, *127*, 3.

Hungerford, T. L. (2001) Just how productive are elderly adults? Evidence from Australia. Washington DC: Office of Policy, Social Security Administration.

International Labour Organization (ILO- 2013). Resolution concerning statistics of work, employment and labour underutilization. Geneva.

Ironmonger, D. (1996). "Counting Outputs, Capital Inputs and Caring Labor: Estimating Gross Household Product," Feminist Economics, Volume 2, 1996, pp. 37-63.

Ironmonger, D. (2003, September). There are only 24 hours in a day! Solving the problematic of simultaneous time. In *Proceedings of the 25 th IATUR Conference on Time Use Research*.

Ironmonger, D. (2004). Bringing up Bobby and Betty. *Family Time: The Social Organization of Care" London: Routledge*, 93-109.

Jahoda, M., Lazarsfeld, P. F. and Zeisel, H. (1933), Marienthal: The Sociography of an Unemployed Community – Translated by the authors with John Reginait and Thomas Elsaesser. Chicago 1971: ldine-Atherton.

Javid, F., & Varney, A. (2007). The great seduction of multitasking. *ABC News*.

Joyce, M., & Steward, J. (1999). What can we learn from time-use data. *Monthly Lab. Rev.*, *122*, 3.

Judge, T. A. (1993). Does affective disposition moderate the relationship between job satisfaction and voluntary turnover?. *Journal of Applied Psychology*, *78*(3), 395.

Juster, F. T., Ono, H., & Stafford, F. P. (2003). An assessment of alternative measures of time use. *Sociological Methodology*, *33*(1), 19-54.

Kahneman, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *The journal of economic perspectives*, *20*(1), 3-24.

Kan, M. Y. (2008). Measuring housework participation: the gap between "stylised" questionnaire estimates and diary-based estimates. *Social Indicators Research*, *86*(3), 381-400.

Kammann, R., & Flett, R. (1983). Affectometer 2: A scale to measure current levels of general happiness. Australian Journal of Psychology, 35, 259–265.

Kan, M. Y. (2008). Measuring housework participation: the gap between "stylised" questionnaire estimates and diary-based estimates. *Social Indicators Research*, *86*(3), 381-400.

Karabenick, S. A., Woolley, M. E., Friedel, J. M., Ammon, B. V., Blazevski, J., Bonney, C. R., & Kelly, K. L. (2007). Cognitive processing of self-report items in educational research: Do they think what we mean?. *Educational Psychologist*, *42*(3), 139-151.

Kelly, Paul, Emma Thomas, Aiden Doherty, Teresa Harms, Órlaith Burke, Jonathan Gershuny and Charlie Foster. (2015) "Developing a Method to Test the Validity of 24 Hour Time Use Diaries Using Wearable Cameras: A Feasibility Pilot" *PLoS ONE* 10(12): e0142198.

Kenyon, S. (2008). Internet use and time use: The importance of multitasking. *Time & Society*, *17*(2-3), 283-318.

Kenyon, S. (2010). What do we mean by multitasking?--Exploring the need for methodological clarification in time use research. *Electronic International Journal of time use research*, *7*(1).

King, L. A., & Napa, C. K. (1998). What makes a life good?. *Journal of personality and social psychology*, *75*(1), 156.

Kitamura, R. and S. Fujii (1996) Two computational process models of activity-travel behavior. Paper presented at the Theoretical Foundations of Travel Choice Modeling Conference held in Stockholm, August 7-11.

Kitterød, R. H. (2001). Does the recording of parallel activities in time use diaries affect the way people report their main activities?. *Social Indicators Research*, *56*(2), 145-178.

Kitterød, R. H., & Lyngstad, T. H. (2005). Diary versus questionnaire information on time spent on housework–The case of Norway. *Electronic International Journal of Time Use Research*, *2*(1), 13-32.

Kitzinger, J. (1995). Qualitative research. Introducing focus groups. *BMJ: British medical journal*, *311*(7000), 299.

Kraan, M. (1996) Time to travel?; a model for the allocation of time and money. Unpublished Ph.D. dissertation, Department of Civil Engineering, University of Twente, The Netherlands

Kracht, M. (2006) "Using Combined GPS and GSM Tracking Informationfor Interactive Electronic Questionnaires" In Peter Stopher and Cheryl C. Stecher (eds) *Travel Survey Methods: Quality and Future Directions* Elsevier.

Krenn, Patricia J., Sylvia Titze, Pekka Oja, Andrew Jones and David Ogilvie. (2011) "Use of Global Positioning Systems to Study Physical Activity and the Environment: A Systematic Review" *American Journal of Preventative Medicine* 41(5): 508-515.

Krueger, A. B., & Schkade, D. A. (2008). The reliability of subjective well-being measures. *Journal of public economics*, *92*(8), 1833-1845.

Krueger, A. B., Kahneman, D., Fischler, C., Schkade, D., Schwarz, N., & Stone, A. A. (2009). Time use and subjective well-being in France and the US. *Social Indicators Research*, *93*(1), 7-18.

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly*, *72*(5), 847-865.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity*, *47*(4), 2025-2047.

Larson, R., & Csikszentmihalyi, M. (1983). The experience sampling method. *New Directions for Methodology of Social & Behavioral Science*.

