

European Data Watch

This section offers descriptions as well as discussions of data sources that are of interest to social scientists engaged in empirical research or teaching courses that include empirical investigations performed by students. The purpose is to describe the information in the data source, to give examples of questions tackled with the data and to tell how to access the data for research and teaching. We focus on data from German speaking countries that allow international comparative research. While most of the data are at the micro level (individuals, households, or firms), more aggregate data and meta data (for regions, industries, or nations) are included as well.

The SOEP Innovation Sample (SOEP IS)

By David Richter and Jürgen Schupp

1. Introduction

The Socio-Economic Panel (SOEP) research infrastructure at the German Institute for Economic Research (DIW Berlin) offers researchers worldwide the opportunity to use the longitudinal SOEP Innovation Sample (SOEP-IS) for their research projects. Besides containing a relatively short set of core questions, the SOEP-IS longitudinal panel survey incorporates innovative content that is purely user-designed. The SOEP-IS is suited to short-term experiments, but it is particularly useful for longer-term surveys that are not suitable to the SOEP-Core survey, whether because they pose an unacceptably high risk of refusal or because the questions deal with very specific research issues. SOEP-IS was also conceived as a tool for collecting measurements that go beyond the scope of standardized questionnaire formats: for example, genetic studies on the interaction between social and genetic factors in human development, implicit association tests (IAT), or the use of mobile devices as part of experience sampling. Since 2012, the SOEP has accepted proposals for the Innovation Sample from users and assessed the submissions in an annual competitive re-

fered process to identify top-quality research questions and operationalizations.

The annual fieldwork of SOEP-IS runs from September to December of each year. SOEP-IS started in September 2011 with a newly developed core questionnaire, “SOEP Innovations”, containing about 44% of the yearly questions in SOEP-Core and integrating all questions on the household, adult household members, their biographies, and the information on children into a single instrument. Rotating modules (i.e., sets of questions that are asked only at specific intervals) were added, following the rotation scheme of SOEP-Core, along with the innovative modules from the particular SOEP-IS survey wave. SOEP-IS is therefore in many respects a shortened version of the ongoing longitudinal household panel study SOEP-Core, enriched with innovative new content and a chance to test new content for future waves of the SOEP.

2. On the Structure of the SOEP-IS

In 2014, 5,868 respondents (52 female) in 3,721 households participated in the SOEP IS (see Table 1). Some of these women and men have been answering SOEP-Core questions since 1998 as part of an extension sample to the SOEP (SOEP Sample E), while others entered SOEP-Core in 2009 (SOEP Sample I). These individuals serve as the foundation for the SOEP-IS and provide a wealth of longitudinal data and rich representation on phenotype data. Within the framework of the SOEP Innovation Sample, additional samples were recruited in 2012, 2013, and 2014 (see Table 1, Figure 1A and Figure 1B).

Table 1
**Number of households and respondents (in brackets)
of SOEP-IS (2011–2014)**

Sample	Survey Year			
	2011	2012	2013	2014
Sample E		339 (551)	310 (511)	298 (484)
Sample I	1040 (1701)	928 (1501)	864 (1407)	798 (1278)
Supplementary Sample 2012		1,010 (1644)	833 (1383)	772 (1269)
Supplementary Sample 2013			1166 (1840)	929 (1467)
Supplementary Sample 2014				924 (1370)
Households total (respondents total)	1040 (1701)	2277 (3696)	3173 (5141)	3721 (5868)

