Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 32 (2018) 229–241

Transportation Research Procedia

www.elsevier.com/locate/procedia

International Steering Committee for Transport Survey Conferences

Does supervision in multi-day travel surveys lead to higher quality? A comparison of two independent surveys

Martin Kagerbauer [a]*, Juliane Stark [b]

[a] *Karlsruhe Institute of Technology – Institute for Transport Studies, Kaiserstrasse 12, 76131 Karlsruhe, Germany*
[b] *Institute for Transport Studies, University of Natural Resources and Life Sciences Vienna, Peter-Jordan-Strasse 82, 1190 Vienna, Austria, juliane.stark@boku.ac.at*

Abstract

The aim of this paper is to analyze whether the level of supervision influences quality of response and results of multi-day surveys. For that reason, we compare a multi-day survey (seven consecutive days) with children and high level of supervision in Austria (UTS) with data of the same age group out of the Mobility Survey in the Greater Stuttgart Region (MOSt). The results show that supervision during the survey process has nearly no influence on fatigue effects. Nevertheless, the quality of answers in supervised surveys is be better in terms of reporting intermodality or trip stages.

Keywords: Multi-day surveys; supervision; children; travel diary; fatigue; data quality

1. Introduction

The quality of reported travel behavior is highly influenced by survey methods and survey processes. Especially in multi-day surveys, the willingness of reporting for the whole survey period is important to obtain reliable results. From literature, we know that in case of multi-day surveys fatigue is observed. This is reflected, for example, by a declining reporting accuracy and completeness of questionnaires in the course of the survey period.

In this paper, we analyze whether the level of supervision influences the quality of response and results of multi-day surveys. For that reason, we compare a multi-day survey (seven consecutive days) on teenagers in the age group

* Corresponding author. Tel.: +49-721-608 47734; fax: +49-721-608 46777.
E-mail address: martin.kagerbauer@kit.edu

10.1016/j.trpro.2018.10.043

of 12 to 14 years and high level of supervision in Austria (Unterwegs travel survey, UTS) with the data of the same age group out of a Mobility Survey in the Greater Stuttgart Region (MOSt). The MOSt survey is a multi-day household travel survey in the course of one week with very little supervision and is very similar to the German Mobility Panel (MOP) (Zumkeller et al., 2011) without panel character.

Our analyses are based on the hypothesis that persons in this age group are less conscientious in filling in a questionnaire over a period of one week – compared to adults. We assume that a higher level of supervision leads to less fatigue and higher data quality. However, given the fact that the surveys were not specifically set up for this methodological analysis, the challenge is to find reliable indicators and isolate the effects of the survey approach from other impacts.

This paper highlights whether supervision during the survey process can help to reduce fatigue and enforce the participants' motivation. In more detail, it describes indicators such as trip rates, share of mobile persons, number of travel modes to analyze the implication of supervision on response and data quality. As a main aspect, paper examines the temporal biases in seven-day diaries and analyzes advantages and disadvantages in both the forced and the self-motivated survey process.

The paper is structured as follows: Section 2 gives a brief overview about literature concerning fatigue effects in multi-day surveys. Section 3 includes a description of the data bases used. Section 4 presents the results of the explorative data analysis. The paper closes with a brief conclusion section (Section 5).

2. Literature and research question

Multi-day surveys collects information of intrapersonal variance travel behavior of people. Such surveys are appropriate, for example, for the exploration of the variability of travel behavior within a week, of specific journey types or target groups, the separation of day-to-day effects, or average and "rare" travel behavior (Pendyala et al., 1997). Such longitudinal surveys allow us to measure continuously intrapersonal transport behavior taking account of day-to-day variability in trip patterns. Multi-day surveys hold the risk of fatigue effects. According to Pendyala et al., (1997) fatigue is caused by under-reporting due to extensive exposure to continued surveying questions. This could result in declining completeness and reporting accuracy (Chlond et al., 2013). In-wave-attrition manifests itself in the end of reporting due to participants' declining motivation. Such temporal bias effects were identified in multi-day travel diaries. Former studies used declining mobility key figures such as indicators of fatigue effects on relative scale, e.g., trip rates per day in the course of one week, the share of people undertaking at least one trip a day on consecutive days, dropout rates, etc. Golob et al. (1986) described the total daily travel time, the number of non-home-based trips, and car-passenger trip stages as "temporal biases in multi-day travel diaries". Golob et al. (1986) analyzed the Dutch National Mobility Panel and found out that the primary reason of temporal biases is the increasing tendency over time for respondents reporting no travel on a given day at all. Further, the authors examined a day-to-day increase in the under-reporting of short walking trips; walking trip stages are increasingly under-reported even after accounting for the bias in walking-only trips. As experiences from the German Mobility Panel (MOP) show, survey participants are motivated at the beginning of the survey and complete their trip diaries daily (Chlond et al., 2013). Then, motivation declines, which is reflected in missing trips. At the end of the reporting period, trip rates increase again. Meurs et al. (1989) found that an increase in under-reporting is associated with the number of days the diary is kept. Moreover, shorter trips and less complex chains are more susceptible to under-reporting due to lower importance. Axhausen et al. (2002) stated that reporting fatigue could be interpreted in two ways: negatively, on the assumption that the respondents were unusually diligent and civic-minded, which would or would not influence their observed daily patterns. Positively, on the assumption that persons once recruited and well informed about the task ahead would participate for such extended periods in a reasonable way. In the study Mobidrive, a six-week travel diary in two German cities in 1999, no fatigue was detected in the reported journeys and trips (Axhausen et al., 2002).

