KM 37(1991)
pag 65 - 86

Mode effects in time diary research

Nelly Kalfs and Willem E. Saris *

Abstract

Traditionally, a paper-and-pencil diary is used to examine time use. An alternative to this procedure is an electronic diary in which the coding of activities is done by a computer-assisted tree-structured questionnaire.
This paper reports on a comparison between the use of this diary in the framework of a self-registration panel (Telepanel) and an interviewer-administered panel (CATI).
The results show that the procedures differ with respect to the response rate, validity and costs. In general, our hypotheses concerning the procedures used and their consequences are confirmed.

This paper was presented at the XII World Congress of Sociology.
Madrid, July 9 - 13, 1990

* Department of Statistics and Methodology PSCW, University of Amsterdam, O.Z. Achterburgwal 237, 1012 DL Amsterdam, The Netherlands. EARN: A715IOPS@HASARA01.BITNET.

This study was made possible by the SWOKA, the Institute for Consumer Research and the IOPS, the Interuniversity Institute for Psychometrics and Sociometrics.

Special thanks are due to Ko Oudhof, Ineke Stoop, Frank van der Pol and Harm Hartman.

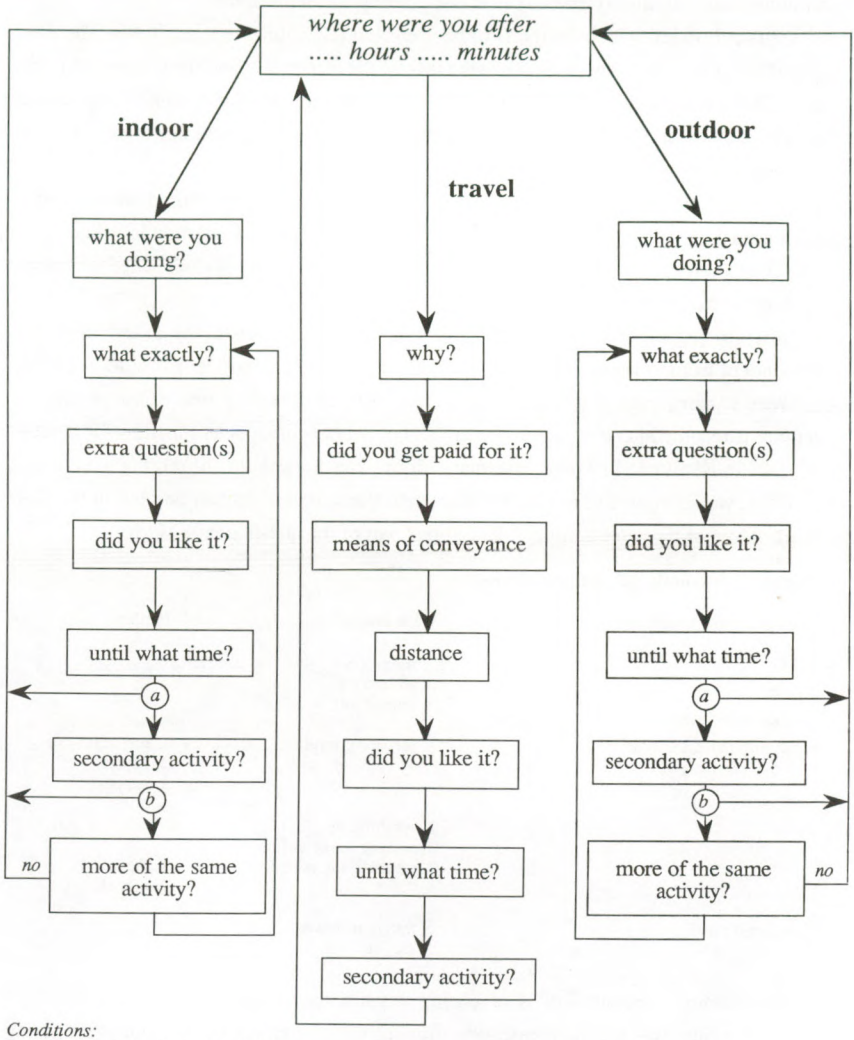
Introduction

In time use studies many researchers make use of a diary to register the activities of respondents (Juster 1985, Robinson 1977). Traditionally, this is done by a paper-and-pencil method. In this paper we report on an alternative procedure: an electronic diary. This diary is used in two different environments. The first data-collection procedure is based on a self-registration method (Telepanel), the second on an interviewer-administered procedure (CATI). A comparison between these two methods is made in terms of response rate, validity and costs. At a later stage its reliability will be examined.
The most obvious measure is the response rate since we know or are able to estimate the maximum number of possible participants. Obviously the preferred procedure is the one with the highest response rate.
Validity is measured indirectly because we normally do not know what real time use looks like (Lyberg, 1990). Often data obtained with a comparable method are used to see whether systematic differences exist. If differences are found, researchers try to explain them as being due to possible dissimilarities in the design and will evaluate their quality by certain accepted aspects of validity: '...other things equal, it seems reasonable to assume that a less valid diary tends to report fewer activities, a smaller variety of activities, fewer secondary activities, more diary time 'not ascertained' and more activities starting on the hour or half hour' (Juster, 1986). These criteria implicitly assume that more information is better: greater detail in reporting and the ability to account for time are associated with more valid reporting.
The most difficult criterion, costs, is estimated in terms of time. Because the environment in which the data are gathered varies considerably, it is impossible to make a comparison in terms of money. The procedure which takes less time is considered preferable.
Before we can evaluate the quality of the procedures, two topics need to be discussed. Since design characteristics influence the precision of the data, something has to be said about the 1) methodology and 2) implementation of the electronic diary. This will clarify the design of our diary and explain the differences between the Telepanel and CATI procedure.

Methodology of the electronic diary

The information gathered with a diary '...can show for an individual what activities were done during the defined period, how many times, in what order, at what time, for how long, where and other objective and subjective information connected with the activities' (Harvey, 1984).

