DOI: 10.12758/mda.2016.004

Retrospective Measurement of Students' Extracurricular Activities with a Self-administered Calendar Instrument

Peter Furthmüller German Youth Institute (DJI)

Abstract

With the expansion of all-day schooling in Germany, students' extracurricular activities are being brought into greater focus in educational and social sciences. However, the diverse range of activities and individual biographies makes it difficult to gather data on the variety and periods of extracurricular activities in classroom-based surveys. This paper introduces a tailored calendar instrument that was applied by the Study on the Development of All-Day Schools (StEG) to retrospectively survey the activities of senior students since the fifth grade. Unlike other calendar applications, the calendar was not filled in by trained staff but self-administered by the students in a group setting. We discuss methodological issues regarding this procedure by examining the current state of research and by sharing experiences of tests of the instrument prior to the survey. By further analysing the survey data, we find no indication that the calendar task induces higher non-response as a result of overburdening the respondents. Calendar elements with an open-ended format resulted in heterogeneous reports, which were nonetheless mostly suitable for further analysis. According to our findings, the number of reported activities does not vary for students with longer intervals of retrospection. From our results, we conclude that a calendar instrument can be successfully applied in classroom-based surveys but should be implemented with a step-by-step procedure under a supervisor's guidance.

Keywords: calendar, retrospective, questionnaire, extracurricular activities, StEG



1 Introduction

Since the introduction and expansion of all-day schooling in Germany, new approaches to informal and non-formal learning have found their way into standard educational institutions. This expansion of educational content has been accompanied by an increase in autonomy and freedom at both an institutional and an individual level: schools are almost unconstrained by curricular, organisational and pedagogical prescripts when conceptualising their all-day programmes. Students, on the other hand, can now choose from a wider range of extracurricular activities not only outside the institutional context but also in school. Overall, these developments have led to more heterogeneous and individualised combinations of educational activities among adolescents. Students can pursue different activities at different ages and for different lengths of time. Investigating the complexity of extracurricular activity participation in large samples has been identified as an important but rare practice (see e.g. Feldman & Matjasko, 2005). The scarcity of appropriate studies might also be due to a lack of suitable research tools. It is a methodological challenge to survey extracurricular activities by theme in a precise and differentiated manner while at the same time recording their biographical sequencing including interruptions and overlaps.

In the *Study on the Development of All-Day Schools (StEG)*, researchers from the German Youth Institute (DJI) are investigating how participation in extracurricular activities affects the transition from school to vocational training at the "first threshold". Extracurricular activities are viewed from a biographical perspective and surveyed *retrospectively* among students in the ninth and tenth grade at secondary schools. Taking into account the heterogeneity of educational biographies and with a view to improving the quality of retrospective reports, StEG applied a tailored calendar instrument to study extracurricular activities since the fifth grade. Unlike most calendar applications, the calendar is not filled in by trained staff during face-to-face-interviews; instead, whole classes are sampled and the

Direct correspondence to

Peter Furthmüller, German Youth Institute (DJI), Study on the Development of All-Day Schools (StEG), Nockherstr. 2, 81541 Munich, Germany E-Mail: furthmueller@dji.de

Acknowledgements: The author would like to thank Bettina Arnoldt and Christine Steiner for their support and assistance with this paper.

¹ The DJI's project only conducts interviews at secondary schools that are not "Gymnasiums".

Different terminology is used for calendar instruments in the relevant literature, e.g. event history calendar, illustrated life history, life events calendar, life history calendar, life history matrix, month-by-month calendar, time axes or timeline (see Glasner & van der Vaart, 2009).

questionnaire is self-administered by the students. Thus, relatively large amounts of data can be collected in a short time, but it is not possible to check every retrospective report individually during the collection stage.

This paper seeks to introduce the calendar instrument for retrospective measurement of extracurricular activities and carries out analyses that are relevant to assessing data quality. In the first section, current methodological views on calendar instruments are summarised. The second and third sections contain background information on the study design and explain the construction and pretesting of the calendar instrument. The extent to which the calendar instrument is appropriate for collecting data about extracurricular activities is discussed in the concluding sections.

2 **Current State of Research**

2.1 The Methodological Challenges of Recalling and Dating

Compared to panel studies, retrospective surveys enable less expensive collection of life history information and faster availability of biographical data, and do not suffer from panel attrition. On the other hand, retrospective studies are considered risky in terms of memory bias (see Dex, 1995; Solga, 2001). Interdisciplinary research since the 1980s has led to a better understanding of how cognitive processes (e.g. for storing and retrieving memories), characteristics of events (e.g. passed time, salience), individual traits (e.g. gender, age) and contextual factors (e.g. interviewing situation, social desirability of topics) have a positive or negative effect on recall (see e.g. Dex, 1995; Sudman, Bradburn & Schwarz, 1996; Schwarz, 2007). It is essential, in this regard, that memories generally have to be reconstructed by mentally linking together multiple notions and that retrospection can be guided by cues (see Pohl, 2007, p. 34). Dating of events is an especially difficult task since only a few events are memorised with a "time stamp" (see Glasner & van der Vaart, 2008, p. 3) and respondents are usually required to recall an event as well as its context before they can date it (see Reimer, 2001, p. 16). Recall and dating can also be distinguished in terms of their error dimensions: while recall of events is mainly associated with memory gaps, the dating of events is also prone to timing errors (see ibid.; Auriat, 1993; Glasner, 2011).

How can a survey instrument be designed to minimise these risks? Balán et al. (1969) reported that the quality of retrospective data was improved through the use of a chronologically structured schedule. Freedman et al. (1988) referred to this idea when they developed the life history calendar after thorough pretesting. Like a coordinate system, the life history calendar provides a grid for dating biographical details by specifying axes for times and themes. The theme axis displays all

domains, issues and events on which respondents are asked to provide information. In this way, the time axis enables each of the topics to be dated with predefined time units. Events can be related to each other and the recording of associated dates is simplified and standardised by the grid (see ibid., p. 41; for recent examples see e.g. Das, Martens & Wijnant, 2011; Rudin & Müller, 2013).