Except the studies mentioned above, there are not that many studies on the effect of different survey designs of multi-day surveys on fatigue. In this paper, we compare two survey approaches of two independent multi-day travel surveys. Whereas one approach refers to a totally self-administered survey, the other one corresponds to a compulsory – not self-motivated – participation of respondents with on-site supervision.

Target group of our analyses are children in the age group of 12 to 14 years. In household travel surveys of German-speaking countries, usually, persons in this age group are treated like adults: This means that they have to fill in the

travel diaries on their own† and do not receive a special treatment and a special questionnaire respectively adapted to their age. This applies, for example, to the Austrian national wide travel survey (which covers all persons of a household from the age of 6) and the German Mobility Panel (travel diary from the age of 10). However, in German national household travel survey MiD the wording in the trip diary is adapted to children (10 to 13 years). According to De Leeuts et al. (2008) standardized questionnaires similar to questionnaires for adults can be used for the target group of pupils in early adolescence (12 to 16 years), because their cognitive functioning is already well developed and the memory capacity fully grown, but memory speed is not. However, the authors state that only adolescents aged 16 years and older can be fully regarded as adults with respect to cognitive development and information processing. Therefore, we hypothesize that persons in this age group are less conscientious in filling in a multi-day questionnaire compared to adults.

This paper analyzes biases in seven-day diaries and summarizes advantages and disadvantages in two survey approaches in Germany and Austria with regard to children in the age group of 12-14 years. It leads to the following questions without considering the issue of correction factors:

- How large are in-wave fatigue effects in this age group?
- Is it possible to prevent fatigue effects by using intensive supervision compared with compulsory approaches?
- Does supervision lead to higher data quality?

3. Data

Germany and Austria are neighbor countries and have more or less the same living conditions. Nevertheless, there are special cultural differences. This paper, however, does not focus on the absolute values of the travel behavior key figures, but on the level and the development of the values due to the seven-day survey period in each survey. In this paper, we analyze the Mobility Survey in the Greater Stuttgart Region (MOSt), which is self-administrated, and the Unterwegs Travel Survey (UTS) which has a high level of supervision. Both the MOSt and UTS survey are seven-day surveys with trip diaries and therefore suitable for this research issue.

3.1. Mobility Survey Stuttgart "MOSt"

The association of the Stuttgart Region (Verband Region Stuttgart) had commissioned a household travel survey in the city of Stuttgart including six surrounding districts (Greater Stuttgart Region) in the year 2009/2010. All household members were asked to complete a trip diary for one week, containing all relevant information about their trips, such as distances, modes and durations. (Kagerbauer et al., 2013). The data were used as input for a macroscopic travel demand model with VISUM as well as a microscopic multi-agent travel demand model with mobiTopp (Mallig et al., 2013). The Greater Stuttgart Region has 2.7 Mio inhabitants (610,000 inhabitants in the City of Stuttgart). The multi-day survey was conducted during one week between September to November 2009 and February to April 2010. In addition to a paper-and-pencil-based survey method, data on households, household members, and trips made, were also collected by telephone interviews and a web-based survey. The participants had chosen the survey method themselves. (Kagerbauer et al., 2013). The survey included a total number of 5,561 households with all household members. This survey was very similar in design to the KONTIV design with questionnaires similar to those for the German Mobility Panel (Kagerbauer et al., 2013). The survey participants were asked to fill in a questionnaire including information on the household (number of persons and especially children living in this household, number of cars in this household, information about cars, data about parking situation, public transportation connection, etc.) as well as person-based information for all persons in the household (age, gender, profession, travel behavior, etc.). The main part of the survey was a trip diary in which participants reported details on each trip. The content of trip diaries used in MOSt is quite similar to others and based on KONTIV-design (Brög et al., 1983): typical trip characteristics such as trip purpose, trip distance, trip duration, or travel mode were collected in the course of one week.

---

† It may be that parents supervise their children, but we do not have any information about that.

The communication method was a mix of paper-based and web-based or by telephone. Each household got a sheet with instructions as well as the questionnaires. This survey was fully self-administrated; the only way for the respondents to receive support was calling a hotline, but this did not happen very often. In the paper- and web-based approach, the participants decided at what time they filled in the questionnaires; day-by- day or at the end of the week. In the telephone approach the participants were called every second day in order to report their trips. To improve comparability of the data sets, we limited the data to those persons aged between 12 and 14 years (N = 625) in the following analysis.

3.2. Unterwegs Travel Survey "UTS"

The study has collected data of 192 children from four different schools in Austria and Germany, representing four socio-spatial settings (big city center, city outskirts, and two provincial towns). The pupils filled in personal travel diaries in the course of one week in spring, in consecutive years. The surveys were embedded in a research project that aimed at measuring the impact of a travel awareness campaign. For the analysis in this paper, we considered data of Austrian children (N = 156) in two following years (2013 and 2014), who did not take part in the campaign (control group). The German data were not taken into account because supervision for these children was not as intensive as for the Austrian ones. The design of the travel diaries was also based on KONTIV-design; however, it has been adapted for this age group: not all of the changes referred to problems of comprehension, but improvements in terms of wording, layout and content (Stark et al., 2015). The questionnaire contained usual trip information such as origin, destination, trip purpose, starting and ending times, etc. Some of the information (such as mode choice and trip duration) was collected on the level of trip stages. Because reporting travel modes in a chronological manner of their use was easily comprehensible to the pupils. Ticking all modes of transportation used on a trip (see conventional trip diaries) seemed too abstract for the pupils.