Figure 1. Flow chart of the questionnaire



Conditions:
a: if the reported activity is not sleeping
b: if the reported activity is doing the household, obtaining goods and services, work or job, media

The diary itself can be designed in many different ways. The activity categories may be precoded or open, the time interval may be fixed (periods of 10, 15 or 30 minutes are the most common) or open (asking until what time an activity lasted), the activity code itself

can be varied and the diaries may provide space for recording only one or multiple simultaneous activities (Gershuny et al. 1986, Stoop & Oudhof 1989).

It is difficult to tell which design is to be preferred (Lingsom, 1979 and 1980). The most important choice seems to lie between the task of the respondent and the processing of the data. This problem has been the major reason for designing an electronic diary. In our opinion this form of diary is less demanding on the respondent and deals with the data processing very well.

In our electronic diary, activities are coded by answering a tree-structured questionnaire. In figure 1 a flow chart of the questionnaire is given. The central question is: where were you after .....hours .....minutes? The answer to this question determines whether indoor, outdoor or travel activities are shown.

In order to register the activity, at least two questions need to be asked. First the respondent has to choose between a number of primary categories: for indoor activity between 17 categories, and for outdoor activity between 15 categories. Subsequently the activity is recorded in more detail. The amount of detail is often determined by the number of questions asked. In some instances: doing the household, obtaining goods and services, work or job and media, we shortened the structure, as can be seen in the flow chart: 'More of the same activity?'. In figure 2 part of the questionnaire is given.

Figure 2. Example of the questionnaire

What were you doing?	What exactly?	Further?
1. job	1. main job 2. second job 3. unpaid job	
2. running the household	1. preparing food	1. breakfast 2. dinner 3. supper 4. other
	2. washing up 3. making tea or coffee 4. cleaning the house	1. dust 2. etc, etc.
3. personal care	1. taking a shower 2. etc, etc	

If, for example, the activity is preparing supper, the activity code is 213. First the respondent chooses the main category 'running the household' (2), secondly 'preparing food' (1) and then 'supper' (3). The code is registered by the computer program. We call this procedure coding by tree-structured questions.

With this procedure, it is easy to ask for specific additional information on the primary activity and it is possible to distinguish more activities than with a coding list because the task for the respondent is easier.

With respect to the time measure, we have chosen an open interval. The main reason for this choice is the expectation that a fixed time interval would put too much burden on the respondent, especially in a telephone interview.
Inadequate information is reduced by having immediate checks on the upper and lower boundary of answers, the time measure and sequences of activities, e.g. report of travel between indoor and outdoor activities (Verwey et al. 1987).
Two drawbacks of our procedure are the use of an open time interval and the fact that respondents are not provided with a list of activities. These are expected to cause a lack of understanding regarding the level of detail of the diary (Lingsom, 1979). To compensate, an example is given as an introduction to the diary exercise. We also put an extra question after each activity in order to check whether the answer given represents the performed activity and its duration. Possible mistakes could be reported in response to an open question. To get the time use precise, the respondent is asked whether he or she is certain that all activities are registered if the duration of the reported activity exceeds a certain amount of time[1].
With this electronic diary, data can be gathered in different ways. The questionnaire can be filled out either by a respondent or an interviewer. In this paper these two methods of data collection are compared. At a later stage the quality of the electronic diary will be compared to the conventional paper-and-pencil diary.
In both procedures the same questionnaire is used. Therefore the reporting and coding structure of activity and time, the number of primary and secondary activities, and the extra information asked about the activities, are all the same. The difference is in the registration method and some aspects of the implementation. The registration is considered to be the most important difference. But in order to get an idea of this difference we must first look at the implementation characteristics.

Implementation of the diary

Telepanel

One of the possible environments in which the electronic diary is used is the NIPO Telepanel (the Dutch Gallup Organization). This panel consists of a sample of about a thousand households which has been randomly drawn from an *address directory*. The NIPO has provided these households with a home-computer and a modem. With these facilities data are collected entirely automatically so that interviewers are no longer necessary. For more details about this panel we refer to Saris and de Pijper (1986), Van Doorn (1987) and Saris (1989).

---

[1] An interval of four hours or more is used where 'work' or 'job' is reported and three hours or more for all other activities.

The time budget survey in the NIPO Telepanel had to be conducted within *one week*. We chose one week in *November* because earlier research has shown that time use in the period October/November is closest to the annual average (Niemi, 1983).
At the end of October all individual Telepanel members of *12 years and older* were asked whether they wanted to participate in the survey. Two attempts were made to elicit cooperation from the household. In order to increase the willingness a lottery with one prize was announced. Out of the willing respondents only *one person per household* was randomly selected. This person was asked to fill out three questionnaires for *three different days* because Niemi (1983), Gershuny et al. (1986), and Oudhof et al. (1988) showed that after 2 or 3 days either the response or the quality of the diary decreased. These days had to be sampled in such a way that all different types of days were presented. In the Telepanel procedure six combinations of three diary days cover an equal amount of diaries per day of the week.
The quality of the data is also expected to be dependent on the time span between the diary day and the filling out of the questionnaire (Juster, 1985). Obviously, it is easier to remember activities and their duration the more frequently or recently the diary is filled out. Therefore, respondents in the Telepanel were asked to fill out the diary as soon as possible after each given day[2].
Demographic information on the panel is gathered from the household members on a regular basis. We did not ask for this information in the time use survey.

CATI

The second data collection method is a telephone panel[3]. The sampling frame consisted of a *telephone directory*. Out of this registry 1054 telephone numbers or households were selected at random[4]. To optimize the procedure every household was called at least ten times to reach the respondent for the first interview. To increase the response, an experiment was done whereby about 28% of the households received a letter in advance, and 49% of the respondents who refused at the first call were called back later. Only respondents who received no letter and refused because they had 'no time', 'did not want to participate in the survey', 'age' or 'unknown reason' were called back.
Interviews took place in *October and November*. Again *one person per household of 12 years or older* was randomly selected to report his or her activities for *three days*. During the first call the respondents were asked about their previous day's activities and about their background. Afterwards attempts were made to make an appointment for an

---

2 It is not a requirement that Telepanel respondents fill in the electronic diary several times a day because during that time the computer must be switched on.
3 For this research we made use of 19 students, most of them inexperienced. They were trained by ourselves. The interviews were performed on 7 days a week, during the afternoon and evening.
4 60 numbers were not taken into account because they represented disconnected numbers (45) or business numbers (15).

interview two days later and four days later. If the respondent could not be contacted on these days the interviewer tried to make an appointment respectively three and five days later for the same *diary day* as before. If even this was impossible, another diary day was selected because we expected that the *recall period* would otherwise become too long.