Based on findings from the cognitive sciences, R. F. Belli (1998) showed how calendar instruments support the process of memory reconstruction. Memory is organised like a network where recalled events can serve as cues to stimulate further memories. A calendric presentation of life events encourages respondents to retrieve information via different pathways, namely thematic "top-down retrieval", cross-thematic "parallel retrieval" and temporal "sequencing" (see ibid., p. 394; Matthes, Reimer & Künster, 2007, p. 72). In these ways, calendar instruments promote a contextualisation of recall (Reimer, 2001, p. 100) since respondents can not only use temporal bounding strategies, but also relate events to each other.

2.2 Findings on the Quality of Calendar Data

The extent to which the calendar method leads to better data quality has been evaluated with non-experimental and quasi-experimental designs (for a synopsis see Glasner & van der Vaart, 2009). A comparison of results is difficult, however, because instruments and procedures for data collection differ considerably from each other in these studies. The calendar methods vary in terms of research themes, retrospective periods (from a few weeks to several years), the time scales that are used (e.g. years, months, days), graph designs (e.g. grids or timelines), survey techniques (e.g. face-to-face or CAT interviewing) and samples (e.g. adolescents or adults). Although mostly small and in some cases ambivalent effects on data quality are reported, the authors in general draw positive conclusions: memory performance seems to be improved by the use of calendar tools, which were successfully applied to the recollection of various events over the course of respondents' lifetimes in different studies (see e.g. Freedman et al., 1988; Caspi et al., 1996; Martyn, 2009). Specifically with regard to educational activities, Dürnberger, Drasch & Matthes conclude that a contextualised approach supports recall for retrospective periods of five years (see ibid., 2011). Calendar instruments produce more complete reports, particularly if events are of the distant past or difficult to remember (see Goldman, Moreno & Westhoff, 1989; Becker & Sosa, 1992; van der Zouwen, Dijkstra & van der Vaart, 1993; Engel, Keifer & Zahm, 2001; Belli et al. 2004; van der

³ For example: a top-down retrieval strategy could result in a (fictional) statement like "When I attended all-day school, I always attended the computer courses". A cross-thematic parallel retrieval could result in "In addition to all-day school I attended a private music school once per week". A temporal sequencing strategy could result in "After I stopped taking guitar lessons, I joined our school band".

Vaart, 2004; Yoshihama et al., 2005). Depending on their topic, calendric retrospections show medium to high consistency with data that was collected earlier from the same respondents (see Freedman et al., 1988; Caspi et al., 1996; Lin, Ensel & Wan-Foon, 1997; Belli, Shay & Stafford, 2001). It was also found that calendar instruments lead to more accurate dating, although the advantage over conventional question lists is sometimes small (see Becker & Sosa, 1992; van der Zouwen et al., 1993; Belli et al., 2007; Sayles, Belli & Serrano, 2010). Calendar instruments are also supposed to reduce heaping, i.e. respondents rounding time periods to typical values such as 6, 12 or 24 months rather than reporting the precise date (see Goldman et al., 1989; Becker & Diop-Sidibé, 2003). However, it was not always possible to confirm the reduction of heaping (see van der Vaart, 2004). The calendar method is particularly recommended if complex biographies with overlapping events need to be reconstructed (see Engel et al., 2001; Belli et al., 2007). On the other hand, complicated histories have also proven to be one cause of low reliability in calendar data (see Callahan & Becker, 2012). Collection, coding and clearing of calendar data is more demanding and advantages over conventional question lists depend on the survey topic. But Glasner and van der Vaart (2009) maintain that calendar instruments have never led to poorer quality of retrospective data than question lists (see ibid., p. 343). More recent findings suggest, however, that this may not be the case if events can be recalled very easily by the respondents (see Belli, Bilgen & Baghal, 2013).

It is repeatedly pointed out in the relevant literature that interviewers and respondents described calendar instruments as helpful tools for recall and conversation. Furthermore, respondents were often motivated to record their biographies as accurately and completely as possible with such a template (see Freedman et al., 1988; Hoppin et al., 1998; Engel et al., 2001; Belli et al., 2004; Martyn, Reifsnider & Murray, 2006; Belli et al., 2007). Some authors reason, however, that calendar instruments may increase non-response because they look particularly demanding at first sight and could discourage respondents (see Glasner & van der Vaart, 2009). Prior to their telephone interviews, van der Vaart and Glasner (2005) sent a calendar instrument to some of the survey participants as a memory aid during the conversation. While the response rate was only 39 percent in the group which had received a calendar, it reached 67 percent amongst those who had no access to such a supplement (quoted in Glasner & van der Vaart, 2009, p. 344). Belli et al. (2001), on the other hand, did not find an essential difference between the response rates of traditional question list surveys and calendar interviews (see ibid. p. 52). Taking additional results into consideration, the findings on non-response in calendar instruments do not allow a definite conclusion (see e.g. Mortimer & Johnson, 1999; Yoshihama et al., 2005; Martyn et al., 2006; Cotugno, 2009).

In most of the present studies, the calendars were filled in by trained interviewers and not respondents themselves. Likewise, CATI designs contain calen-

dric grid views primarily for the interviewers to reveal gaps and inconsistencies in biographical reports and to clarify them.⁴ There only appear to be a few studies in which calendar instruments were administered by the respondents themselves. In a panel study, Mortimer and Johnson (1999) sent out a life history calendar annually to collect data about important events and activities in the respondents' lives. Cotugno (2009) successfully applied a self-administered calendar in a question list paper-and-pencil-interview with more than 200 participants. Martyn and Martin (2003) interviewed adolescents about sensitive topics like drug abuse and sexual behaviour using an event history calendar (EHC). To fill out the form, the participants could choose the support of a trained interviewer or administer the EHC themselves. About 86 percent (n=43) decided to fill in the EHC autonomously. A self-administered calendar instrument seems to comply with the need for confidentiality and therefore reduces social desirability bias when intimate and private domains are being surveyed. "In addition, when one-on-one interviews are not required, the EHC can be administered to groups of participants like large-scale surveys are administered, obtaining comprehensive data while saving time and money" (see Martyn, 2009, p. 73). Based on comparative data from prior surveys, personal interviews and plausibility analyses, Martyn views the self-administered calendar method as a suitable instrument for youth research (see ibid. 2009).