The survey was conducted using a paper-and-pencil questionnaire. The children were motivated to fill in their travel diaries by calling attention to the fact that they were contributing directly to significant scientific research. Researchers introduced the questionnaires (household questionnaire, travel diary); parts of the household questionnaire were filled in together. Something special about the UTS-approach is that the survey data were collected during regular school lessons. This was intended to prevent the children from being forced to fill in the questionnaire at home as part of their homework. However, if the respondents wanted to fill in their travel diaries at home, they were encouraged to do so – especially over the weekend. Researchers offered adequate support to the children in terms of filling in their travel diaries: In the course of seven days, children received support four times. This approach was designed to prevent the reporting procedure from being postponed until the end of the reporting period, which would affect the accurateness of the report. Special on-site supervision was required to report addresses and alternative modes of transportation per trip. The researchers also offered the opportunity to contact them by phone or by e-mail at any time during the survey period.

3.3. Similarities and differences of the surveys

Table 1 gives an overview of main characteristics of our data sources. The most important similarities of both surveys are seven consecutive reference days, the target group, and the use of travel diaries based on KONTIV-design. Moreover, there is a huge overlap in terms of the content of the travel diaries. In contrast to MOSt the respondents of UTS are the same persons in the consecutive years with only a very few exceptions. In the MOSt survey all participants are once in the sample. We assume that the different years of data collection have no influence on the in-wave effects.

Table 1: Overview of similarities and differences of data sources

	MOSt	UTS
Years of data collection	2009 & 2010	2013 & 2014*
Age of respondents in analyzed sample	12 to 14 years	
Survey design	PAPI/CATI/CAWI	PAPI
	Trip diary based on KONTIV-design	
Survey period	7 consecutive days	
Questionnaire	Self-administered questionnaire	
Time of survey	Fall and spring	Spring
Delivery of questionnaires	Sent by postal mail	Handed out in class
First reference day	Randomly over the week	Wednesday
Supervision	Self-administered survey, support by hotline and e-mail	Supervision face-to-face by researchers in class (four times a week), support by hotline and e-mail
Motivation for participation	On a voluntary basis without incentive	In class (as part of lessons) without incentive
N (persons)	652	64 (2013); 62 (2014)
Location	Greater Stuttgart Region (random sample drawn from official statistics including all inhabitants)	Selected schools in Austria (Vienna, Lower Austria)
Survey questionnaires	Household and person questionnaire, trip diary	Household questionnaire, address questionnaire, trip diary
Adaptations to the age-group	Standardized questionnaire	Adaptations of wording and visual design, trip duration instead of distance, travel modes in chronological sequence
Variables in trip diary	Trip purpose	
	Time of start and arrival	
	Travel modes**	
	Duration**	
	Address of origin and destination	
	Weather conditions	
	Trip length	-
	-	Accompaniment
	-	Decision on mode choice**
	-	Alternatives

**) A special analysis was conducted based on data collected in 2015*

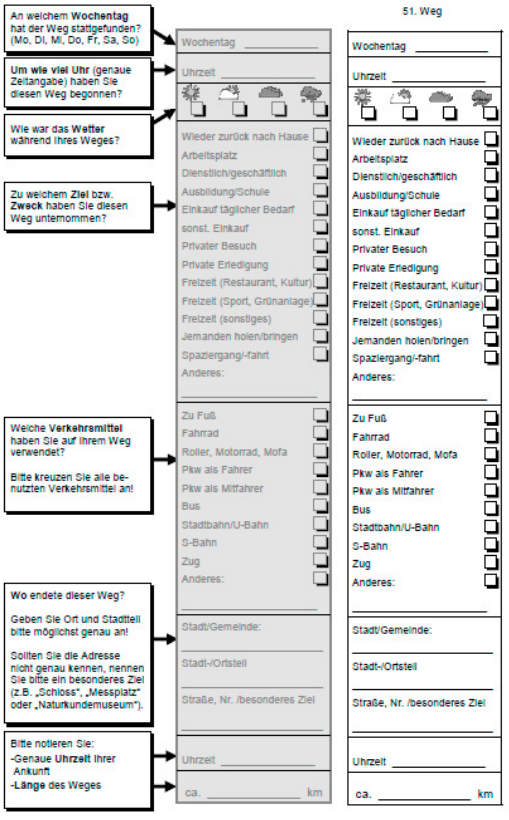
***) UTS collected this information per trip stage.*

The main difference between the survey approaches is the kind of supervision during the data collection process. The answer process of MOSt was totally self-administered without on-site support. This means that all information about the study as well as instructions needed to be transmitted by the questionnaire itself respectively by accompanying cover letters. This referred to the questions, response categories, and meta-information (study objectives, instructions how to respond to the questions, etc.). The questionnaires were sent by postal mail. Presumably, parents encouraged their children to fill in the questionnaires as part of a household survey or conducted proxy interviews.