Difference in implementation between Telepanel and CATI and the effect on the data
From the above it becomes clear that dissimilarities in the implementation exist. In table 1 the consequences of these differences are summarized.

Table 1. Impact of implementation differences between Telepanel and CATI

Implementation differences	*Effect on*
coverage sampling frame ***H1***	coverage of personal response ***more complete coverage with Telepanel***
attempts to reach the respondent ***H2***	level of contact rate ***higher level of contact rate for CATI***
sampling frame and response stimulation ***H3***	level of personal response rate ***higher response rate for Telepanel***
determination of diary days ***H4***	coverage of day response and time use of activities ***better coverage of days and better estimates of out-of-home activities with Telepanel***
control over recall period ***H5***	time use of activities ***better estimates of activities with CATI***

The sampling frame can bias the results due to the fact that certain members of the population are not in the frame. When using a telephone directory one does not reach the population of non telephone owners (6%), nor the people who are not in the directory for other reasons (12%). This non covered population can be selective and therefore the results can be biased (Kerssemakers 1985, Kersten and Moning 1985)[5]. Since the most complete register is the address directory we expect the coverage, and thus the generalization of the results to the population, to be better in the Telepanel procedure[6].
The second dissimilarity, the number of attempts to reach the respondent, varies from a maximum of two (Telepanel) to a minimum of ten (CATI). Because we work with a panel that normally responds to questionnaires every week in the Telepanel we did not expect a very different contact rate.

5 In the Netherlands non coverage is expected to be higher for households with a young head, single person and large households, the unemployed, students, low income earners and skilled or unskilled labourers.

6 The reason for using the (incomplete) telephone directory is the possibility of obtaining, with relative ease and low costs, the address of the household and these people can then be sent a letter in advance.

However, using an existing panel must have an effect on the response rate. The expectation that respondents in the NIPO Telepanel would participate in the time use survey was very high because they already participate in the panel, and their cooperation would be further influenced by the lottery. Therefore, we expected a much higher response rate in the Telepanel than in CATI.

With respect to the choice of days, we have mentioned above that in the telephone survey the diary days were not completely determined in advance. Consequently, representativity according to the days of the week covered may be less. More importantly it can affect the time use. Kinsley and O'Donnell (1983) found differences in time use of activities when convenient day interviewing was used instead of designated day interviewing.

With the designated day system it is less likely that the number of hours spent on out-of-home activities will be overestimated. On the other hand, the designated day system can bias the time use of activities because of its extended recall period. In our study this recall period is more tightly controlled in CATI than in the Telepanel procedure. Telepanel respondents have the possibility of filling out the diary in a wider time range after the diary day than is possible in CATI because of interviewer working time.

If there are differences between the procedures with respect to implementation aspects, it is advisable to correct for these differences. Correction can generally be made by using weighting procedures if (specific) people do not respond or certain days of the week are badly covered.

It is less obvious how the researcher should react to differences in time use on activities due to the determination of diary days and control over the recall period. Since these effects exist, as we will see later, we had to make a decision. In order to analyze the effect of the registration method, we omitted the cases in which either the diary day had been changed or the recall period was extended beyond 24 hours. The possible registration effect is the topic of the next section.

Registration method

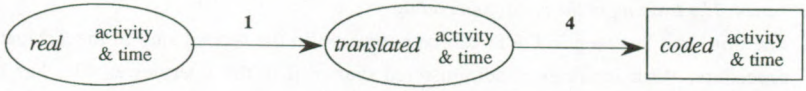
By now it is probably clear that the electronic diary is self-administered in the Telepanel procedure and interviewer-administered in CATI. We have tried to visualize the consequences of the difference between these administration forms. The result is presented in figure 3.

1. The training of the respondent

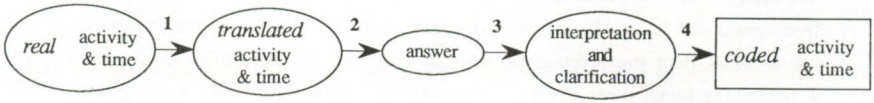
In the first step in both procedures the respondent's task is to translate his or her activity and its duration to our specified and desired level of report. In the Telepanel this means respectively choosing between 368 activities and reporting travel or activities that last for 10 minutes or more.

Figure 3. Process of activity and time administration

TELEPANEL



CATI



Of course, these activities are not displayed on one screen, but respondents can go back to the central question and change the direction in which the activity must be scored. Thus the respondents have a (visual) reference frame during the filling out of the questionnaire. In CATI the procedure is quite different. The activity reporting is open ended and the respondent is less aware of how detailed his/her report has to be (Lingsom, 1979). This will negatively affect the precision of the report. Also, the pace of the interview may have a negative influence on the quality. The natural pace of a telephone interview is somewhat faster than the natural pace of a self-administered interview. Juster (1986) found some evidence that slowing down the interview improved the quality of the time use data.
Given these two differences (reference frame and pace of the interview) in CATI we expected less activities, less variation in activities, more activities starting on the hour or half hour and more activities lasting more than 3 or 4 hours[7].

2. The presence of an interviewer

The presence of the interviewer in CATI causes a possible second effect. The respondent may try to create a positive impression on the interviewer by reporting only socially desirable activities. According to Robinson (1985) this source of bias is not supposed to be large in time use diaries because of the neutral way in which the diary interview is conducted. The chronology of such a questionnaire has a matter-of-fact character and is assumed to be far less directive or suggestive than questions like 'How many hours did

[7] The validity rule 'amount of not ascertained time' does not directly exist in the electronic diary because we opted for an open time interval.

you work last week?' or 'Please record all time that you watch television'. Nevertheless some activities appear to be underreported in diaries, probably due to social desirability effects. Examples are activities such as sex, gambling, fighting or stealing.