Lawson, C. (1996) Household activity-travel decisions. Draft dissertation proposal, Urban Studies and Regional Science Program, Portland State University.

Lee, Y. S., & Waite, L. J. (2005). Husbands' and wives' time spent on housework: A comparison of measures. *Journal of Marriage and Family*, *67*(2), 328-336.

Lee, Y., Hofferth, S. L., Flood, S. M., & Fisher, K. (2016). Reliability, validity, and variability of the subjective well-being questions in the 2010 American Time Use Survey. *Social indicators research*, *126*(3), 1355-1373.

Lindquist, J. D., & Kaufman-Scarborough, C. (2007). The polychronic—monochronic tendency model: PMTS scale development and validation. *Time & Society*, *16*(2-3), 253-285.

Marini, M. M., & Shelton, B. A. (1993). Measuring household work: Recent experience in the United States. *Social Science Research*, *22*(4), 361-382.

McGillivray, M., & Clarke, M. (2006). *Understanding human well-being*. United Nations University Press.

Meier, S., & Stutzer, A. (2008). Is volunteering rewarding in itself?. *Economica*, *75*(297), 39-59.

Merz, J. (2009). Time use and time budgets–Improvements, future challenges and recommendations.

Michelson, W. (2009). On adding affect to time-diary accounts. *Social indicators research*, *93*(1), 31-32.

Minnen, Joeri, Ignace Glorieux, Theun Pieter van Tienoven, Sarah Daniels, Djiwo Weenas, Jef Deyaert, Sarah Van den Bogaert and Sven Rymenants. (2014) "Modular Online Time Use Survey (MOTUS) – Translating an Existing Method in the 21st Century" Electronic International Journal of Time Use Research 11(1): 73-93.

Muis, K. R., Duffy, M. C., Trevors, G., Ranellucci, J., & Foy, M. (2014). What were they thinking? Using cognitive interviewing to examine the validity of self-reported epistemic beliefs. *International Education Research*, *2*(1), 17-32.

Niemi, I. (1993). Systematic error in behavioural measurement: Comparing results from interview and time budget studies. *Social Indicators Research*, *30*(2-3), 229-244.

OECD Better Life Initiative. (2013). OECD guidelines on measuring subjective well-being.

Offer, S., & Schneider, B. (2011). Revisiting the gender gap in time-use patterns: Multitasking and well-being among mothers and fathers in dual-earner families. *American Sociological Review*, *76*(6), 809-833.

O'Neill, G., & Dixon, J. (2005). Nonresponse bias in the American Time Use Survey. Proceedings of the American Statistical Association, Survey Research Methods Section. Alexandria, VA: American Statistical Association.

Otterbach, S. (2010). Mismatches between actual and preferred work time: Empirical evidence of hours constraints in 21 countries. *Journal of Consumer Policy*, *33*(2), 143-161.

Otterbach, S., & Sousa-Poza, A. (2010). How accurate are German work-time data? A comparison of time-diary reports and stylized estimates. *Social indicators research*, *97*(3), 325-339.

Park, H., Sha, M. M., & Pan, Y. (2014). Investigating validity and effectiveness of cognitive interviewing as a pretesting method for non-English questionnaires: Findings from Korean cognitive interviews. *International Journal of Social Research Methodology*, *17*(6), 643-658.

Pentland, W. E., Harvey, A. S., Lawton, M. P., & McColl, M. A. (Eds.). (1999). *Time use research in the social sciences* (pp. 259-268). New York: Kluwer Academic/Plenum Publishers.

Pember Reeves, M. (1913), Round About a Pound a Week, Bell, London; Neue Auflage: Verlag Virago 1990.

Phipps, P. A., & Vernon, M. K. (2009). Twenty-four hours. *Calendar and time diary: Methods in life course research*, 109-128.

Ravanera, Z., & Raiulton, F. P. Turcotte (2003), Youth integration and social capital–An analysis of the canadian general social surveys on time use. *Youth and Society*, *35*(2), 158-82.

Ricroch L.(2011) "The enjoyable moments of day to day life", Insee Premiere n°1378.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of experimental psychology: General*, *124*(2), 207.

Robinson, J. P. (1999). The time-diary method. Structure and uses. In Pentland, W. E. et al. (Eds.) Time Use in the Social Sciences. New York: Kluwer Academic Publishers, pp. 47–89.

Robinson, J. P., & Godbey, G. (1999). Time for Life. State College.

Robinson, J. P. (2013). Americans less rushed but no happier: 1965–2010 trends in subjective time and happiness. *Social indicators research*, *113*(3), 1091-1104.

Rydenstam, K. (1999). The EUROSTAT Project On Harmonising Time Use Statistics, Merz, J. and M. Ehling (eds.), Time Use–Research, Data and Policy, FFB-Schriftenreihe.

Rydenstam, K. (2001, October). Paid and unpaid work, the apples and pears of time use statistics. In *IATUR Conference, Oslo.*

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. Journal of Experimental Psychology: Human Perception and Performance, 27, 763–797

Ruuskanene, O. P. (2004) Essay 4. More than two hands: is multitasking an answer to stress? In Ruuskanene, O. P. An econometric analysis of time use in Finnish households. Unpublished PhD thesis, Helsinki School of Economics, Acta Universitatis Oeconomicae Helsingiensis. 189-229.