In terms of UTS, (i) detailed instructions for the questionnaires were provided by the researchers at the beginning of each survey period, (ii) the travel diary was used in a group setting (classroom) and special supervision was offered by researchers at four days during the period of one week. More precisely, the questionnaire was in principle self-administered, but researchers provided assistance on-site; this also allowed individual explanations of uncertainties and immediate data checks. However, the questionnaires itself contained all instructions and information like usual self-administered paper-and-pencil questionnaires. The questionnaires were handed out in class. As the survey was embedded in a research project, for which the teachers' and parents' declaration of consent had been obtained, the children's participation was not self-motivated. This means that different levels of motivation were expected among the children. In the case of MOSt survey, persons with low motivation are likely to drop out in the course of the survey week.

In both surveys, the opportunity was offered to contact a researcher at any time during the survey period by phone or by e-mail. The trip information in the UTS consisted of nine characteristics of which three characteristics had to be reported on stage level. The duration for filling in the questionnaire is a little bit lower for MOSt than for UTS. As outlined above, the questions in the UTS questionnaire were adapted to the age group; the instructions were more detailed compared to MOSt. In terms of the visual design of the travel diary, there are differences with regard to the use of colors and the spatial arrangement of text and response categories (Figure 1). Trip information of consecutive trips are depicted in vertical columns. Overall, we assume that these differences in the visual design are not determining factors for the data quality and fatigue effects.

a



b

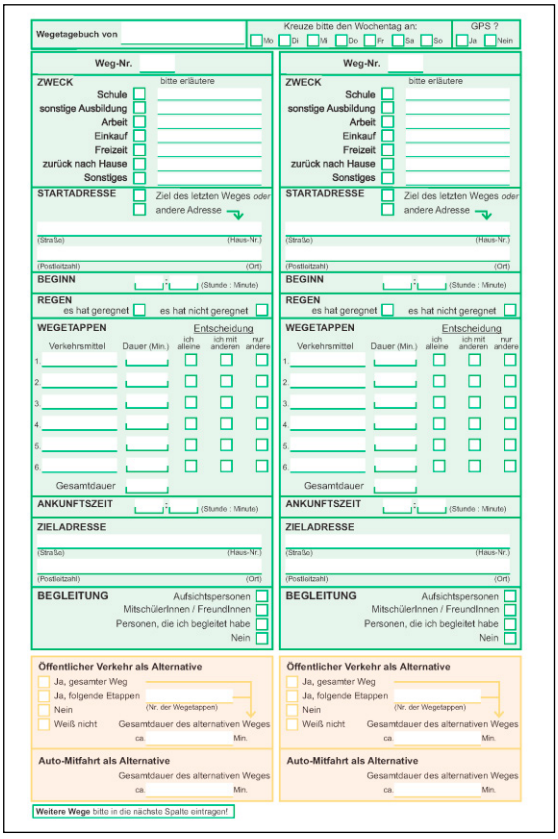


Figure 1: (a) Travel diary (MOSt); (b) travel diary (UTS), in German

4. Analyses

We used a descriptive data analysis to explore the impact of supervision on in-wave fatigue. Different indicators were calculated based on the two data sets. As mentioned before, differences in the absolute values of our indicators are less appropriate to derive conclusions on data quality respectively on advantages and disadvantages related to the data collection procedure. Therefore, relative differences (in the course of the survey period) have to be considered in order to make a statement how supervision affect the survey.

4.1. Trip rates

Figure 2 shows trip rates – measured in trips per person and weekday – as well as the related linear regression functions over the reporting period. Weekend days were excluded because the absolute numbers of trips differ between weekend und working day due to different activities. Again, for the comparison of MOSt and UTS, the level of trip rate per day was not of interest but the curve progression over the survey period of seven days.

The results show for both data sets a decline of the overall trip rate over the reporting period, but the differences are very low. The UTS data set shows some variation over the weekdays: it reveals a declining development over all days with slight increases on the second day and at the end of the period. The trips rates differ significantly between the days (Chi-square(4)=46.90, p=.000, n=126). Especially, they differ between weekday 4 and 1 (z=3.944, p=.001), 4 and 2 (z=4.701, p=.000), and between 5 and 1 (z=3.008, p=.026), 5 and 2 (z=3.765, p=.002). The regression model (F(1,628)=21.036, p=.000) and the coefficient (t=-4.586, p=.000) are statistically significant. The graph of MOSt is more constant with slight increases at the end and in the middle of the period. Overall, the reported number of trips per day has a higher stability for MOSt: The regression model (F(1,3258)=3.708, p=.054) and the coefficient (t=-1.925, p=.054) are not statistically significant. The differences between trip rates are significant, but on low level (Chi-square(4)=10.26, p=.036, n=652, mainly based on differences of day 4 and 3). From these analyses of MOSt it can be concluded, that there is only a slight tendency of fatigue effect. Unmotivated or low motivated people in the MOSt survey did not participate, thus, answers of extremely unmotivated participants might missing. As described above, for UTS it was assumed that respondents' low level of motivation was expressed in a decline or discontinuity of trip rates in the course of the reporting period because a total drop out was not possible.