3 and 4. The training of the coder and coding

The third and last step in CATI can be compared to the second step in the Telepanel procedure. With interviewer-administered diaries it is the interviewer who has the overview of the kind of information required. The interviewer is well trained in the possible activities, and he or she can make significant contributions to increased data quality by posing additional questions, clearing up misunderstandings, etc. The interviewer is in a better position to make difficult classifications due to his or her knowledge of the respondent and the diary (Lingsom, 1979). Therefore, we expect that an interviewer will find and code the performed activities in the tree-structured questionnaire more easily than the respondent who has not done it frequently. Consequently, we expected the interviewer to report less mistakes with respect to the activity and time in open questions than the respondents in the Telepanel.

We have now formulated our expectations with regard to the registration differences. These hypotheses are presented in table 2.

Table 2. Impact of registration differences between Telepanel and CATI

Registration	*Effect on*
reference frame and pace of the interview ***H 6***	precision of activity and time report ***for Telepanel, more primary and secondary activities, more variation in activities, less activities starting on the hour/half hour, less activities lasting more than 3 or 4 hours***
presence of an interviewer ***H 7***	social desirability ***for Telepanel, fewer socially desirable answers***
coder ***H 8***	use of activity categories, number of mistakes ***for CATI, more use of activity categories, less mistakes***

In summary, we expect the Telepanel procedure to be preferable with respect to the coverage of persons and days, the response rate, the precision of the activity and time report, and social desirability. This procedure is expected to be less advantageous in the contact rate, the control over the recall period, the coverage of activity categories and the amount of mistakes made. In the next section we concentrate on the outcome of the hypotheses.

Results

Implementation

We first wanted to ascertain which procedure has the highest response rate. The basic results are presented in table 3[8]. Response rates can be calculated in different ways. Groves (1989) defines the most common rates, making a distinction between rates to evaluate field activities and a rate to estimate quantities that are related to the non-response error[9].

The response rates with respect to the field activities are the contact rate and the cooperation rate. The contact rate describes to what extent the sample is reached, and the cooperation rate, how well the field staff has persuaded those contacted and able to respond. The response rate is the proportion of the total eligible sample which provided three diary days.

Table 3. Response to Telepanel and CATI procedures

	TELEPANEL		*CATI*	
	abs	*%*	*abs*	*%*
Willing respondents	810	82.9	583	58.7
At least one diary	781	79.9	568	57.1
At least two diaries	775	79.3	533	53.6
Full response	748	76.6	520	52.3
Initial refusals	97	9.9	356	35.8
terminated during interview			15	1.5
Not contacted	70	7.2	41	4.1
Language difficulty			14	1.4
Available respondents	977	100.0	994	100.0

In CATI the contact rate is slightly higher than in the Telepanel: 95.9% versus 92.8% (supports H2). This can be due to the fact that in CATI at least 10 or more attempts were made to reach the respondent for an initial interview whereas only two were made in the Telepanel: the two weekends in October that the household was asked for its cooperation. The cooperation rate is, as expected, much higher in the Telepanel; 82.5% provided three diaries versus 55.4% in CATI. The most important rate, the response rate, is also much higher in the Telepanel; 76.6% versus 52.3% in CATI (supports H3). This difference in response rate is not caused by the partial non response (the percentage of people with less

8 These figures are cleaned for diaries with less than four activities and not relating to a special day and diaries with implausible activities that last more than 6 hours.

9 The rates are based on full response:
Contact rate (I+P+R+NI) / (I+P+R+NI+NC)
Cooperation rate I / (I+P+R)
Response rate I / (I+P+R+NC+NI)

I = completed interviews
P = partial interviews
NC = not contacted
R = refused interviews
NI = other non interviewed units

than three diaries is respectively 3.3% and 4.8%), but by the fact that in the Telepanel we made use of an existing panel.

In order to examine the quality of the data it is important to know whether the response is representative for the Dutch population or if the non response is selective. Therefore, for both samples we compared a demographic variable that is expected to be strongly related to time use and is known for the population[10]: the perceived position on the labour market. The results are summarized in table 4.

Table 4. Distribution of perceived position on the labour market

%	*TELEPANEL*	*CATI*
employed	41.0	43.5
unemployed, disabled etc.	12.8	8.6
scholar or student	9.2	10.9
housekeeper	19.8	21.1
retired, other	17.2	24.4
Total	100.0	100.0
n: number of respondents	781	568

An important difference between the Telepanel and CATI can be seen with respect to the unemployed (higher in the Telepanel). It is possible that this difference is caused by the difference in sampling frame (supports H1).

Another result is that in both the Telepanel and CATI too many retired people cooperated in the survey. This probably has more to do with the topic of our research than with the sampling frame since older people are slightly under-represented in the NIPO Telepanel. It is possible that these respondents have more free time and find it more important to report on their time use.

In both surveys we asked non respondents why they did not want to cooperate as we wanted to know whether their refusal had anything to do with their time use. Related reasons for refusals were 1) no time or too much work to do, and 2) not at home or not much at home. In the Telepanel procedure the former is the main reason for refusal (6%). In CATI 6% of the refusals were based on this reason. The second reason was mentioned equally frequently: in both procedures it was 2%. The differences due to time or place related reasons for refusal is thus also about the same.

Another feature of the response is the distribution of days over the week. In the Telepanel the diary days are, as expected, fairly equally distributed over the days of the week (χ^2 3.8, degrees of freedom 6). In CATI however, this distribution shows significant deviations (χ^2 22.5). Thursdays and Saturdays are underrepresented whereas Sundays are

10 These figures are based on statistics of the Central Bureau of Statistics, see Imbens (1990).

overrepresented. It was difficult to reach respondents on weekend days by telephone. Consequently, the diary day changed for some respondents (supports H4). The results are presented in table 5.