Sauvain-Dugerdil, C., Ritschard, G., & Dieng, A. W. (2006). Socio-démographie des phases tardives de la vie familiale.

Sayer, L. C. (2010). Trends in housework. *Dividing the domestic: Men, women, and household work in cross-national perspective*, 19-38.

Schulz, F., & Grunow, D. (2007). *Comparing Time Diary Data and Stylized Time Use Estimates*. Working Paper 2007-05, The Center for Research on Inequalities and the Life Course, Yale University.

Schulz, F., & Grunow, D. (2011). Comparing diary and survey estimates on time use. *European Sociological Review*, *28*(5), 622-632.

Schwartz, L. K. (2002). The American time use survey: Cognitive pretesting. *Monthly Lab. Rev.*, *125*, 34.

Schwarz, N. (1987). Stimmung als Information: Untersuchungen zum Einfluβ von Stimmungen auf die Bewertung des eigenen Lebens. Heidelberg: Springer Verlag; 1987.

Sonnenberg, B., Riediger, M., Wrzus, C., & Wagner, G. G. (2012). Measuring time use in surveys–Concordance of survey and experience sampling measures. *Social science research*, *41*(5), 1037-1052.

Stiglitz, J., Sen, A., & Fitoussi, J. P. (2009). The measurement of economic performance and social progress revisited. *Reflections and overview. Commission on the Measurement of Economic Performance and Social Progress, Paris.*

Stopher, P. R., Wilmot, C. G., Stecher, C., & Alsnih, R. (2006). Household travel surveys: proposed standards and guidelines. In *Travel survey methods: Quality and future directions* (pp. 19-74). Emerald Group Publishing Limited.

Sullivan, O. (1997). Time waits for no (wo) man: An investigation of the gendered experience of domestic time. *Sociology*, *31*(2), 221-239.

Sullivan, O., & Gershuny, J. (2017). Speed-Up Society? Evidence from the UK 2000 and 2015 Time Use Diary Surveys. Sociology, 0038038517712914. Chicago

Szalai, A. (1972). The use of time: Daily activities of urban and suburban populations in twelve countries.

The United Kingdom 2000 Time Use Survey Technical Report. Available from http://www.statistics.gov.uk/downloads/theme_social/UKTUS_TechReport.pdf, accessed at 26 October 2007.

Tourangeau, R. (1984). Cognitive sciences and survey methods. *Cognitive aspects of survey methodology: Building a bridge between disciplines*, 73-100.

UNIFEM (2009) United Nations Development Funds for Women 2008, Who Answers to Women? Progress of the World's Women 2008/2009 –Gender & Accountability (http://www.unifem.org/progress/2008/)

United Nations Economic and Social Council (2013). Time use surveys, policy relevance, international harmonization initiatives and new guidelines. Economic commission for Europe, Conference of European Statisticians. ECE/CES/2013/6.

United Nations Statistics Division (2016). International Classification of Activities for Time Use Statistics. UNSD – 13 February 2017.

Wildy, H., & Clarke, S. (2009). Using cognitive interviews to pilot an international survey of principal preparation: A Western Australian perspective. *Educational Assessment, Evaluation and Accountability*, *21*(2), 105-117.

Wallis, C. (2006). The multitasking generation. *Time Magazine*, *167*(13), 48-55.

Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design*. Sage Publications.

Wilson, W. R. (1967). Correlates of avowed happiness. *Psychological bulletin*, *67*(4), 294.

Zick, C. D., & Bryant, W. K. (1996). A new look at parents' time spent in child care: Primary and secondary time use. *Social Science Research*, *25*(3), 260-280.

Zuzanek, J. (2009). Students' study time and their "homework problem". *Social indicators research*, *93*(1), 111-115.

Zuzanek, J. (2009). Time use research in Canada–History, critique, perspectives. *electronic International Journal of Time Use Research*, *6*(2), 178-192.

Appendices

Appendix A: 5 and 7-point scales used in the focus group with mid-points 3 and 4 respectively

1	2	3	4	5
Not at all				Very much

1	2	3	4	5	6	7
Not at all						Very much

Appendix B: 5 and 7-points scales used in the focus group with 0 as the mid-point

-2	-1	0	+1	+2
Not at all				Very much

-3	-2	-1	0	+1	+2	+3
Not at all						Very much

Appendix C: Potential words to measure well-being presented to the focus group

How *pleasant* was this moment?
1   2   3   4   5
Not at all   Very much

How much did you *enjoy* your time?
1   2   3   4   5
Not at all   Very much

How much did you *value* your time?
1   2   3   4   5
Not at all   Very much

Appendix D: Types of activity and specific examples presented to the focus group

Group	Examples
Personal care	Sleeping, eating
Employment	Main job
Study	Classes, lectures
Household and family care	Gardening, pet care, shopping
Volunteer work and meeting	Informal help to other households
Social life and entertainment	Cinema, telephone conversation
Sport and outdoor activities	Walking, hiking, biking
Hobbies, games and computing	Chess, Bridge
Mass media	Reading, watching TV, listening to radio
Travel and unspecified Time Use	Travelling to work

Appendix E: Pre-completed example diary page and completion instructions

Example

- Record your main activity for each 10-minute period
- Only record one main activity on each line! Distinguish between travel and the activity that is the reason for travelling.
- Do not forget the mode of transportation.
- Distinguish between first and second activity, if any.