As an interim summary, we can note that calendar instruments are not only recommended because they improve recall and stimulate more complete reports, but also because they are an effective survey technique for recording a multitude of events in a compact format. Therefore, calendar instruments seem particularly suited for retrospective collection of data about complex biographies of extracurricular activities. Self-administered calendars have already been employed successfully in the past and their application in a classroom setting seems promising.

3 Data Base

StEG is a research programme being carried out by a consortium of several institutions that conducts surveys on all-day schooling in Germany on a regular basis. The participating institutions are the German Institute for International Educational Research (DIPF), the Institute for School Development Research (IFS), Justus Liebig University Giessen (JLU) and the German Youth Institute (DJI). StEG is sponsored by the German Federal Ministry of Education and Research (BMBF).

Based on a representative survey of head teachers from all over Germany in 2012, the participating institutions of StEG carry out in-depth studies on the effects

⁴ No significant differences could be found in face-to-face and CATI calendar interviews (see Freedman et al., 1988, p. 65).

of all-day schooling. The DJI's in-depth study is conducted with a subsample of secondary schools ("Schulen der Sekundarstufe I") that are not grammar schools ("Gymnasiums") and participated in the survey of head teachers.⁵ 65 all-day schools from 12 federal states agreed to in-depth studies as part of the StEG programme. 6 This article contains the data of 1,901 students in graduating classes who were interviewed with a self-administered paper-and-pencil questionnaire in spring 2013. Of these adolescents, 608 were in the ninth grade and 1,292 were in the tenth grade. The questionnaire consisted of a tailored calendar instrument and a classic question list on different topics, e.g. school careers and family background.

4 **Pretesting and Tailoring the Calendar**

The calendar instrument was intended to enable analysis and profiling of extracurricular activities, i.e. to collect data on the kinds of activities carried out by students as well as the biographical periods of activity. Like the question list section of the questionnaire, the calendar was supposed to be filled in by the respondents as autonomously as possible. To identify problems, the instrument was pretested with students in three different classes prior to the survey.⁸ The test gave no indication that the respondents were overwhelmed by the task of dating events with a calendar instrument in principle. It became clear, however, that respondents often had to deal with subtasks that were implicitly imposed or assumed by the instrument, which made it more difficult to fill in the grid as a whole (e.g. converting recalled dates to the calendar scale). Pretesting provided valuable feedback and suggestions for how to improve the calendar instrument. The most important steps for creating the final instrument (see Appendix A-1 and A-2) are explained in the following sections.

Theme axis: open question about activities

While other calendar instruments typically cover many domains on the theme axis, the calendar in StEG focusses on extracurricular activities to keep the task as simple as possible. The theme axis of the pretested calendar included 12 categories

The DJI's survey was carried out in cooperation with the IEA Data Processing and Research Center (DPC) in Hamburg.

The participating schools are from Bavaria (n=15), Baden-Württemberg (n=7), Brandenburg (n=4), Bremen (n=4), Hesse (n=4), Mecklenburg-West Pomerania (n=4), Lower Saxony (n=7), North Rhine-Westphalia (n=4), Rhineland-Palatinate (n=7), Saxony (n=3), Saxony-Anhalt (n=4) and Thuringia (n=2).

Depending on their individual school career, students in the ninth and tenth grades are typically 15 to 16 years old.

One ninth-grade class at a lower secondary school and two preparation classes at a vocational school ("Berufsgrundbildungsjahr" and "Berufsvorbereitungsjahr" respectively).

which brought together a wide range of activities from different domains. To match each recalled activity to a date, the respondents first had to assign them to a specific category by reading and interpreting the labels provided. If the adolescents were simultaneously engaged in various activities within a category and at a certain date, they were supposed to fill in the number of parallel activities for that point in time. *Matching* and at the same time *counting* activities turned out to be a very difficult task and the calendar was substantially revised in this regard for the main survey: instead of asking about the type of activities with a matching task, predefined categories were abandoned in favour of an open-ended format (see Appendix A-2). Since the respondents could list every single activity separately with the new format and did not need to summarise them in any way, counting also became unnecessary in the final instrument.

Time axis: grades as a personalised time scale

In the pretest, students preferred the calendar's time axis to be scaled according to school grades. Grades seem especially suitable for dating extracurricular activities at schools since they correspond to the institutionalised schedule there. They structure every individual's school career and provide further cues for recall. It cannot be assumed, however, that every respondent passes through school grades in the same sequence: while some of the students have had a regular career, other students may have repeated or skipped specific grades. To take individual school careers into account, the calendar's time axis was not completely labelled with preset values in its final implementation. Instead, the grades had to be filled in by the students according to their school career after the fifth grade. They were specifically instructed to write down a grade multiple times if it had been repeated.

Narrowing the reference frame of relevant activities

Asking about activity types using an open-ended question involves the risk of respondents ignoring activities they consider inappropriate and reporting trivial ones instead (for further discussion see also Sudman et al., 1996, p. 55). To convey the range of relevant activities more clearly to the respondents, an additional task preceded the calendar instrument in the main study: the respondents were simply asked to recall what they did on a regular basis apart from attending school lessons and to record their thoughts on a separate sheet of the questionnaire ("memo", "memorandum" or "Merkzettel", see Appendix A-1). The memo contained one column for recalled activities from the in-school domain, and another one for activities from the non-school domain. In order to establish boundaries for the scope of activities that were of interest to StEG, a *reference frame* was outlined in simple terms: the respondents were instructed to only write down activities that were not lessons and that they had attended regularly (at least once per week) and steadily (for the duration of at least one term) at any point since fifth grade. Moreover, the

range of relevant activities was specified by the supervisors through explanations and examples. It turned out to be very helpful to add an index to the columns' rows, since it encouraged respondents to fill in their activities more consistently.