Table 5. Determination of diary days in CATI

First day: number of calls		*Difference between two diary days*	*first-second*	*second-third*
one call	45.8	one day	84.4	81.0
two calls	27.3	two days	7.9	9.6
three calls	13.7	three or more days	7.6	9.4
four calls	5.1			
five calls or more	8.0			
Total	100.0		100.0	100.0
n: number of respondents	568		526	524

In the CATI procedure respectively 84% and 81% of the respondents had the intended interval of one day between the diary days. For the first interview the amount of telephone calls distinguished respondents whose diary day was determined at random from those who were not at home on the given day. About three efforts were made each day to reach the respondent. Thus in 87% of the cases the first diary day was chosen at random.

If we compare the time use estimates of activities on random and non random diary days we find that more time on outdoor activities is reported on non random days. The activities for which time use estimates increased are personal needs, socializing, work or job and leisure. Less time was reported on indoor personal needs and indoor household activities.

However, in table 6 we see that for 76% (4.0+72.3%) of the respondents in the Telepanel (T) the time span between the filling out of the diary and the reported day consists of one day or less. In CATI (C) this figure is much higher: 88% (supports H5). The figures are about the same for the third day. The reason for this is that the normal response period in the Telepanel ends for all people after that third diary day. The control over filling out the diary is, as expected, less in the self-administration procedure.

Comparison of time use on activities between a recall period of 24 hours and a longer period shows that for the latter less time is reported on indoor media-related activity and more time on outdoor education and work or job activities (supports H4).

We conclude that all our hypotheses with regard to the implementation differences can be accepted. The Telepanel is the preferred procedure on the basis of the response rate. The methods are approximately equal good in gathering data for three diaries if cooperation is promised, the non response is selective and shows roughly the same pattern for both

procedures. In CATI the distribution of diary days is selective but the control over the recall period is better.

Table 6. Recall period (figures are based on report of the respondent and the interviewer)

	First day		*Second day*		*Third day*		*Total*	
Recall period	*T*	*C*	*T*	*C*	*T*	*C*	*T*	*C*
less than 24 hours	1.8		4.2		6.0		4.0	
1 day	69.7	94.8	66.3	86.4	81.0	82.9	72.3	88.2
2 days	13.1	1.8	14.6	9.4	9.5	11.6	12.4	7.5
3 or more days	15.4	3.4	14.9	4.2	3.5	5.5	11.3	4.3
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

As the non response and the (different) coverage of days of the week can bias the result and since these differences are not too large, we could weight the data in order to compare the quality of the time use[11]. The differences due to a non random diary day and extension of the recall period were taken into account by selecting only those respondents for whom both aspects were as intended by the experimental design. In the next section the time use on activities and its quality are compared after these corrections had been made.

Registration of time use

First of all, we looked at the validity aspects noted earlier. In table 7 a summary of these criteria is given.

In CATI the first interview yielded significantly less activities than the second and third, which yielded approximately the same number. Closer inspection of the participation rate per diary day shows that personal care activities in particular are reported less on the first day (71.5% versus 87.5% on the second and third day).

In the Telepanel procedure an approximately equal amount of activities is reported in the first and second interviews. The Telepanel tends to cover 5 to 10% (one or two) more activities on these interviews than CATI. The third diary covers significantly less activities. This diary consists mainly of weekend days. Since the average number of activities on weekdays is expected to exceed the weekend average, because people are more active on weekdays than on weekends (Lingsom, 1980), we regard the accuracy of the number of reported activities as good. With respect to the variation in activities, CATI is a bit better: respondents report about 8 different activities as against 7 in the Telepanel.

[11] For this research the weighting has been done by the Central Bureau of Statistics. Their method is based on an iterative proportional fitting algorithm and can be described by multiplicative models. Diary days are weighted seperately from demographic variables. For more information on this, see van de Pol et al. (1988).

Table 7. Validity criteria * weighted figures	*DIARY DAY TELEPANEL*			*DIARY DAY CATI*		
	first	*second*	*third*	*first*	*second*	*third*
# OF PRIMARY ACTIVITIES *						
% of respondents						
less than 11	2.6	4.3	5.3	12.3	5.0	3.3
11 - 15	23.3	22.3	31.2	28.8	22.9	24.6
16 - 20	28.9	31.4	28.9	30.0	37.4	38.9
21 - 25	25.0	23.5	24.5	15.0	20.0	20.3
26 - 30	12.1	10.7	7.8	8.7	8.0	8.2
31 - 35	5.9	5.8	2.1	3.8	5.6	4.3
36 or more	2.2	2.0	0.2	1.4	1.1	0.4
total	100.0	100.0	100.0	100.0	100.0	100.0
mean number	**20.4**	**19.9**	**18.2**	**17.7**	**19.3**	**19.0**
standard deviation	6.6	6.6	5.7	6.7	6.3	5.4
n: number of respondents	563	561	629	463	385	353
VARIATION IN ACTIVITIES *						
mean number	**7**	**7**	**7**	**7**	**8**	**8**
standard deviation	2	2	2	2	2	2
n: number of respondents	563	561	629	463	385	353
SECONDARY ACTIVITIES *						
% of activities						
listening to the radio	17.6	19.2	17.2	11.3	11.1	12.0
watching television	4.6	3.2	3.9	2.9	3.1	3.1
socializing	8.9	10.8	11.8	13.0	15.0	14.5
productive activities	1.0	0.9	0.8	0.5	0.3	0.5
other	4.5	3.7	3.3	5.3	5.6	5.1
total	**36.6**	**37.8**	**37.0**	**33.0**	**35.1**	**35.2**
n: number of activities	10456	10209	10445	7372	6719	6038
ACTIVITIES ON HOUR/HALF						
% of activities	**42.2**	**40.8**	**43.1**	**53.1**	**50.0**	**50.3**
n: number of activities	11555	11578	11516	8293	7468	7274
DURATION OF ACTIVITY						
% of activities						
more than 3 hours (most activities)	2.3	2.4	3.4	3.4	2.6	2.3
.. of which forgotten activity	0.5	0.5	0.5	1.1	1.0	0.8
more than 4 hours (work or job)	1.1	1.1	0.4	1.1	1.1	1.2
.. of which forgotten activity	0.3	0.4	0.2	0.6	0.6	0.7
total	**3.4**	**3.5**	**3.8**	**4.5**	**3.7**	**3.5**
n: number of activities	11555	11578	11516	8293	7468	7274
USE OF DIFFERENT ACTIVITIES						
% of categories (over three days)			**77.2**			**73.9**
n: number of categories			368			368
MISTAKES						
% of activities						
report of activity	4.6	3.4	3.0	1.2	1.4	1.1
report of time	1.0	0.5	0.4	0.2	0.1	0.1
total	**5.6**	**3.9**	**3.4**	**1.4**	**1.5**	**1.2**
n: number of activities	11555	11578	11516	8293	7468	7274