- For each 10-minute period, please indicate how much you enjoyed this time by circling a number on a scale of 1 to 7, with 1 meaning you did not enjoy it at all and 7 being that you enjoyed it very much.
- For example, if you didn't enjoy an activity at all then you would circle 1 in the box.

Afternoon & Evening   Please fill in all three of the separate hour periods
Times: 15:00 - 16:00, 19:00 - 20:00 and 23:00 - 00:00

Time	What were you doing? Please write down one main activity.	If you did something else at the same time, what else did you do?	Did you use a smartphone tablet, or computer?	Where were you? Location, or mode of transport	Alone	If you were with your own child, note whether or not they live with you.	How much did you enjoy this time? On the scale below from 1 to 7 Circle the appropriate rating
15:00-15:10	Walked to children's school	Listened to radio	■	on foot	■		1 2 3 (4) 5 6 7
15:10-15:20	Picked children up from school	Talked to friend	☐		☐	Friend	1 2 3 4 (5) 6 7
15:20-15:30	Walked home	Talked with my children	☐		☐	My youngest child, live with	1 2 3 4 5 (6) 7
15:30-15:40	↓	--"--	☐		☐	--"--	1 2 3 4 5 (6) 7
15:40-15:50	Helped the children change	--"--	☐	At home	☐	--"--	1 2 3 (4) 5 6 7
15:50-16:00	Prepared the dinner	Set the table	☐	--"--	■		1 (2) 3 4 5 6 7

Were you alone or with somebody you know? If you were alone tick the box or if you were with someone else, write in who this was. Note their relationship e.g. your child, friend, partner, husband etc

1 = not at all   7 = very much

Use an arrow, citation marks or the like to mark that an activity lasts longer than 10 minutes.

Appendix F: Written instructions covering diary procedures as provided to interviewers

Time Use Survey	

Client: ONS
Target Group: All respondents who agree to participate in future research (England and Scotland only).

For interviewers in England and Scotland– You will have been provided with up to 6 'Time Use' packs containing a diary and a pre-paid envelope. Each pack also contains an information sheet for you to give to respondents, both explaining the survey and how to fill out the diaries.

If, at the end of the questionnaire, the respondent consents to participate in future research, please ask if they would be willing to take part in the 'Time Use Survey. If so, please leave the pack and ask them to complete the diary and post it back within a week. A prompt response will ensure that the research team meet their deadline.

The data provided will be used to develop and monitor policy across Europe. For example, understanding the amount of time that parents spend working and with their children may help to develop child benefit policies.

The front of the diary asks for the OPN case number. This is the quota/address number for the interview. Please ensure you write this in **before** handing it over to the respondent.

You have been provided with two versions of the diary – A and B. Please alternate which one you give to respondents. This will ensure a roughly 50/50 split.

The research team will be conducting follow-up, telephone interviews with some of the respondents who complete the Time Use Survey so it's really important that you take a contact number for each respondent who agrees to complete a Time Use diary.

Appendix G: Respondent letter

Office for National Statistics

www.ons.gov.uk/surveys

Segensworth
Road Titchfield,
Fareham
PO15 5RR

Dear Sir or Madam,

August 2017

Thank you for agreeing to take part in future research for the Office for National Statistics (ONS). We would like you to help us develop a future study about how people choose to spend and enjoy their time.

Why should I take part?

We are running a small survey that will take about 20 minutes to complete. The study will capture how you spend your time completing day to day activities. The purpose of this study is to gain information that will help design and improve future research. This is important as time use studies measure national well-being and economic growth which can influence policy making.

What happens next?

We ask that you complete the diary within 7 days of receiving this letter. Once you have completed the diary, please post it back to us in the pre-paid envelope provided by our interviewer. If you have any queries, you can phone us free on 0800 298 5313. More information about this Time Use Study is also available on our website www.ons.gov.uk/surveys.

Who can take part?

We ask that the adult who took part in your Opinions and Lifestyle Survey interview should complete the diary. The information you give us will be treated as confidential as directed by the code of practice for Official Statistics.

Thank you for your time.
Yours faithfully

Tim Vizard, Research Lead

Appendix H: Example diary page

Morning   Please fill in all three of the separate hour periods
Times: 04:00 - 05:00, 07:00 - 08:00 and 10:00 - 11:00