Detailed instructions and step-by-step progress

Written guidance for the calendar was often not read by the pretest participants or was perceived as being too difficult to understand. If students tried to clarify uncertainty at all, they did so by asking their seatmates or the supervisor. This created a disturbance, irritated other students and distracted the whole class from filling in the questionnaire. Thus, a different approach was chosen for the main survey: written instructions were almost completely replaced by guidelines for verbal directives given by the supervisors. The supervisors instructed the respondents step by step and presented an example of each task on a poster-sized calendar. Care was taken to ensure that the students began each subtask together and got to ask questions if necessary. The whole procedure was organised as follows:

- 1. Filling in the memo
- 2. Labelling the calendar's time axis with completed grades
- 3. Transcribing 9 relevant activities from the memo to the calendar's theme axis
- 4. Dating each activity and completing the calendar

Analogous to the memo's separate columns for in-school and non-school activities, two calendar grids were included on different pages of the final questionnaire. Such a step-by-step process may seem contrary to the rationale of a typical calendar, which integrates all recall tasks into one instrument. However, the memo was not only supposed to convey an idea about the relevant activities, but also to provide the basis for the dating task, since the respondents have to recall some activities first before they can date them (see also Reimer, 2001, p. 16) and are thereby provided with cues for further activities.

The designation "Merkzettel" was supposed to connote that the memos' content is only temporary and auxiliary for the respondents. In the main survey, the instructors merely asked the students to copy their activities from the memo to the calendar ("Bitte übertragt alle schulischen Angebote aus dem Merkblatt auf S.4 in die leeren Zeilen auf der linken Seite unter der Anleitung 'Trage hier deine Aktivitäten ein'"). The respondents were not asked to keep the entries on the memo and calendar in sync nor were they told not to add further entries to the calendar. For future applications of the procedure, however, it might be beneficial to invite respondents more explicitly to add further activities while filling out the calendar.

5 Assessment of Data Quality

In the following, an attempt will be made to assess the quality of the calendar data and to evaluate the validity of the school careers and extracurricular activities recorded with the instrument. With calendar instruments the *gold standard* of validity assessment includes a comparison of retrospective data and similar information gathered from the same individuals at an earlier date (see Alwin, 2009, p. 283). An example of this rare type of evaluation can be found in Belli et al. (2001). However, an alternative strategy has to be pursued if retrospective interviewing with a calendar instrument is applied due to the fact that no data was gathered in the past, as is usually the case. "Establishing the validity of survey measurement is difficult because within a given survey instrument there is typically little available information that would establish a criterion for validation" (Alwin, 2009, p. 282).

In StEG too, no longitudinal data is available that would allow a *gold standard* analysis. But verifiable quality criteria can be derived from the current state of research and experiences in the pretest: firstly, it can be assumed that overburdening students induces frustration or uncooperativeness with regard to the calendar task that, in the end, will lead to higher non-response. Therefore, high non-response would indicate that the calendar is too complicated and not suited for classroombased self-administered surveying. Secondly, the quality of the calendar must be judged by the degree of activities entered in the calendar that are relevant, plausible and applicable for the research topic. Thirdly, the calendar would be insufficiently qualified for ascertaining extracurricular activities if biographical patterns can be ascribed to shortcomings of memory and recall. These issues will now be examined in greater detail.

5.1 Willingness to Fill in the Calendar

Despite the adjustments that were made after the pretest, the calendar instrument confronts respondents with an unfamiliar and relatively complicated task that demands a high level of cooperation. Due to the classroom-based nature of the survey, there is no individual supervision and few ways to deal with frustration and refusal to cooperate. With a self-administered instrument, data quality primarily depends on a sufficient number of completed calendars, but the data does not allow direct identification of uncooperative respondents: if students did not fill in the calendar, the reason may be deliberate refusal or the simple fact that no activities were carried out or could be recalled. A clearer picture emerges if the step-by-step progress is taken into account when assessing non-response (see Appendix B-1).

The differences between the filled-in segments on the memo page show that more respondents (n=1,901) reported activities for the non-school domain (92.7%) than for the school domain (88.5%). Only a few of the adolescents who filled in

the memo omitted to list activities on the corresponding calendar page (in-school: 0.4%; non-school: 1.3%) and almost all of the students also dated the activities they entered on the theme axis of the calendar. All students were asked by the supervisors to fill in the time axis of the calendar with their completed grades, regardless of whether they could recall activities or not. This instruction was followed by 97.5 percent of all 1,901 students for the in-school calendar and 95 percent for the nonschool calendar respectively.

If lack of cooperation was the main cause for missing data, the proportion of respondents who filled in neither the calendar page for activities in school nor the page for non-school activities (no table) should be high: 2.5 percent of the respondents did not supply dated activities for any domain, but only 0.9 percent repeatedly neglected the instruction to fill in the time axes with grades. This implies that there are almost no students who refused to cooperate at every step of the procedure. For comparison: the average proportion of non-response in the question list part of the questionnaire is 1.9 percent (n_{max}=1,901 with 358 items). Hence, willingness to fill in the calendar instrument can be regarded as positive. Missing data seems instead to be caused by a lack of activities being carried out or recalled within the specified boundaries.

5.2 Thematic Classification of Recalled Activities

The calendar was expected to yield heterogeneous textual data on a broad range of different activities that needed to be standardised in some way. Prior to the survey, a coding scheme was developed based on the scope of relevant activities, earlier findings and additional Internet research. The scheme was intended to enable individual activities to be recorded distinctively whilst mapping them to thematic categories. It includes 372 detailed activities in 15 categories. In a sense, the coding scheme also represents an ex ante explication of the thematic variety assumed in students' activities.

Appendix A-3 shows the proportions of adolescents (n=1,901) who reported activities from the 15 thematic categories. The distribution parallels known results in some respects (on the dominant role of sports, for example, see Züchner, Arnoldt & Vossler, 2008; Grgic & Züchner, 2013), and differences between the in-school and non-school domains suggest that the calendar data is able to portray social realities. As stated above, only a rudimentary assessment of the "correctness" of reports is possible since no precedent data is available from these respondents.¹⁰ But the StEG questionnaire also included some extra questions on activities that were supposed to be more difficult to survey with a calendar. A comparison of both survey methods demonstrates the limitations of the instrument: with the calendar, only three percent of the adolescents autonomously recorded that they had once attended "homework support", a more formal programme that is typical at all-day schools. However, when they were asked directly, 23 percent of the students confirmed that they had attended "homework support". A third of the students affirmed the direct question "Do you earn money with a side part-time job?", but less than four percent recorded such a job in the calendar. If activities in school are strongly associated with formal lessons or if non-school activities are weakly associated with leisure, they may have been cued less often by the calendar. As expected prior to the survey a specific query about activities that may not be fully covered by the boundaries of relevant activities presented in the memo or the calendar seems a useful precaution to prevent missing data.