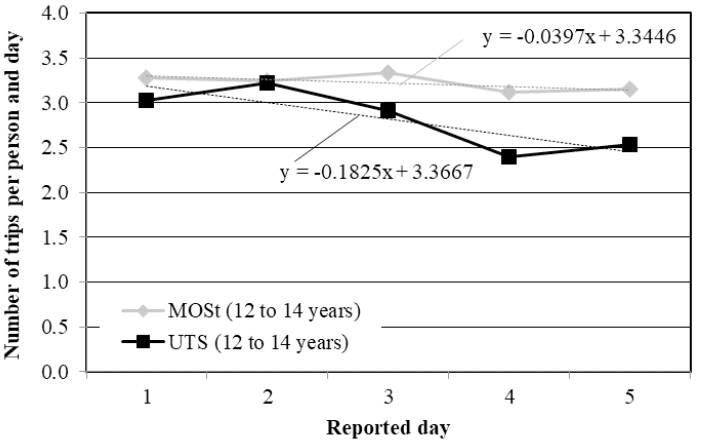


Figure 2: Trip rates by reported day and linear regression function, weekdays

In the UTS data set, the first day of the reporting period (day 1) is a Wednesday because the reporting period was the same for all of the participants, whereas for MOSt the starting days were randomly distributed. As weekend days were excluded from these analyses, day number 4 and 5 were the last two days of the reporting period in any case in the UTS dataset. In the case of MOSt this could also be day 4 and 5 within the seven-day period. To consider this aspect and to check for differences in the data sets, we analyzed trip rates in MOSt reflecting both the sequence day of each diary entry (day 1 through 5) and the starting day within the survey period: each starting day subsample was

considered separately for each weekday on which diary recordings were requested. On doing so, it is possible to observe for both sequence-day and day-of-the-week effects. Golob et al. (1986) also recommended this approach.

Table 2 shows that, as planned, about one-seventh of the sample starts on each day of the week in MOSt. The trip rates are similar over the starting day with a slight decrease. Nevertheless, the trip rates in MOSt are more stable over the reporting period compared to UTS, although starting day Wednesday shows the highest decline. As a partial outcome, these results lead to the conclusion that an intensive supervision does not obviously lead to a better reporting in trip rates over the survey period.

In order to verify this outcome, a special analysis was made using UTS-data of the year 2015. In this year, the multi-day survey had been repeated – under the same conditions (sample, time of survey, questionnaires, supervision via telephone hotline or e-mail, etc.) except for on-site supervision and incentive (Table 3): Children participated voluntarily and got an incentive (shopping voucher of 15€). Due to administrative reasons, it was not possible to supervise the children within the school lessons like during the two surveys before. It can be assumed that the motivation level was higher in 2015, although researchers tried to convey why the children's participation in the travel survey is very important.

Table 2: Decline in trip rates depending on starting day in the survey period

Data set	Starting day of reporting period	Count	Intercept	Slope
UTS	Wednesday	126	3.3667	-0.1825
MOSt	Monday	103	3.1282	0.0097
MOSt	Tuesday	91	3.4275	-0.0429
MOSt	Wednesday*	101	3.5950	-0.1277
MOSt	Thursday	97	3.3093	-0.0474
MOSt	Friday	85	3.5965	-0.1071

**) corresponds to UTS*

Table 3: Selected characteristics of data sets UTS in three consecutive years

Year	Age	Incentive	Participation	On-site supervision	Assumed level of motivation
2013	12/13	No	Not voluntary	Yes	Mixed
2014	13/14	No	Not voluntary	Yes	Mixed
2015	14/15	Yes	Voluntary	No	High

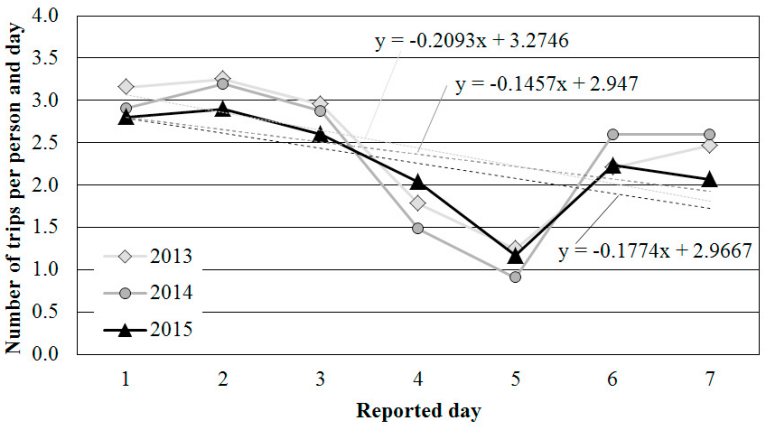


Figure 3: Trip rates by reported day and linear regression functions, UTS data set (1 - Wednesday to 7 - Tuesday)

Figure 3 shows UTS trip rates by reported day for each survey year. We can deduce that the fatigue effect is very similar – whether or not on-site supervision was offered. However, it should be considered that in 2015 the pupils reported the third time. This means that they knew the questionnaires very well and they were older. Moreover, the effect of the incentive cannot be clearly determined. Further, it can be assumed that only highly motivated pupils

took part in the third survey since their participation was optional (response rate 45% compared to 100% before).