Time	What were you doing? Please write down one main activity.	If you did something else at the same time, what else did you do?	Did you use a smartphone tablet, or computer?	Where were you? Location, or mode of transport	Were you alone or with somebody you know? If you were alone tick the box or if you were with someone else, write in who this was. Note their relationship e.g. your child, friend, partner, husband etc		How much did you enjoy this time? On the scale below from 1 to 7 *Circle the appropriate rating* 1 = not at all 7 = very much
					Alone	If you were with your own child, note whether or not they live with you.	
04:00-04:10			☐		☐		1 2 3 4 5 6 7
04:10-04:20			☐		☐		1 2 3 4 5 6 7
04:20-04:30			☐		☐		1 2 3 4 5 6 7
04:30-04:40			☐		☐		1 2 3 4 5 6 7
04:40-04:50			☐		☐		1 2 3 4 5 6 7
04:50-05:00			☐		☐		1 2 3 4 5 6 7
07:00-07:10			☐		☐		1 2 3 4 5 6 7
07:10-07:20			☐		☐		1 2 3 4 5 6 7
07:20-07:30			☐		☐		1 2 3 4 5 6 7
07:30-07:40			☐		☐		1 2 3 4 5 6 7
07:40-07:50			☐		☐		1 2 3 4 5 6 7
07:50-08:00			☐		☐		1 2 3 4 5 6 7
10:00-10:10			☐		☐		1 2 3 4 5 6 7
10:10-10:20			☐		☐		1 2 3 4 5 6 7
10:20-10:30			☐		☐		1 2 3 4 5 6 7
10:30-10:40			☐		☐		1 2 3 4 5 6 7
10:40-10:50			☐		☐		1 2 3 4 5 6 7
10:50-11:00			☐		☐		1 2 3 4 5 6 7

Row group labels (vertical): 04:00 - 05:00 (AM), 07:00 - 08:00 (AM), 10:00 - 11:00 (AM)

Appendix I: Diary checklist question

Checklist

1 When did you fill in the diary?
- [ ] Now and then during the day
- [ ] At the end of the diary day
- [ ] The day after the diary day
- [ ] Later, __ days after the diary day

2 Were you at home or somewhere else at the start of the diary?
- [ ] At home
- [ ] Somewhere else

3 Were you at home or somewhere else at the end of the diary?
- [ ] At home
- [ ] Somewhere else

4 Did you feel rushed this day?
- [ ] Yes
- [ ] No

5 Was this an ordinary or unusual day?
- [ ] An ordinary day
- [ ] An unusual day

6 If you are either employed or a student: What kind of day was this for you?
- [ ] An ordinary workday
- [ ] An ordinary school day
- [ ] A day off due to weekend/ holiday or work schedule
- [ ] A sick leave day
- [ ] A vacation day
- [ ] On leave for another reason

Before you post me back...

Time Use research is important for shaping policies that relate to how people divide their time between work and home life.

We may be interested in talking to a few people to ask how you found completing the diary.

Please tick here if you are willing to be contacted: [ ]

Please note down a number we can contact you on:

Please go through the diary and check the following:

- Have you noted one main activity in each line and checked that there are no empty time periods?
- Have you marked clearly when you were working, even if that was at home?
- Have you recorded all travel and modes of transport?
- Have you marked the duration of parallel activities, if any?
- Have you checked that every row has either "alone" marked or the relationship to the person you were with noted?

Many thanks for filling in the diary

If you have any questions, please contact us on our survey enquiry line: 0800 298 5313

Or visit our website: www.ons.gov.uk/surveys

Please return completed diaries to:
Segensworth Road, Titchfield, Fareham, PO15 5RR

With thanks to NatCen and the Centre for Time Use Research for providing the UK Time Use diary.

Appendix J: Mean enjoyed and satisfied ratings for context variables, differences between them and statistical information for those cases where significant differences were identified

		Mean enjoyed rating	Mean satisfied rating	Mean difference
Alone or with strangers?	No, not selected	5.9	5.6	0.3
	Yes, alone	5.4	5.6	0.2
Smart device used	No, not selected	5.6	5.6	0.0
	Yes - selected used smart device	5.7	5.4	0.3
Location	Own home	5.6	5.7	0.0
	Workplace or school	5.5	4.7	0.8
	Home of others	6.0	6.1	0.1
	Restaurant, bar, cafe	6.6	6.3	0.3
	Other locations	5.7	5.6	0.0
Mode of transport	Not travelling	5.6	5.6	0.1
	Travel by car, van	5.6	5.2	0.3
	Public transport or taxi	5.3	4.8	0.6
	Walk, other physical transport	5.7	6.1	0.4
	Other/unspecified transport	6.1	5.5	0.6
Felt rushed completing diary	Yes	4.8	5.1	0.4
	No	5.7	5.7	0.1

Appendix K: Showing the main activities within which, smart devices were used

Main activity	Smart device status	Percentage of respondents (%)
Sleep, naps, in bed sleepless or ill	No - not selected	96
	Yes - selected used smart device	4
Eating or drinking	No - not selected	75.6
	Yes - selected used smart device	24.4
Wash, dress, other care for self	No - not selected	87.8
	Yes - selected used smart device	12.2
Employment in main or second job	No - not selected	49.7
	Yes - selected used smart device	50.3
Breaks, job search, other employment related, education, training	No - not selected	40.7
	Yes - selected used smart device	59.3
Food preparation, cooking	No - not selected	85.8
	Yes - selected used smart device	14.2
Gardening	No - not selected	95.3
	Yes - selected used smart device	n/a
Pet and animal Care	No - not selected	90.5
	Yes - selected used smart device	9.5
Shopping and services	No - not selected	90.6
	Yes - selected used smart device	9.4
Other domestic work and household management	No - not selected	79.8
	Yes - selected used smart device	20.2
Read to, teach, play with child, help with homework	No - not selected	78.9
	Yes - selected used smart device	21.1
Physical, medical other child, adult care	No - not selected	82.5
	Yes - selected used smart device	17.5
Voluntary, civic, organisational act,	No - not selected	87.5