Furthermore, Appendix A-3 shows the proportion of students' statements that were difficult or could not be assigned to one of the 15 thematic categories. The Miscellaneous category includes all activities that could be identified as some sort of informal or non-formal practice, but were not more precisely attributable (e.g. statements like "project group" or "all-day programme"). Around one in ten students recorded in-school activities that were completely unusable and were therefore classified as invalid (mostly because they could not be read or were crossed out). Non-school activities were treated as invalid mostly due to deviations from the reference frame (e.g. statements like "meeting friends", "chilling out", "parties" or "shopping"). Approximately 16 percent of all respondents made an entry for either the in-school or non-school domain which was not usable for further analysis. The invalid and vague activities can also be related to the number of total records: in total, 5,181 entries were made for in-school activities of which 345 (6.7%) were considered invalid and 149 (2.9%) could not be assigned to the scheme. The respondents entered a total of 4.598 records in the calendar for non-school activities. Of these, 200 (4.3%) were invalid and 96 (2%) were too inaccurate for further classification. Hence, the vast majority of the information provided could be assigned to either a specifically defined activity or to at least one thematic category.

¹⁰ It seems worth noting that, inter alia, the comparison of activity distributions from different studies depends heavily on the specified boundaries and the reference frame that is applied: StEG was gathering data on extracurricular activities that were regularly practised from the fifth grade for at least one term. In contrast, other studies sometimes concentrate on activities that are relevant at the time of the survey, and hence also include shorter periods of activity (for example).

5.3 **Number of Recalled Activities**

The calendric survey is supposed to reduce the risk of activities being reported incompletely, particularly for grades that are dated far in the past. The calendar would perform unsatisfactorily in this regard if fewer activities were reported for larger retrospective intervals.

The students reported significantly more activities for the non-school domain than the in-school one (in-school: M=0.81 (SD=0.97); non-school: M=1.38 (SD=1.25); n=1,901; p<.001)¹¹ and grades without activities are more frequent within the in-school domain (see Appendix A-4). Based on the calendar data, a distinct trend can be noted with regard to the extracurricular activities in both domains. While respondents entered fewer in-school activities for higher grades (r=-.11; n=1,901; p<.001), the sample differs with regard to non-school activities: on the one hand, the proportion of adolescents who did not report any activity in higher grades at all also increases slightly; on the other hand, the number of students who carried out multiple activities at the same time is higher. The non-school domain tends to feature more numerous activities in higher grades (r=.07; n=1,901; p<.001). The calendar data seems consistent with earlier findings on extracurricular activities by adolescents in this respect. Using a previous data base, StEG has already demonstrated an age-dependent decline of extracurricular activities at allday schools by analysing longitudinal data of students from fifth to ninth grade (see Arnoldt, Furthmüller & Steiner, 2013; Züchner & Arnoldt, 2011). The developing activity pattern in the non-school domain corresponds to the findings of the study "Medien, Kultur und Sport bei jungen Menschen (MediKuS)" which revealed a common shift in activities, meaning that growing up does not necessarily imply a withdrawal from activities, but rather involves a change of contexts (see Grgic & Züchner, 2013, p. 258).

Neither the in-school nor the non-school data show a linear growth of inactive students for grades further back in the past. However, remarkably few activities were recorded for the fifth grade as the lower boundary of the retrospective interval. The sampling design can be utilised to examine whether this deviation should be ascribed to a lack of activity or rather to the passing of time, difficulties of recall and deficiencies of the instrument: the sample is composed of students for whom different periods of time have elapsed since the fifth grade. Adolescents who were in the ninth grade at the time of the survey and followed a regular career in school were fifth-graders four years prior to the interview. But respondents in the tenth grade or students who needed to repeat a grade must look back at least five years in order to recall the events of the fifth grade. If the calendar instrument could not provide sufficient support for recall, respondents with larger retrospective intervals should have recorded fewer activities for the fifth grade. This assumed relation

¹¹ In the following, the number of activities in repeated grades was not accounted for.

between activity count and retrospective interval size is examined with a Poisson regression model (see Appendix B-2 and B-3). To control other influences on extracurricular practice, the model includes independent variables that have been shown to be relevant for the attendance of all-day programmes (see e.g. Steiner, 2011; Steiner & Fischer, 2011; Züchner et al., 2008).

In school, adolescents from former East Germany carried out more extracurricular activities in the fifth grade, as did students with a migration background and those of high socioeconomic status. School context seems to be relevant, since students recorded fewer activities if they had attended fifth grade at a different school. The higher activity count of students who participated in all-day programmes in the fifth grade is unsurprising, but a promising sign of the retrospections' validity. Respondents who attended the fifth grade six years prior to the survey reported fewer activities than the reference group with a retrospective interval of four years, but not to a significant level. Thus, despite some other reasonable findings, no statistically significant influence of retrospective interval lengths was observed with regard to the recalled number of in-school activities. For the non-school domain, adolescents from the eastern states of Germany, with a migration background or of low socioeconomic status recorded significantly fewer activities than the reference group. Students who had attended all-day programmes in the fifth grade were also more active outside of school. While these controlled variables exhibit different patterns in relation to the number of in-school and non-school activities in the fifth grade, the retrospective intervals do not: the number of recalled non-school activities does not differ significantly depending on how far the students had to look back.