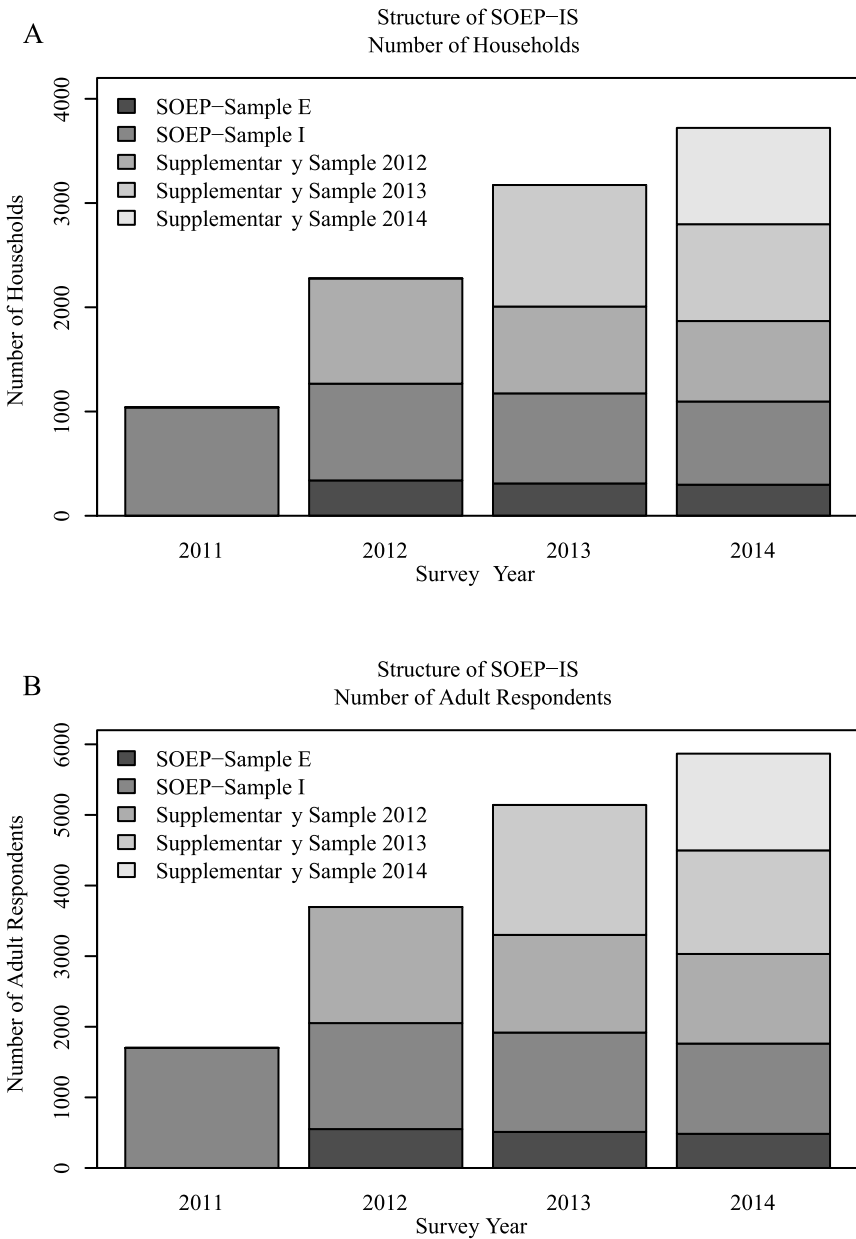


Figure 1A: Structure of SOEP-IS: Households.

1B. Structure of SOEP-IS: Respondents.

The SOEP-IS spans the whole adult age range from 17 to 96 years and is a randomly drawn representative sample of private households in Germany. The sample consists of respondents with differing levels of education (19% low education: ISCED-Scores 0, 1, and 2; 56% medium education: ISCED-Scores 3 and 4; and 25% high education: ISCED-Scores 5 and 6; ISCED-1997; OECD, 1999), work situations (36% working full-time, 18% working part-time, 30.0% retired, 16% not working), and marital statuses (26% single, 57% married, 10% divorced, 8% widowed). In addition, 28% of the respondents live in households with children aged 16 years or younger and provide yearly information on the development of their children reported by the primary caregiver, usually the mother.

3. Application Procedure

A two-step module of governance has been established in SOEP-IS to regulate the selection of topics and question modules. First, SOEP survey management conducts a basic methodological test in close cooperation with the survey institute to establish whether the size, format, and survey mode outlined in a proposal seem appropriate for implementation in the SOEP-IS. Then, experts in the respective field – in many cases longtime SOEP users or colleagues recommended by the SOEP Survey Committee – are asked to evaluate the proposals as external reviewers. On the basis of these reviews, the SOEP executive team, consisting of the director and deputy directors of the SOEP, prioritize the proposals for implementation. Criteria for the selection of proposals are: scientific quality and relevance, potential for scientific breakthrough, sound analytical framework and empirical design, appropriateness of the proposed research methodology, and finally, scientific merit of the researcher or research group. Survey questions from accepted proposals are introduced into SOEP-IS at no additional cost. However, if additional funding is needed to carry out a research proposal (e.g., for financial incentives in the case of a high response burden, or costs of conducting experiments), the proposer is expected to raise external funding. To have proposals considered for evaluation and possible implementation in any given year, they must be submitted by December 31 of the year prior to study implementation.

Complex research proposals involving additional software or extra work by our survey institute (such as interviewer training to collect biomarkers) are evaluated by a subcommittee of the SOEP Survey Committee. As we can only implement one complex research proposal per survey year, the SOEP Survey Committee evaluates and rates the relevance and significance of these proposals. If the proposed research project already has funding from the DFG or other peer-reviewed funding agencies, this step of external evaluation is skipped and the SOEP executive team decides on the timing of implementation.

Since projects funded by the DFG and similar funding bodies also require a guarantee that the project is feasible from the standpoint of survey methodology, interested researchers should contact SOEP Survey Management before submitting an application for funding.