Main activity	Smart device status	Percentage of respondents (%)
worship and religion	Yes - selected used smart device	n/a
Social leisure	No - not selected	83
	Yes - selected used smart device	17
Rest, relax, think, do nothing	No - not selected	76.9
	Yes - selected used smart device	23.1
Sport or exercise, outdoor activities	No - not selected	77.1
	Yes - selected used smart device	22.9
Arts, hobbies, culture and events	No - not selected	92.3
	Yes - selected used smart device	n/a
Computing and social media	No - not selected	7.4
	Yes - selected used smart device	92.6
Games	No - not selected	53.3
	Yes - selected used smart device	46.7
Read	No - not selected	78.7
	Yes - selected used smart device	21.3
Watch TV, video, DVD, visual content, listen to radio, music, podcast, audio	No - not selected	65.3
	Yes - selected used smart device	34.7
Travel to/from work	No - not selected	86.2
	Yes - selected used smart device	13.8
All other travel	No - not selected	92.6
	Yes - selected used smart device	7.4

1. Where activities have contained fewer than 3 cases, the categories have been combined or suppressed to avoid disclosure. Suppression is indicated by "n/a" in the table.

107

Appendix L: Table showing correlations between standardized versions of the OPN well-being scales and diary well-being scales

England and Scotland, 2017

		Overall, how satisfied are you with your life nowadays?	Overall, to what extent do you feel that the things you do in your life are worthwhile?	Overall, how happy did you feel yesterday?	Overall, how anxious did you feel yesterday?
Diary enjoyed scale	Spearman's rho	0.198	0.212	.280*	-0.197
	Sig (2-tailed)	0.082	0.064	0.013	0.084
	N	78	77	78	78
Diary satisfied scale	Spearman's rho	0.226	.278*	.364**	-.478**
	Sig (2-tailed)	0.051	0.016	0.001	0.000
	N	75	75	75	75

** significant at p<0.01

**significant at p<0.05

Appendix M: Glossary of shortened activity descriptions

Activity description as originally coded in the diary study	Shortened activity description for main body of report
Sleep, naps, in bed sleepless or ill	Sleeps/naps
Eating or drinking	
Wash, dress, other care for self	Personal hygiene
Employment in main or second job	
Breaks, job search, education, training, other employment related	Employment related activities
Food preparation, cooking	
Gardening	
Pet and animal care	Animal care

Activity description as originally coded in the diary study	Shortened activity description for main body of report
Shopping and services	
Other domestic work and household management	
Read to, teach, play with child, help with homework	Childcare (read/teach/play)
Physical, medical other adult/child care	
Voluntary, civic, organisational act, worship, religion	Voluntary/religion
Social leisure	
Rest, relax, think, do nothing	Rest and relaxation
Sport or exercise, outdoor activities	Sport/exercise
Art and hobbies, other indoor leisure, events, culture	Arts/hobbies
Computing and social media	
Games	
Read	
Watch TV, video, DVD, visual content or listen to radio, music, podcast or radio	Audio/visual media use
Travel to/from work	Commuting
All other travel	

1. In figure 4.7 Breaks, job search, education, training, other employment related is split into two categories: Breaks, job search, other employment related and education and training. The former is referred to in figure 7 as employment related activities.

Appendix N: <u>Cognitive Interview Topic Guide – Time Use Diary Study</u>

Background for interviewers

Introduction

Respondents completed time-use diaries at specified times on a particular day, in which they recorded:

- Their main activity
- Anything else they were doing at the time
- Any technology used
- Their location
- People they were with
- The extent to which they enjoyed this activity or were satisfied with it (as an activity-level measure of well-being)

Diary study design
A sample of 500 OPN respondents were invited to take part in the study and 135 completed diaries were returned. A between groups design was employed in which respondents were allocated to different groups according to the day of the week that they completed the diaries (i.e weekday/weekend) and the well-being measure used (i.e. they were either asked about how much they enjoyed or were satisfied with an activity as a measure of well-being). Some of the questions they will be asked in the Cognitive Interview will differ in wording slightly depending on whether respondents were asked about the extent to which they enjoyed or satisfied with the activities reported in the diary.

Previous ONS research

We first pre-tested the diaries using focus groups. The focus groups identified several issues that might have caused difficulty for respondents. These included:

1. Ratings of enjoyment/satisfaction – are they based on current mood vs mood recall?, are they based on events during the day?, are they influenced by how enjoyable their lives are in general or are satisfied with them?, what do their responses mean?
2. Anchoring – extent to which ratings might change based on the ratings of other activities
3. Temporal effects - events happening earlier in the day likely to affect ratings of events later in the day
4. Meaning of scales – what would be the mid-value on a scale?, what do numbers mean in terms of well-being terms? (i.e. satisfaction, enjoyment).

We would like to probe these issues in the cognitive interviews and we have devised a range of potential questions designed to elicit responses about these issues.