6 Summary

Self-administering a calendar instrument not only involves recalling and dating, but a wide range of secondary tasks that are sometimes carried out simultaneously by the respondents. A step-by-step approach and verbal guidance by supervisors proved to be the most important measure to enable the application of a self-administered calendar in a classroom-based survey. By gradually labelling the calendar's axes, the students created a personalised scheme which relieved them of most of the work of reading and interpretation. However, while this decreased difficulty for the respondents, it increased complexity for the researchers. With its open-ended components, the calendar requires greater effort to code and process the data, since variability usually needs to be standardised in some way prior to analysis. By far the biggest challenge in surveying extracurricular experiences with an open-ended format is to convey an idea of the relevant activities the respondents are supposed to enter. The preceding memorandum turned out to be a practical solution that further facilitated the calendar procedure for the students. Although most of the recorded

activities could be interpreted and categorised after the survey, the accuracy of listed activities and dates remains undetermined since no data is available for comparison. However, the non-response rates give no indication of major problems due to frustration or a lack of cooperation. Furthermore, no differences in the number of recalled activities were found for respondents with longer retrospective intervals up to seven years. These results also seem to be an affirmation that respondents were supported in their main task of recalling and dating activities, and that, in contrast to the pretest, they were no longer overwhelmed by the instrument or diverted by complicated secondary tasks. Against this background, the calendar instrument has proven successful in gathering retrospective data about the complex biographies of extracurricular activities.

References

- Alwin, D. F. (2009). Assessing the validity and reliability of timeline and event history data. In R. F. Belli, F. Stafford & D. F. Alwin (Eds.), Calendar and time diary methods in life course research (pp. 277-301). Thousand Oaks: Sage.
- Arnoldt, B., Furthmüller, P. & Steiner, C. (2013). Ganztagsangebote für Jugendliche Eine Expertise zum Stellenwert von Ganztagsangeboten für Schüler/innen ab der 9. Klasse. München: Deutsches Jugendinstitut (DJI).
- Auriat, N. (1993). My wife knows best a comparison of event dating accuracy between the wife, the husband, the couple, and the Belgium population register. Public Opinion Quarterly, 57(2), 165-190.
- Balán, J., Browning, H. L., Jelin, E. & Litzler, L. (1969). A computerized approach to the processing and analysis of life histories obtained in sample surveys. Behavioral Science, 14(2), 105-120.
- Becker, S. & Diop-Sidibé, N. (2003). Does use of the calendar in surveys reduce heaping? Studies in Family Planning, 34(2), 127-132.
- Becker, S. & Sosa, D. (1992). An experiment using a month-by-month calendar in a family planning survey in Costa Rica. Studies in Family Planning, 23(6), 386-391.
- Belli, R. F. (1998). The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. *Memory*, 6(4), 383-406.
- Belli, R. F., Shay, W. L. & Stafford, F. P. (2001). Event history calendars and question list surveys: A direct comparison of interviewing methods. *Public opinion quarterly*, 65(1), 45-74.
- Belli, R. F., Lee, E., Stafford, F. P. & Chou, C.-H. (2004). Calendar and question-list survey methods: Association between interviewer behaviors and data quality. Journal of Official Statistics, 20(4), 185-218.
- Belli, R. F., Smith, L. M., Andreski, P. M. & Agrawal, S. (2007). Methodological comparisons between CATI event history calendar and standardized conventional questionnaire instruments. Public Opinion Quarterly, 71(4), 603-622.
- Belli, R. F., Bilgen, I. & Baghal, T. A. (2013). Memory, communication, and data quality in calendar interviews. Public opinion quarterly, 77(S1), 194-219.

- Callahan, R. & Becker, S. (2012). The reliability of calendar data for reporting contraceptive use: evidence from rural Bangladesh. *Studies in Family Planning*, 43(3), 213-222.
- Caspi, A., Moffitt, T. E., Thornton, A., Freedman, D., Amell, J. W., Harrington, H., Smeijers, J. et al. (1996). The life history calendar: A research and clinical assessment method for collecting retrospective event-history data. *International Journal of Methods in Psychiatric Research*, 6, 101-114.
- Cotugno, M. Z. (2009). Creating a self-administered event history calendar. *Survey Practice*, 2(9). Retrieved from www.surveypractice.org
- Das, M., Martens, M. & Wijnant (2011). Survey Instruments in SHARELIFE. In M. Schröder (Ed.), *Retrospective data collection in the survey of health, ageing and retirement in Europe. SHARELIFE Methodology* (pp. 20-28). Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).
- Dex, S. (1995). The reliability of recall data: A literature review. *Bulletin de Methodologie Sociologique*, 49(1), 58-89.
- Dürnberger, A., Drasch, K. & Matthes, B. (2011). Kontextgestützte Abfrage in Retrospektiverhebungen. *Methoden Daten Analysen*, 5(1), 3-35.
- Engel, L. S., Keifer, M. C. & Zahm, S. (2001). Comparison of a traditional questionnaire with an icon/calendar-based questionnaire to assess occupational history. *American Journal of Industrial Medicine*, 40, 502-511.
- Feldman, A. F. & Matjasko, J. L. (2005). The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of Educational Research*, 75(2), 159-210.
- Freedman, D., Thornton, A., Camburn, D., Alwin, D. & Young-DeMarco, L. (1988). The life history calendar: A technique for collecting retrospective data. *Sociological Methodo-logy*, 18, 37-68.
- Glasner, T. (2011). Reconstructing event histories in standardized survey research: cognitive mechanisms and aided recall techniques. University of Amsterdam, Amsterdam. Retrieved from www.researchgate.net
- Glasner, T. & Vaart, W. van der. (2008). *Cognitive processes in event history calendar interviews: A verbal report analysis.* In S. Balbi, G. Scepi, G. Russolillo & A. Stawinoga (Eds.), Proceedings of the Seventh International Conference on Survey Methodology. Retrieved from library.wur.nl
- Glasner, T. & van der Vaart, W. (2009). Applications of calendar instruments in social surveys: A review. *Quality and Quantity*, 43(3), 333-349.
- Goldman, N., Moreno, L. & Westoff, C. F. (1989). Collection of survey data on contraception: An evaluation of an experiment in Peru. Studies in Family Planning, 20(3), 147-157.
- Grgic, M. & Züchner, I. (Eds.). (2013). *Medien, Kultur und Sport. Was Kinder und Jugendliche machen und ihnen wichtig ist. Die MediKuS-Studie*. Weinheim/Basel: Beltz Juventa.
- Hoppin, J. A., Tolbert, P. E., Flagg, E. W., Blair, A. & Zahm, S. H. (1998). Use of a life events calendar approach to elicit occupational history from farmers. *American journal* of industrial medicine, 34(5), 470-476.
- Lin, N., Ensel, W. M. & Wan-Foon, G. (1997). Construction and use of the life history calendar: Reliability and validity of recall data. In I. H. Gotlib & B. Wheaton (Eds.), Stress and adversity over the life course: Trajectories and turning points (pp. 249-272). Cambridge: Cambridge University Press.