If we look at the secondary activities, we see that for the Telepanel in 37% of all primary activities, another activity is simultaneously performed. In CATI this figure is somewhat lower: 34%. This difference is merely due to the first diary day. Larger dissimilarities are found in the way the time is spent. In the Telepanel relatively more radio listening was recorded, in CATI more socializing.

In the Telepanel the activity started on the hour or half hour in 42% of all reported activities. In CATI this figure is 51%. Especially in the first diary day, the duration of many activities is rounded to half or full hours. We think that this difference is mainly caused by the pace of the interview[12].

The number of activities lasting more than three or four hours is about the same for both procedures. The only exception to this is (again) the first diary day in CATI. In CATI respondents also answered more often that they had forgotten to report an activity during that time interval.

In summary, the respondents in the Telepanel report their activities and time more precisely (supports H6). In the Telepanel more primary and secondary activities are recorded, there are less activities starting on the hour or half hour and less activities are forgotten.

The use of the different categories tells us something about the ease of finding the activity respondents performed. On the aggregated level, in the Telepanel survey 77% of the categories were chosen at least once, and in CATI 74% (does not support H8). Dividing the categories into those which appear on the first screen and those on the second screen, and comparing the results for the two procedures, we see few differences between the Telepanel and CATI.

On the individual level we see that the respondents in the Telepanel report significantly more mistakes with respect to the activity and/or the time than the interviewers in CATI (supports H8). The number of mistakes decreases after the first diary day, but does not reach the level of CATI even after three days. For time use on activities for a day, in the Telepanel 63 minutes and in CATI 21 minutes are incorrectly recorded.

To get an better understanding of the problems, we looked at the activities which were not recorded successfully. We analyzed the rates of participation in activities differentiated by education and age. On the basis of our experience, we assumed that respondents with higher education and middle-aged people are more able to fill out the diary themselves. In other words, their participation rates were expected to be the same for the two procedures.

---

12 The interview-time was in the Telepanel on average 35 minutes for the first diary (standard deviation: 15), 29 minutes for the second (sd=14) and 25 minutes for the third (sd=12). In CATI the figures were respectively 20 (sd=8), 17 (sd=7) and 15 (sd=6).

From this data it became clear that the activities which were most frequently incorrectly chosen were comparable with those for which differences are found in time use. It is also remarkable that although interviewers showed less mistakes, they report them for roughly the same activities. Most mistakes are made in relation to travel, indoor personal needs, indoor household, indoor leisure, indoor media, outdoor work or job and outdoor leisure activities.

The differences between the education groups of both procedures is smallest for the respondents with higher education. Especially participation in leisure, obtaining goods and services and socializing differ greatly between the education groups and favour the higher education group. If we look at age differences, we see in the Telepanel that young respondents make more errors with respect to household, socializing and leisure activities; middle-aged persons with obtaining goods and services and leisure activities and old-aged respondents with obtaining goods and services, leisure, socializing and travel activities.

Given the above reported differences between the procedures, we did not, however, expect to find large differences between the time use estimates of activities.

We first compared the time use of both procedures by using a common classification (Robinson 1977, 1985, Lyberg 1990). The results are shown in table 8 (T=estimates of the Telepanel, C=estimates of CATI and T-C=difference between Telepanel and CATI).

Table 8.
Time use estimates in minutes for a day (total time) and percentages of time allocated to major activity categories (weighted figures)

* Travel is classified differently in our time use survey

	TOTAL TIME			*PERCENTAGE OF TOTAL TIME*		
	T	*C*	*T-C*	*T*	*C*	*T-C*
Work related	121	146	**-25**	8.4	10.1	**-1.7**
Housework	131	119	**12**	9.1	8.3	**0.8**
Child care	19	15	**4**	1.3	1.0	**0.3**
Shopping	13	26	**-13**	0.9	1.8	**-0.9**
Personal care	680	669	**11**	47.3	46.5	**0.8**
(Adult) Education	49	58	**-9**	3.4	4.0	**-0.6**
Organizations	11	9	**2**	0.8	0.6	**0.1**
Social entertainment	75	98	**-23**	5.2	6.8	**-1.6**
Active leisure	40	40	**0**	2.8	2.8	**0.0**
Passive leisure	195	171	**24**	13.6	11.9	**1.7**
Travel *	104	88	**16**	7.2	6.1	**1.1**
TOTAL (24 hours)	1438	1439		100.0	100.0	
N: number of diary days	1753	1201				

Time use estimates are based on all available diary days and represent the time use over all persons for a day and the percentage of time allocated to the activity categories for this day. Although the differences in percentage of total time are not large, major difference between the two procedures are found if we compare the time use in minutes.

The results of our own main classification are presented in table 9[13]. In this table the time use estimates for a day, the percentage of respondents who actually participate in the activity and the time spent by them are shown for both procedures. These figures are, again, based on the total amount of diaries and provide us with more detailed information. The table shows that in the Telepanel more respondents report on indoor personal care and indoor pet care activities. With respect to indoor leisure, outdoor obtaining goods or services and outdoor socializing, the participation is less for the Telepanel.