The Time Use Diaries

The diaries covered six hour-long blocks in which activities and accompanying contextual information was recorded by respondents. Keep in mind that this diary covers a lived experience over two months ago, and the respondent may have little precise recall of that exact day now, but

they will remember the exercise. What we are seeking to do now is to elicit more information about how relevant and meaningful the process was for them and to gain more insight into the drivers of different levels of well-being associated with activities. You can use the diary instrument as a guide to talk about their experiences yesterday, what kinds of activities or combinations of circumstances increase or decrease their daily well-being.

Methodology

Design

20 participants will be selected from a convenience sample of 135 diary respondents who signaled their intention to participate in continued research on the time use survey. They will be cognitively interviewed over the telephone.

Procedure

Setting up the telephone interview

1. *Find a quiet room/desk for the telephone interview*
2. *Set up telephone recording equipment and test it*
3. *Set out scripts and instructions.*

Note to interviewers:
- Have a blank copy of the diary available to write down what the respondent tells you during probing.
- Try to motivate the respondent whilst minimizing feedback that could be construed as positive or negative. Avoid use of 'yes', 'no' and any other term of agreement or disagreement. Instead use 'okay' and 'that's really interesting' for example.

TOPIC GUIDE

Introduction

Hello [R name]/ My name is [name] and I work at the Office for National Statistics. You completed a diary for us a few weeks ago in which you recorded your activities over the course of a day. On that diary you indicated that you were willing for us to contact you to ask you a few questions about your experiences in completing it. This will help to make sure the diary is the best it can be for people to fill in.

Would it be convenient for you to spend 15-20 minutes answering some questions right now?

I'm just going to ask you a few questions about what you wrote, what you thought and how you understood some of the wording. I didn't design the diary so feel free to share your honest thoughts; you won't offend me. There are no right or wrong answers. Some of the questions I will ask might seem really obvious. I'm not trying to catch you out, I just don't want to presume I know what you're thinking. I'm here to learn from you.

If it's ok I'm going to record this conversation so I don't have to make any long notes and I can properly listen to you. Everything you say will be kept confidential and you won't be identifiable in any way.

Do you have any questions before we start?

Section 1: Diary completion

Aims to encourage spontaneous answers and to ease the respondent in

"Firstly I just wanted to ask you a couple of questions about how you found the diary task" (interviewer may have to remind R – the diary asked you to fill in your main activity within certain time periods….)

Condition	Probe	Prompt
All	How did you find completing the diary?	Did you find it easy or difficult? If difficult: Were there any particular aspects you struggled with?
All	How did you decide what to write as the main activity?	
All	How easy or difficult did you find it to decide what the main activity was and which were the secondary activities?	
All	How did you decide what to write for the location?	

Section 2: General feeling the day before the interview

Aims to understand their general levels of satisfaction/ enjoyment on that day

I know that it might be difficult to think back to the detail you wrote in the diary a couple of months ago, so I wanted to do a similar exercise looking at your day yesterday.

Condition	Probe	Prompt
All	If you were to write a diary for yesterday, what would you include? (Interviewer to write down a few activities to probe on later)	What did you do? Was it a normal day or did something unusual happen? How did this make you feel?
All	Can you tell me a little about how you felt yesterday?	Were you feeling particularly satisfied, particularly unsatisfied or just normal?

Condition	Probe	Prompt
All	How would you rate your day in general, using the diary scale ranging from 1 (not at all satisfied) to 7 (very satisfied)?	Why was this? How did you come up with that answer?
Satisfied	What was the activity you felt most satisfied with yesterday?	Why was this? Which aspects made it particularly satisfying? Is this a normal activity for you or something unusual? Who were you with, if anyone, when doing this activity? What rating would you give that?
Satisfied	What was the activity you felt least satisfied with yesterday?	Why was this? Which aspects made it particularly unsatisfying? Is this a normal activity for you or something unusual? Who were you with, if anyone, when doing this activity? What rating would you give that?

Section 3: Providing a rating

I'd like to ask you a couple of questions about how you went about rating the activities.

Condition	Probe	Prompt
All	Which, if any, of the activities from yesterday would you find easiest to rate?	Why?
All	Which, if any, of the activities from yesterday would you find most difficult to rate?	Why?
Satisfied	Can you tell me about another activity from yesterday (which was not your most satisfying or unsatisfying)?	
All	What rating would you give that moment? Why	What were you thinking about when you came up with that rating?

Condition	Probe	Prompt
All	How easy or difficult did you find it to come up with an answer?	Why was this?
All	What does a 3 on that scale mean to you?	The scale ranges from 1 (not at all satisfied) to 7 (very satisfied). Can you give it a name or description? (i.e. if you were to give it a label what would it be?).
All	What do you see as the difference between number 5 and 6 on this scale?	The scale ranges from 1 (not at all satisfied) to 7 (very satisfied).
All	What does the midpoint (number 4) on this scale mean to you?	The scale ranges from 1 (not at all satisfied) to 7 (very satisfied).
Satisfaction	Does your satisfaction with this activity stay the same or does it change on different types of days?	If you'd had a really good day would you answer have changed or been the same? If you'd had a really bad day would your answer have changed or stayed the same?