- Martyn, K. K. (2009). Adolescent health research and clinical assessment using self-administered event history calendars. In R. F. Belli, F. Stafford & D. F. Alwin (Eds.), Calendar and time diary methods in life course research (pp. 69-86). Thousand Oaks: Sage Publications.
- Martyn, K. K. & Martin, R. (2003). Adolescent sexual risk assessment. Journal of Midwifery & Women's Health, 48(3), 213-219.
- Martyn, K. K., Reifsnider, E. & Murray, A. (2006). Improving adolescent sexual risk assessment with event history calendars: A feasibility study. Journal of Pediatric Health Care, 20(1), 19-26.
- Matthes, B., Reimer, M. & Künster, R. (2007). Techniken und Werkzeuge zur Unterstützung der Erinnerungsarbeit bei der computergestützten Erhebung retrospektiver Längsschnittdaten. Methoden - Daten - Analysen, 1(1), 69-92.
- Mortimer, J. T. & Johnson, M. K. (1999). Adolescent part-time work and postsecondary transition pathways in the united states. In W. R. Heinz (Ed.), From education to work. cross-national perspectives (pp. 111-148). New York: Cambridge University Press.
- Pohl, R. (2007). Das autobiographische Gedächtnis: Die Psychologie unserer Lebensgeschichte. Stuttgart: Kohlhammer Verlag.
- Reimer, M. (2001). Die Zuverlässigkeit des autobiographischen Gedächtnisses und die Validität retrospektiv erhobener Lebensverlaufsdaten: Kognitive und erhebungspragmatische Aspekte. Materialien aus der Bildungsforschung (Vol. 71). Berlin: Max-Planck-Institut für Bildungsforschung.
- Rudin, M. & Müller, C. (2013). Kann es sein, dass ich das Beginndatum falsch erfasst habe? Methoden - Daten - Analysen, 7(3), 433-463.
- Sayles, H., Belli, R. F. & Serrano, E. (2010). Interviewer variance between event history calendar and conventional questionnaire interviews. Public Opinion Quarterly, 74(1), 140-153.
- Schwarz, N. (2007). Cognitive aspects of survey methodology. Applied Cognitive Psychology, 21(2), 277-287.
- Solga, H. (2001). Longitudinal surveys and the study of occupational mobility: Panel and retrospective design in comparison. Quality & Quantity, 35(3), 291-309.
- Steiner, C. (2011). Teilnahme am Ganztagsbetrieb. In N. Fischer, H. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Eds.), Ganztagsschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG) (pp. 57-75). Weinheim/Basel: Juventa.
- Steiner, C. & Fischer, N. (2011). Wer nutzt Ganztagsangebote und warum? Zeitschrift für Erziehungswissenschaft, 14(3), 185-203.
- Sudman, S., Bradburn, N. M. & Schwarz, N. (1996). Thinking about answers: The application of cognitive processes to survey methodology. San Fransisco: Jossey-Bass.
- Vaart, W. van der. (2004). The time-line as a device to enhance recall in standardized research interviews: A split ballot study. Journal of Official Statistics, 20(2), 301-317.
- Vaart, W. van der & Glasner, T. (2005). Enhancing recall accuracy in surveys by cognitively tailored timeline methods: A record check study. Paper presented at the First EASR Conference, Barcelona.
- Yoshihama, M., Gillespie, B., Hammock, A. C., Belli, R. & Tolman, R. M. (2005). Does the life history calendar method facilitate the recall of intimate partner violence? Comparison of two methods of data collection. Social Work Research, 29(3), 151-163.

- Zouwen, J. van der, Dijkstra, W. & Vaart, W. van der. (1993). Effects of measures aimed at increasing the quality of recall data. *Bulletin de Methodologie Sociologique*, 39(1), 3-19.
- Züchner, I. & Arnoldt, B. (2011). Schulische und außerschulische Freizeit- und Bildungsaktivitäten. In N. Fischer, H. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Eds.), Ganztagsschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG) (pp. 267-290). Weinheim/Basel: Juventa.
- Züchner, I., Arnoldt, B. & Vossler, A. (2008). Kinder und Jugendliche in Ganztagsangeboten. In H. Holtappels, E. Klieme, T. Rauschenbach & L. Stecher (Eds.), *Ganztagsschule in Deutschland. Ergebnisse der Ausgangserhebung der Studie zur Entwicklung von Ganztagsschulen (stEG)* (pp. 106-122). Weinheim/München: Juventa.

Appendix A

Deine Aktivitäten

Innerhalb der Schule

Schreibe hier nur Angebote auf

- die du irgendwann seit der 5. Klasse besucht
- die KEIN Unterricht waren
- die du mindestens ein halbes Jahr gemacht
- in denen du regelmäßig aktiv warst (mindestens 1 Mal pro Woche)

Außerhalb der Schule

Schreibe hier nur Aktivitäten auf

- · die du seit der 5. Klasse gemacht hast
- · die du mindestens ein halbes Jahr gemacht
- in denen du regelmäßig aktiv warst (mindestens 1 Mal pro Woche)

	. '
1. FUBBALL	1
2. Foto - AG	2
3. COMPUTERKURS	3
4.	4
5.	5
6.	6
7.	7
8.	8
9.	9
10.	1
11.	1
12.	1
13.	1
14.	1
15.	1

1.	KIRCHENCHOR
2.	TENNIS
3.	KLAVIERUNTERRICHT
4.	TANZEN
5.	
6.	
7.	
8.	
9.	
10.	
11.	
12.	
13.	
14.	
15.	

Figure A-1. Example of a filled-in memorandum for in-school and non-school activities