4.2. Share of trip makers per day

Typically, survey participants start the reporting period with a high motivation (Chlond et al., 2013). Nevertheless, after the first few days, respondents stop filling in the trip diaries promptly and exactly respectively. Some of them stop reporting at all, whereas others fill in the diaries later with the risk of memory gaps. To analyze these fatigue effects in our data sets, we calculated the share of people with at least one trip a day for each data set. The results are depicted in Figure 4, illustrating the share of persons making at least one trip a day by reported day. This analysis only refers to workdays because the share of trip makers is less on weekend days compared to working days. Overall, the declining share of trip makers is quite stable (and on a very high level). For MOSt, the share of mobile persons is the highest for the very first day and declines slightly over the survey period (slope: -0.0031). For UTS, the share is somewhat more stable regarding the regression function, but with some outliers (day 2). However, the trend is similar in both surveys and the differences are not significant. The absolute level is higher for UTS than for MOSt, presumably, because of the compulsory answering process. From this result, it can be concluded that the quality of answers might be better because people normally forget trips instead of thinking up tips.

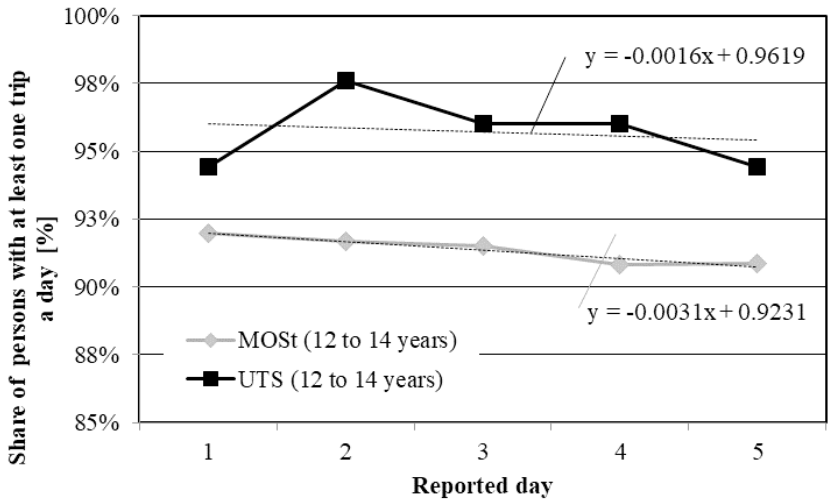


Figure 4: Share of persons making at least one trip a day by reported day (1: first reported working day (Mo-Fri) to 5: last reported working day (Mo-Fri) without weekend days)

4.3. Mode usage

Number of different modes used per trip (Intermodality)

The approach follows the assumption that fatigue effects in multi-day surveys may also come up in form of a decline in reported modes used on a trip – representing a loss of accuracy over time. The UTS-answers refer to trip stages, whereas the used modes per trips in MOSt is aggregated on trip level. In the MOSt questionnaire, different modes used on a trip were reported via ticking boxes. In the UTS questionnaire, the travel modes used were reported in the order of their usage. This questionnaire design, combined with high levels of supervision, was expected to lead to high data quality, although writing off words or mode-specific abbreviations is more time-consuming than ticking boxes.

Therefore, we analyzed the number of different modes used per trip. Figure 5 shows the results of the calculations. A comparison of the two data sets indicates that the absolute number of used modes in the UTS survey is slightly higher than in the MOSt survey, but both curves have a similar horizontal development. Within the data sets, no significant differences between the weekdays can be observed.

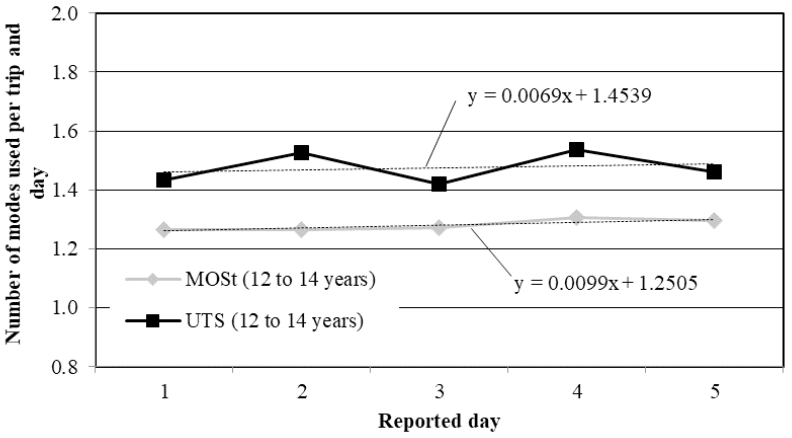


Figure 5: Number of different modes used per trip by reported day and linear regression functions (1: first reported working day (Mo-Fri) to 5: last reported working day (Mo-Fri) without weekend days)

The comparison of the graphs of all UTS-waves in three consecutive years shows scarcely any differences. Following the overall hypothesis, this seems a bit surprising, because there was no on-site supervision in 2015.