Section 4: Understanding of response scale wording

Aims to look at how respondents understood the rating aspect of the activity

And lastly, I'd like to ask you about your understanding of some of the words we have used

Condition	Probe	Prompt
Satisfied	Using your own words, what does satisfied mean to you in this context?	How would you define? Can you think of a word that means the same? Can you think of a word that means the opposite (probe further if they

Condition	Probe	Prompt
		say unsatisfied)?
Satisfied	How would you rate your day in general yesterday, using a similar diary scale ranging from 1 (not at all satisfied) to 7 (very satisfied)?	Why is that?
Satisfied	Do the words enjoyed and satisfied mean the same thing to you or different?	
Satisfied	Would rating your enjoyment be easier or harder than rating your satisfaction with your activities?	Ask with regard to the activities already mentioned from yesterday: Would it be easier or more difficult to rate [pick activity] in terms of how much you enjoyed or did not enjoy this activity? Would your rating be different?

Close by thanking the respondent for their time and asking if they have any questions.

Appendix O: Factors and specific example descriptor terms suggested by participants as representing the points on the satisfied diary scales

General factor	Specific factor	Examples	Represented scale points
Task	Task outcome	Task non-completion	4
		Basic task achievement	5
		Semi-task completion	6
		Task completion	7
	Task process	Not particularly difficult task	4
		Basic task process	5
		Task process	6
		Interesting task(s)/work/unsual work	7
	Task timing & novelty	*"Something has to be done"*	3
		"Sorting out plans/something has to be done"	4
		"Fairly routine"	5
		"Delayed/put-off task"	7
Individual	Emotion only	*"Boring"*	4
		"Good, with frustrating moments"	5
		"A bit frustrating"	6
	Physical & emotional	*"able to relax"*	6
Situational	Unusualness	*"Nobody died"*	2
		"Nothing exciting happened"	3
		"Normal day"	5
		Memorable event	6
		Win money	7
	Environmental	*Poor traffic conditions*	3
Satisfaction	Satisfaction	*"Not at all satisfied"*	1
		"Dissatisfying"	3
		"Unsatisfactory"	3
		"Very satisfactory"	4
		"Slightly satisfied"	6
		"Satisfied"	6
		"Very satisfied"	6
		"Always satsified"	7
		"Continually satisfied"	7
Enjoyment	Enjoyment	*"Unpleasant but not serious experience"*	2
		"Unhappy"	3
		"Not enjoyed"	3
		"Not particularly happy activity"	3
		"Quite happy"	5
		"Chuffed"	5
		"Fairly happy"	6
		"Enjoy task"	7
		"Very happy"	7

General factor	Specific factor	Examples	Represented scale points
Ordinal	Ordinal	*"OK"*	3
		"Not particularly high or low"	3
		"Not particularly high or low"	4
		"Less than average"	4
		"Average"	4
		"Middle"	4
		"Tolerable"	4
		"Ordinary"	4
		"Usual"	4
		"Normal"	4
		"Neutral"	4
		"Either"	4
		"OK"	5
		"Not particularly high or low"	5
		"Medium"	5

1. Quotes in italics in the examples column denote direct quotes from participants on the meaning of scale points. Those examples not in quotes were derived by the research team from participants' responses.

Appendix P: Factors and specific example descriptor terms suggested by participants as representing the points on the enjoyed diary scales

General factor	Specific factor	Examples	Represented scale points
Task factors	Task outcome	Task non-completion/achievement	3
		Task non-completion/achievement	4
		Task completion/achievement	6
		Task completion/achievement	7
	Task timing & novelty factors	*"Not as planned"*	3
		Necessary task	3
		Routine food preparation	5
		"Break from routine"	6
		"Break from routine"	7
		Novel activity	7
Individual	Emotion only	*"Pretty hacked off"*	2
		"Tedious"	3
		"Emotionless"	3
	Physical & emotional	*"Relax"*	5
Situational	Unusualness	*"In a hurry"*	2
		"In a hurry"	3
		"Unusual day"	3
		"Different work colleagues"	5
		"Not doing anything else"	6
		"Nice treat"	7
	Environmental	*"Travelling"*	4
		"Pleaseant day"	6
		"Good day"	6
Satisfaction	Satisfaction	*"Dissatisfied"*	3
		"Satisfactory"	4
		"Satisfying"	5
		"Satisfying"	7
Enjoyment	Enjoyment	*"Stressful"*	2
		"Not enjoyed"	3
		"Upsetting"	3
		"Not happy"	3
		"Stressful"	3
		"Mostly enjoyable"	4
		"Partially enjoyable"	4
		"Quite enjoyed"	5
		"Happy but less than quite happy"	5
		"Less pleasurable"	5
		"Very enjoyable"	6
		"Really enjoyed"	7
		"Enjoyed that"	7
Ordinal	Ordinal	*"Just under halfway"*	3
		"Average"	4
		"Neither"	4
		"Just over halfway"	5
		"Good but not great"	5
		"Meidiocre"	5
		"OK"	5
		"Slightly improved from mediocre"	6
		"Nice"	6

1. Quotes in italics in the examples column denote direct quotes from participants on the meaning of scale points. Those examples not in quotes were derived by the research team from participants' responses.