4. Data Access

To maintain respondent confidentiality, the SOEP adheres to strict security standards in the collection and distribution of the SOEP-IS data. Like the SOEP-Core data, the SOEP-IS data are reserved exclusively for research and provided only to members of the scientific community (Wagner et al. 2007). The SOEP Research Data Center (RDC) distributes the SOEP-IS data as an independent dataset. Individuals and institutions that have signed a SOEP data distribution contract can informally request (by letter or e-mail) a supplemental contract allowing use of the SOEP-IS data. After signing the required contracts, users receive the SOEP-IS dataset by personalized encrypted download. Users can also access GEO-coded small-scale regional data that can be linked to the SOEP-IS data, but these are only available on site at the SOEP-RDC.

The data are provided in user-friendly “long” format: that is, not as wave-specific individual files but with all available survey years pooled. In addition, user-friendly generated SOEP variables are provided for the SOEP-IS data sets as well as for all of the previous years since the first SOEP-IS subsample in 1998.

The data from the innovative modules are provided exclusively to the researchers who proposed the respective ideas for an initial period of 12 months. As soon as the 12-month embargo has ended, they are released and offered for secondary analysis (Siedler et al., 2010) to the entire SOEP user community.

5. Innovative Modules Surveyed 2011 to 2014

Adaptive General Ecological Behavior Scale. The adaptive General Ecological Behavior scale (a-GEB) was designed to provide an economic assessment of a person’s propensity to engage in an ecological lifestyle. As such, it is especially useful for promoting sustainable development in science-based policy making. Its construct validity and external validity were previously established for non-adaptive versions (e.g., Kaiser/Byrka/Hartig, 2010; embargo ended 04/2015).

Anxiety and Depression. (embargo ended 04/2015).

Comparing Measures of Experiential and Evaluative Well-Being. Asking about people’s feelings throughout the day as they go through their lives allows us to get an overall sense of how a person’s life is going by measuring how they

feel on a moment-to-moment basis. Further, this allows us to examine how specific day-to-day activities affect quality of life (embargo until 04/2017).

Computer-Assisted Measurement and Coding of Educational Qualifications in Surveys (CAMCES). To overcome incompleteness in the case of long-list factual questions, a database query was incorporated in the survey procedure. To test different database interfaces, a split-ballot experiment was conducted using the example of educational attainment comparing a dynamic text field, a search tree, and a long list (embargo until 04/2017).

Confusion, Hubbub, and Order Scale (CHAOS). Children growing up in chaotic homes, i.e., homes characterized by noise, crowding, and a lack of structure, are more likely to develop behavioral and other problems. The CHAOS module (e.g., Matheny/Wachs/Ludwig/Phillips, 1995) in the SOEP-IS allows the replication and extension of previous results on the association of home chaos with child development outcomes (embargo until 04/2017).

Conspiracy Mentality Questionnaire (CMQ). The five-item Conspiracy Mentality Questionnaire (CMQ; Bruder et al., 2012) is designed to efficiently assess the general propensity to believe in conspiracy theories. The CMQ examines meaningful associations with personality measures, generalized political attitudes (e.g., right-wing authoritarianism), individual differences (e.g., perceived socio-political control). Measurement equivalence across three language versions has already been established (embargo until 04/2016).

Control Strivings. (embargo ended 04/2015).

Cross-Cultural Study of Happiness. The goal of this project is to explore how culture-specific meanings of happiness (cf. Hitokoto/Uchida, 2015) give rise to differences both across cultures (Germany and Japanese) and within cultures (based on socio-economic status).

Day Reconstruction Method. Experiential measures of subjective well-being assess affective reactions over time, either by sampling experiences as they happen or by retrospectively recreating experiences soon after they occur. In this module, respondents reconstruct a full day, describing what they did, who they were with, and how they felt throughout the day (cf. Kahneman et al., 2004; embargo ended 04/2015).

Decisions from Description and Experience. The SOEP-IS Risk Module consists of two incentive-compatible behavioral risk-taking tasks involving described and experienced risk (cf. Hertwig/Barron/Weber/Erev, 2004) and extends the SOEP by providing an assessment of individual differences that may predict real-world outcomes such as employment, financial, and health decisions that are partly guided by individuals' risk tendencies (embargo until 04/2017).

Determinants of Attitudes to Income Redistribution. We introduce two questions measuring attitudes towards income redistribution and two questions on the beliefs of the causes of low and high income (cf. Fong, 2001). Together

with existing SOEP questions, these make it possible to estimate the relative importance of self-interest, fairness considerations, and various other factors in explaining attitudes towards redistribution (embargo until 04/2017).