Other important differences are found in the amount of time participants spent on the activities. Large differences between activities that are performed rather frequently are found in indoor pet care, indoor and outdoor work or job, indoor education or training, indoor media, outdoor personal needs, outdoor leisure and travel.

Earlier we stated the possibility that social desirability would cause a difference in time use between the two procedures. The most obvious socially undesirable activity is assumed to be television viewing. In the Telepanel and CATI about the same proportion of respondents report this activity. The amount of time spent on it differs: in the Telepanel respondents spent half an hour more of their time on watching television.

Comparison with an 'objective' time use measure, the so-called 'TV-meter'[14], shows that the amount of time spent on watching television per day is closest to the Telepanel estimate (supports H7). Robinson (1985), however, noted that television viewing does not appear to be influenced by respondent bias towards underreporting. He assumed that people tend to overreport the time spent on watching television because they report watching a program in its entirety when in fact they viewed a part of it.

However, a closer look at our data showed that the average time spent on specific programs is about the same for both procedures. The difference could therefore be due to the number of programs that respondents watched.

Another explanation could be a change between primary and secondary activity. Yet, in the Telepanel fewer secondary activities were performed during television viewing. Thus our conclusion for the time being is that television watching is underreported in the CATI procedure due to social desirability.

In summary, all of our hypotheses are (partly) accepted. We conclude that the precision of the activity and time report is a (bit) better in the Telepanel procedure.

---

13 At a later stage standard errors will be calculated. Because we deal with a sample of persons and a sample of days it is not possible to calculate these standard errors in the normal way.

14 This measure is provided by the 'NOS-Kijk en luisteronderzoek'.

Table 9. (weighted figures)
Time use estimates in minutes for a day (total time) and percentages of people who actually participate in the activity (participation rate) and the time spent by these people (participation time)

	TOTAL TIME			*PARTICIPATION RATE*			*PARTICIPATION TIME*		
	T	*C*	*T-C*	*T*	*C*	*T-C*	*T*	*C*	*T-C*
TOTAL INDOOR	1068	1031	**37**						
personal needs	618	616	**2**	99.1	100.0	**-0.9**	623	616	**7**
personal care	36	33	**3**	94.2	81.3	**12.9**	38	40	**-2**
child/adult care	17	13	**4**	19.3	15.6	**3.7**	87	85	**2**
pet care	6	5	**1**	18.3	8.4	**9.9**	35	60	**-25**
household activities	90	84	**6**	63.1	62.5	**0.6**	143	134	**9**
goods/services	0	0	**0**	0.7	0.0	**0.7**	46	0	**46**
home repair etc.	24	17	**7**	16.6	10.7	**5.9**	145	155	**-10**
gardening	4	8	**-4**	4.4	6.2	**-1.8**	92	123	**-31**
work or job	16	20	**-4**	10.5	7.8	**2.7**	153	251	**-98**
education/training	18	21	**-3**	13.6	13.7	**-0.1**	130	150	**-20**
religious practice	1	1	**0**	0.6	0.6	**0.0**	72	76	**-4**
organizations	3	1	**2**	2.7	0.8	**1.9**	122	131	**-9**
socializing	38	41	**-3**	34.7	34.9	**-0.2**	111	116	**-5**
leisure	22	32	**-10**	19.9	34.1	**-14.2**	108	93	**15**
media	168	136	**32**	85.1	83.8	**1.3**	197	162	**35**
home computer	5	3	**2**	4.8	3.0	**1.8**	104	94	**10**
diary	2	0	**2**	2.4	1.4	**1.0**	75	22	**53**
TRAVEL	104	88	**16**	90.2	92.7	**-2.5**	116	95	**21**
TOTAL OUTDOOR	269	322	**-53**						
personal needs	23	17	**6**	20.9	19.7	**1.2**	110	87	**23**
personal care	3	3	**0**	6.1	6.6	**-0.5**	55	51	**4**
child/adult care	3	2	**1**	2.9	3.0	**-0.1**	87	65	**22**
pet care	1	0	**1**	0.9	0.2	**0.7**	62	21	**41**
household activities	1	2	**-1**	1.4	2.4	**-1.0**	104	84	**20**
goods/services	13	26	**-13**	16.7	36.6	**-19.9**	77	70	**7**
home repair etc.	2	3	**-1**	0.7	1.6	**-0.9**	254	181	**73**
gardening	2	1	**1**	0.9	0.5	**0.4**	223	153	**70**
work or job	105	127	**-22**	24.5	28.1	**-3.6**	429	451	**-22**
education/training	31	37	**-6**	10.6	12.5	**-1.9**	295	300	**-5**
religious practice	3	3	**0**	2.8	2.8	**0.0**	92	95	**-3**
organizations	5	4	**1**	2.9	3.1	**-0.2**	154	146	**8**
socializing	38	60	**-22**	23.0	37.4	**-14.4**	163	161	**2**
leisure	36	34	**2**	21.9	24.8	**-2.9**	165	136	**29**
media	3	3	**0**	2.7	2.5	**0.2**	120	124	**-4**
TOTAL (24 hours)	1441	1441							
N:number of diary days	1753	1201							

The expected unawareness of how detailed the reports should be, hypothesized for CATI respondents, only holds for the first diary day. On the other hand, as we expected, more mistakes are reported in the Telepanel. It is possible that some of the differences found in time use can be reduced if we take these mistakes into account. At a later stage more

attention will be given to this point. The fact remains that for some differences we do not have an explanation.

Costs

Earlier we noted that the data for time use are gathered in two very different environments. The NIPO Telepanel is a panel which operates as part of a market research organization. CATI is conducted at the University of Amsterdam. In order to compare the costs, we have estimated the time spent on both surveys. In table 10 the time of professional staff, connect time and interviewer time is presented.

Table 10. Costs in terms of hours

	TELEPANEL	*CATI*
Number of respondents:	700	500
Professional staff	170	160
implementation	20	50
help desk	120	0
interviewer training	0	10
call back	30	0
administration and cleaning	0	100
Connect time	230	470
Interviewers	0	1200
Total	400	1830

The figures are based on a situation in which we have a panel with computers or the facilities to conduct a telephone survey. The result will not be surprising. In the Telepanel the total time required to gather the data is 400 hours. CATI takes much more time, merely due to interviewing.