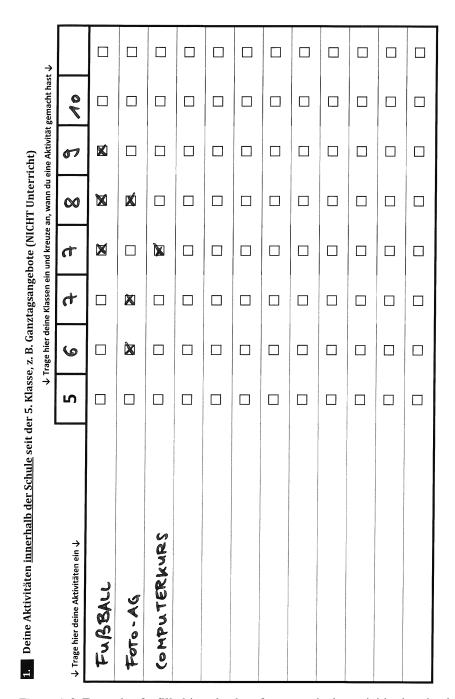


Figure A-2. Example of a filled-in calendar of extracurricular activities in school



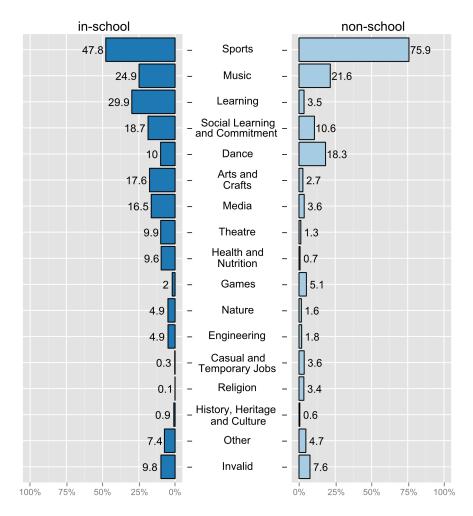


Figure A-3. Percentage of respondents with entered activities by categories and domain (n=1,901)

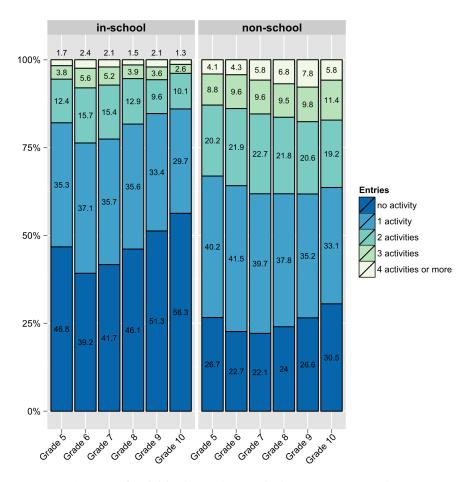


Figure A-4. Count of activities by grades (grade 5 to 9: n=1,901; grade 10: n=1,293; repeated grades are excluded)

Appendix B

Table B-1 Completion of the calendar instrument by separate steps of the guided procedure

Step of procedure	segment/ domain: in-school	segment/ domain: non-school	
1. filled in memo	88.5	92.7	
2. filled in grades	97.7	95.0	
3. filled in activities	88.1	91.4	
4. filled in dates for activities	88.0	91.3	

Note: figures in %. n=1,901

Table B-2 Poisson regression and Incidence Rate Ratios (IRR) on fifth grade activity counts of the in-school calendar

		95% Confidence interval			
	β (SE ¹)	CI _{min}	IRR	CI _{max}	
Constant	-0.59*** (0.13)	0.43	0.55	0.71	
Retrospective intervals (Reference: 4 years)					
5 years	0.01 (0.11)	0.81	1.01	1.25	
6 years	-0.22 (0.14)	0.60	0.80	1.06	
7 years	0.13 (0.18)	0.80	1.14	1.62	
Controlled variables					
Former East Germany	0.27* (0.12)	1.02	1.31	1.67	
Female	0.09 (0.07)	0.95	1.10	1.27	
Intermediate school change	-0.26^{\dagger} (0.14)	0.59	0.77	1.01	
Migration background	0.21** (0.08)	1.07	1.24	1.43	
Lower HISEI ² quartile	0.03 (0.08)	0.87	1.03	1.21	
Upper HISEI ² quartile	0.23* (0.10)	1.03	1.26	1.53	
All-day-participant	0.62*** (0.10)	1.51	1.85	2.27	

Note: n = 1,617. McFadden's Pseudo R² = .05. Wald $\chi^2(10) = 145.27***$. AIC = 3,762.9.

¹ Adjusted standard errors for 65 school clusters. ² Highest International Socio-Economic Index of Occupational Status (HISEI) of parents

[†] p < .10. * p < .05. ** p < .01. *** p < .001.

Table B-3Poisson regression and Incidence Rate Ratios (IRR) on fifth grade activity counts of the non-school calendar

		95% Confidence interval			
	β (SE ¹)	CI _{min}	IRR	CI _{max}	
Constant	0.37*** (0.06)	1.30	1.45	1.62	
Retrospective intervals (Reference: 4 years)					
5 years	0.00 (0.06)	0.88	1.00	1.13	
6 years	0.10 (0.10)	0.91	1.11	1.35	
7 years	0.12 (0.18)	0.79	1.12	1.60	
Controlled variables					
Former East Germany	-0.19*** (0.05)	0.74	0.83	0.92	
Female	0.02 (0.05)	0.93	1.02	1.12	
Migration background	-0.12* (0.05)	0.80	0.89	0.98	
Lower HISEI quartile ²	-0.23*** (0.05)	0.72	0.80	0.87	
Upper HISEI quartile ²	0.03 (0.04)	0.95	1.03	1.11	
All-day-participant	$0.12^{\dagger} (0.07)$	0.98	1.13	1.30	

Note: n = 1,617; Wald $\chi^2(9)$ = 61.18***; McFadden's Pseudo R² =.01; AIC = 4,654.3

Adjusted standard errors for 65 school clusters. ² Highest International Socio-Economic Index of Occupational Status (HISEI) of parents

[†] p < .10. * p < .05. ** p < .01. *** p < .001.