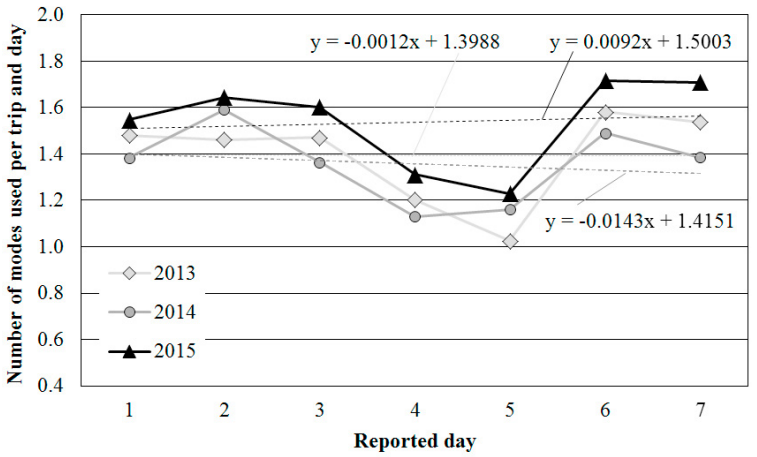


Figure 6: Number of modes used per trip by reported day and linear regression functions, UTS data set (1 - Wednesday to 7 - Tuesday)

Share of public transport trips, on which "walking" is reported

Walking trip stages are often under-reported (like walking trips and short trips in general – see, for example, Vanwolleghem et al., (2016); Golob et al., (1986). However, walking trip stages are part of nearly every public transport trip. On the assumption that the UTS approach leads to more accurate data due to the survey approach with reported stages, we calculated the overall share of public transport trips, on which walking is reported. A comparison of the two data sets indicates a significant higher share for UTS (76.8%) compared to MOSt (24.2%). This clearly indicates that the stage-based differentiation of transport modes is necessary when active modes, especially walking stages, are relevant to the evaluation. This refers, for example, to research on the level of physical activity of children. In contrast to the self-administrated survey, the supervision may lead to a better reporting of the access and egress stages especially for walking. To find out more about fatigue effects in the course

of the survey period, we examined the share of public transport trips, on which "walking" is reported per day (Figure 7). It can be derived that besides the absolute level of share of walking stages the development of the curves is very similar. Comparing the differences between the weekdays within each data set, no significant effects can be observed indicating no fatigue effect.

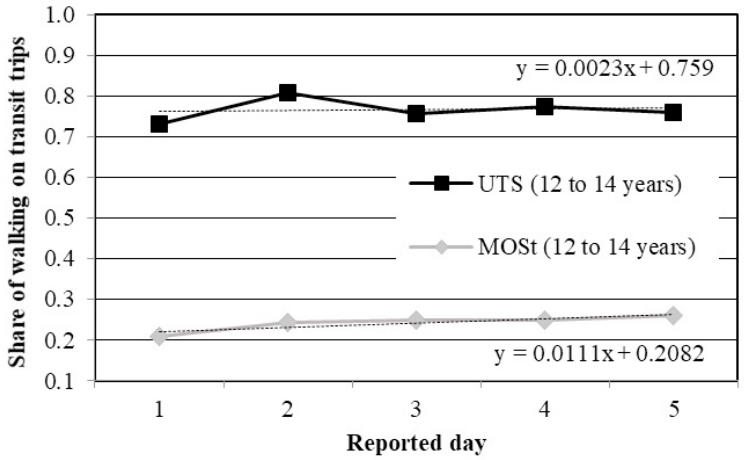


Figure 7: Share of public transport trips, on which "walking" is reported by reported day and linear regression functions

Item response rate

Another indicator for data quality is the item response rate, which is 94.5% for UTS. This high response rate was achieved because the survey was conducted during school hours under the supervision of teachers and researchers. Unfortunately, no comparison with MOSt is possible, as this indicator cannot be derived due to missing information of the plausibility checks.

Summary

With our analyses, we tried to identify fatigue effects and changes in data quality within the two multi-day surveys. Reasons for such effects might be that respondents are less willing or less focused in filling in the questionnaires in the course of the survey period. The observed indicators were (i) trip rates per person and day, (ii) share of mobile persons per day, (iii) number of different modes used per day, (iv) share of public transport trips with walking stages and (v) item response rate. Fatigue would be expressed by a decline for example of trip rates (i) and share of mobile persons (ii) over the period of reported dates. In terms of data completeness, it was also interesting to compare absolute differences between the two data sets (indicators (iv) and (v)).

Though there are some significant differences, we summarize that fatigue effects are on a very low level within both datasets. We observed no significant differences between the two data sets considering the slopes of the regression functions. Thus, supervision during the survey process seems to have no influence on fatigue effects. This is also confirmed by the comparison of UTS-only data using another reference year‡. No significant fatigue effects could be observed for the indicators in the context of mode usage. However, in terms of data completeness, we found the tendency that supervision may lead to more reliable results.

‡ As one limitation of this special analysis, it has to be considered that children in the third survey wave were older and got an incentive.

5. Conclusion

In our analyses, we have compared two independent multi-day surveys of 12 to 14 years old people in Germany and Austria in the course of one week. The main difference is that MOSt (household travel survey in the Greater Stuttgart Region) is self-administrated and UTS (Austrian survey of pupils) is highly supervised by researchers.

The main intention of our data analysis was to examine whether the on-site supervision during the data collection process has an influence on data quality, reporting accuracy, and fatigue. From these two data sets we have derived appropriate mobility key figures such as daily trip rates, share of people with at least one trip a day, or number of travel modes used within a trip. We have compared the survey-specific development of these indicators in the course of the survey period. The results show that the differences with regard to most of the indicators between the supervised and the self-administrated survey approach are not very distinct. This leads to the conclusion that self-administrated travel surveys are appropriate, also in this age group.