Discussion

In general, our hypotheses are confirmed as can be seen in table 11. The results provide clear information on the response and the recall period. As we expected, the Telepanel is better in terms of the cooperation and response rate because it already existed. Also, the coverage of days of the week is better in the Telepanel, due to the choice for designated day interviewing. On the other hand CATI has a more favourable recall period. We also found that the dissimilarities in diary day choice and the control over the recall period had an effect on the time use estimates of activities.

Table 11. Outcome of hypotheses

support H1	*more complete coverage in Telepanel*
support H2	*higher level of contact rate in CATI*
support H3	*higher response rate in Telepanel*
support H4	*better coverage of days and better time use estimates of activities in Telepanel*
support H5	*better time use estimates of activities in CATI*
support H6	*more precise activity/time report in Telepanel*
support H7	*less social desirability in Telepanel*
partly support H8	*no extra use of activity categories in CATI, less mistakes in CATI*

After we had 'controlled' for the differences due to coverage, non response, choice of diary day and recall period, remarkable differences were still found in the time use estimates. Explanations for some of these differences can be found in the difference in precision of the activity and time report, and in the amount of mistakes. The Telepanel seems to be a bit more precise in activity and time reporting. On the other hand more mistakes are made: a certain amount of training seems to be important.

Nevertheless, the small differences in the aspects studied can not explain the large differences in reported time use between the different methods. Further research is needed. On the one hand we think it is important to examine whether interpretation and classification problems in our diary structure exist, and if so, how these influence the time use estimates. On the other hand is it possible that the samples differ on other time use related variables than the ones we have taken into account. In order to examine the former, an experiment is planned in which we will ask respondents to code a number of different activities. The latter is under discussion.

References

Gershuny, J., I. Miles, S. Jones, C. Mullings, G. Thomas & S. Wyatt (1986).
Time Budgets: Preliminary Analyses of a National Survey. The Quarterly Journal of Social Affairs, 1986, 2 (1): 13-39.

Groves, R.M. (1989).
Survey Errors and Survey Costs. John Wiley & Sons, New York.

Harvey, A. (1984).
Proposal for Multinational Cooperation in Time Budget Research. Saint Mary's University.

Imbens, H. (1990).
Bezigheden en voornaamste bezigheid (Principal activities of the dutch population).
In: Supplement Sociaal-economische maandstatistiek, jaargang 1990, no.3.

Juster, F.T. (1985).
Conceptual and Methodological Issues involved in the Measurement of Time Use.
In: Juster, F.T. and F.P. Stafford. Time, goods and well-being. Survey Research Center, University of Michigan.

Juster, F.T. (1986).
Response Errors in the Measurement of Time Use. Journal of the American Statistical Association. June 1986, Vol. 81, No. 394: 390-402.
Kerssemakers, F.A.M. (1985).
Telefonisch enquêteren. (Telephone Interviewing). In: CBS Select 3, Staatsuitgeverij, The Hague.
Kersten, H.P.M. and H.J. Moning (1985).
Differences in estimates due to changes in methods of data collection. Central Bureau of Statistics, Voorburg.
Kinsley, B.L. and T. O'Donnell (1983).
Marking time, Methodology Report on the Canadian Time Use Pilot Study 1981. Exploration in Time Use Volume 1. Department of Communications, Ottowa.
Lingsom, S. (1979).
Open and fixed interval time diaries. A comparison based on a pilot study on time use 1979. Statistisk Sentralbyra, Oslo.
Lingsom, S. (1980).
Advantages and disadvantages of alternative time diary techniques, a working paper. Statistisk Sentralbyra, Oslo.
Lyberg, I (1990).
Sampling, Nonresponse and Measurement Issues in the 1984/85 Swedish Time Budget Survey. Statistics Sweden, Research Methods Development 1990: 2.
Niemi, I. (1983).
The 1979 Time Use Study Method. Central Statistical Office of Finland, Helsinki.
Oudhof, J., I.A.L. Stoop, R. Luttikhuizen (1988).
Pilot study tijdsbestedingsonderzoek 1986 (M33). (Pilot study time use survey 1986). Staatsuitgeverij, The Hague.
Pol, F.J.R. van de, T. Jonkers en I. Stoop (1988).
Herweging tijdsbudgetonderzoek vanaf 1987; met multiplicatieve aanpassing. (Weighting time budget research from 1987; a multiplicative adjustment). Centraal Bureau voor de Statistiek, Voorburg.
Robinson, J.P. (1977).
How Americans use time. A social psychological Analysis of Everyday Behavior. Praeger Publishers, New York, London, page 9-21.
Robinson, J.P. (1985).
The Validity and Reliability of Diaries versus Alternative Time Use Measures. In: Juster, F.T. and F.P. Stafford. Time, goods and well-being. Survey Research Center, University of Michigan.
Saris, W.E. (1989).
A technological revolution in data collection. Quality & Quantity 23: 333-349.
Saris, W.E. and M. de Pijper (1986).
Computer assisted interviewing using home computers. European Research 14: 144-152.
Stoop, I.A.L., J. Oudhof (1989).
Designing a time use survey on informally productive activities. In: CBS Select 5. Staatsuitgeverij, The Hague.
Van Doorn, L. (1987).
Het gebruik van microcomputers in panelonderzoek. (The use of microcomputers in panel surveys.) In: Recente ontwikkelingen in het marktonderzoek. Jaarboek '87-'88 van de Nederlandse Vereniging van Marktonderzoekers.
Verwey, M., N. Kalfs, W.E. Saris, M. de Pijper (1987).
Tijdsbestedingsonderzoek middels Tele-Interviewing en de mogelijkheden voor segmentatie. (Time budget research by means of Tele-Interviewing and possibilities for segmentation). Massacommunicatie 4: 353-369.

Ontvangen: 15-01-1991
Geaccepteerd: 18-03-1991