However, as outlined above, supervision leads to very low item non-response rates and helps to prevent missing information on trip chains, e.g., walking trip stages on public transport trips, which could be relevant for specific research questions (for example, health impacts and active mobility of children). Some quality indicators were not comparable such as item non-response; therefore, no conclusion can be drawn regarding this issue. This also refers to the content-related accuracy of self-reported data. For example, difficulties were observed during the UTS supervision process when children reported addresses for origin and destination of their trips and for reporting trip-specific alternatives.

As outlined above, it has to be considered that the sample size is very low. Another limitation is that some differences between the two surveys could not be completely eliminated, although most of the basic conditions of the surveys are comparable. Despite the country, this mainly refers to some differences in the questionnaires in terms of wording and layout. We do also not know, who exactly filled in the questionnaire within the MOSt survey. Further, but less relevant differences are, for example, the aggregation level of some information (trip stages versus trips) or the starting day of the reporting period. Due to a lack of data, it was not possible to derive every indicator for both surveys (e.g. item non response).

To sum up, intensive on-site supervision may have positive effects, if detailed information is required (e.g. intermodality, often forgotten trips or trip stages - particularly walking). Nevertheless, looking at fatigue effects in multi-day surveys, supervision has nearly no influence in dropouts. It can accordingly be concluded that – at least for this specific target group – survey designs without supervision in traditional household travel surveys deliver sufficiently accurate results for most purposes.

References

Axhausen, K. W., Zimmermann, A., Schönfelder, S., Rindsfüser, G., Haupt, T., 2002. Observing the rhythms of daily life: A six-week travel diary. Transportation, 29(2), 95-124. doi:10.1023/a:1014247822322.

Brög, W., Meyburg, A.H., Wermuth, M.J., 1983. Development of survey instruments suitable for non home activity patterns. Transportation Research Record 944, 1-12.

Bundesministerium für Verkehr, Innovation und Technologie (BMVIT), 2016. Österreich unterwegs 2013/2014 - Ergebnisbericht zur österreichweiten Mobilitätserhebung „Österreich unterwegs 2013/2014".

Chlond, B.; Wirtz, M,: Zumkeller, D., 2013. Data quality and completeness issues in multiday or panel surveys. In: Transport Survey Methods: Best practice for decision making. By Zmud, J.; Lee-Gosselin, M.; Munizaga, M. Emerald Group Publishing Limitied. ISBN 978-1-78190-287-5.

De Leeuts, E. D., Hox, J., 2008. Self-administered questionnaires: mail survey and other applications. In: International handbook of survey methodology. By de LeeUTS, Hox, Dillman. 239-263.

Golob, T. T., Meurs, H., 1986. Biases in response over time in a seven-day travel diary. Transportation, 13(2), 163-181. doi:10.1007/bf00165546.

Goulias, K. G., Pendyala, R. M., Kitamura, R., 1991. Updating a Panel Survey Questionnaire. *University of California Transportation Center*. UC Berkeley: University of California Transportation Center. Retrieved from: http://escholarship.org/uc/item/2g93m0gg.

Kagerbauer, M.; Bricka, S., 2015. Panel, Continuous, and Cross-Sectional Travel Surveys – Germany's Experience. 94th Annual Meeting of the Transportation Research Board, January 11-15, Washington.

Kagerbauer, M., Manz, W., Zumkeller, D., 2013. Analysis of PAPI, CATI, and CAWI Methods for a Multiday Household Travel Survey. In Transport Survey Methods - Best Practice for Decision Making, J. Zmud, M. Lee-Gosselin, M. Munizaga and J.A. Carrasco, eds. Emerald Group Publishing, pp. 289-304.

Kitamura, R., Bovy, P.H.L., 1987. Analysis fo attrition bisases and trip reporting errors for panel data. Transportation Research, 21A, 287-302.

Meurs, H., Van Wissen, L., Visser, J., 1989. Measurement bias in panel data. Transportation 16, 175-194. Doi:10.1007/BF00163114

Pendyala, R. M., Pas, E. I., 1997. Multi-Day and Multi-Period Data for Travel Demand Analysis and Modeling. Paper presented at the International Conference on Transport Survey Quality and Innovation, Grainau, Germany. Resource Paper retrieved from: TRB Transportation Research Circular E-C008: Transportation Surveys: Raising the Standard, Proceedings of an International Conference on Transport Survey Quality and Innovation.

Stark, J., Bartana, I. B., Fritz, A., 2015. Examining Mobility Behavior among Youth - A Progress Report. Transportation Research Procedia, 11, 481-491. doi:http://dx.doi.org/10.1016/j.trpro.2015.12.040.

Vanwolleghem, G., Schipperijn, J., Gheysen, F., Cardon, G., De Bourdeaudhuij, I., Van Dyck, D., 2016. Children's GPS-determined versus self-reported transport in leisure time and associations with parental perceptions of the neighborhood environment. International Journal of Health Geographics, 15(1), 16. doi:10.1186/s12942-016-0045-9.

Zumkeller, D., M. Kagerbauer, T. Streit, P. Vortisch, B. Chlond, Wirtz, M., 2011. Deutsches Mobilitätspanel (MOP) wissenschaftliche Begleitung und erste Auswertungen Bericht 2011. Institut für Verkehrswesen (KIT), Retrieved from: http://mobilitaetspanel.ifv.kit.edu/Downloads.php#Wissenschaftliche%20Berichte (16/03/2018).