Determinants of Not Wanting to Know. Would you like to know the exact day of your death? The goal of this SOEP-IS module is to examine the determinants of people's knowledge preferences, and more specifically, the reasons behind the puzzling and sometimes strong desire for explicitly not wanting to know (cost-free) information (cf. Ehrlich/Irwin, 2005; embargo until 04/2017).

Expected Financial Market Earnings. The module is concerned with the determinants of stock market participation and the (perceived) risks people expose themselves to in their investments. At the module's heart is an elicitation and an experimental manipulation of respondents' beliefs about the return on the German stock market (embargo ended 04/2015).

Fear of Dementia. (embargo ended 04/2015).

Future Life Events. Many life events do not happen out of the blue but can be anticipated beforehand (e.g., Luhmann/Hofmann/Eid/Lucas, 2012). To investigate these kinds of anticipation effects, participants were asked to indicate how likely it was that different events would occur in the next 12 months. Events from various life domains (e.g., family, work, mobility) were included (embargo until 04/2017).

Inattentional Blindness. Inattentional blindness (IB) characterizes the failure to notice salient events while one is paying attention to something else (Chabris/Simons, 2010). IB has been widely replicated but has never been measured in a nationally representative sample. We measure IB's prevalence and its association with ADHD and autism, and with socio-demographic background (embargo until 04/2017).

Internalized Gender Stereotypes Vary Across Socioeconomic Indicators. We propose that one important cause of gender inequality is the presence of gender stereotypes in society. We describe two approaches to measure gender stereotypes: an explicit questionnaire based on rating scales and an Implicit Association Test. Findings indicate that gender stereotypes are related to socioeconomic and social variables (Dietrich et al., 2013; embargo ended 04/2014).

Job Preferences and Willingness to Accept Job Offers. The factorial survey module on job preferences and willingness to accept job offers provides multi-dimensional, experimental measurements of job preferences and demanded compensations for (un-)favorable working conditions (e.g., Auspurg/Abraham, 2007). It enables deeper insights into the impact of gender, household structures, and job conditions on inequalities in the labor market (embargo until 04/2016).

Job Task Survey. (embargo until 04/2016).

Just Sustainable Development Based on the Capability Approach (GeNECA). The GeNECA data collected in 2012 covers a broad range of issues: quality of

life, capabilities, expectations about the preservation of living conditions for subsequent generations, attitudes towards justice and the environment, actors for sustainable development, regional currency, and sustainable consumption with regard to car use and organic food (cf. Leßmann/Masson, 2015; embargo ended 04/2015).

Justice Sensitivity. Individuals differ systematically in how readily they perceive situations to be unjust and how strongly they react to subjective injustice – cognitively, emotionally, and behaviorally (e.g., Schmitt, 1996). Justice sensitivity from the perspectives of a victim, observer, beneficiary, and perpetrator can be measured reliably with two items per perspective (embargo ended 04/2014).

Lottery Play: Expenditure, Frequency, and Explanatory Variables. This SOEP-IS module measures lottery play (expenditure, frequency, individual vs. social play) plus three possible explanatory variables: A “daydreaming” scale, measuring individual tendencies to indulge in fantasies about positive future states; an “alienation” scale, measuring senselessness and dissatisfaction with daily (work) routines, and a “work ethic” scale, measuring attitudes toward work and effort (cf. Beckert/Lutter, 2013; embargo until 04/2017).

Narcissistic Admiration and Rivalry Questionnaire (NARQ-S). Our SOEP-IS module is concerned with the investigation of narcissism, its development and interpersonal, intrapersonal, and institutional consequences. Inclusion of a theoretically sound, reliable, and efficient measure of narcissism (NARQ-S; Back et al., 2013) enables researchers to investigate a variety of focal research questions with wide-ranging implications on individual, social, and societal levels (embargo until 04/2016).

Regional identification. (embargo until 04/2016).

Self-Evaluation and Overconfidence in Different Life Domains. We measure overconfidence in various domains of life and over several waves. These measures, in combination with the rich panel structure of the SOEP, allow us to examine the prevalence of overconfidence in different domains and study how overconfidence responds to life events (cf. Arni/Goette/Ziebarth, 2013). We also investigate how overconfidence affects labor-market outcomes (embargo until 04/2017).

Self-Regulated Personality Development. We are interested in identifying why and through what processes personality changes throughout adulthood by analyzing whether individuals actively change their own personality traits (cf. Hennecke/Bleidorn/Denissen/Wood, 2014). Specifically, we measure (a) the subjective desirability of personality change, (b) the subjective feasibility of personality change, and (c) actual personality change (embargo until 04/2018).

Separating Systematic Measurement Error Components Using MTMM in Longitudinal Studies. In this project we develop a new research design that

enables us to measure and disentangle multiple types of error: method, social desirability, acquiescence (tendency to select first category), and random error. Furthermore, we investigate how these errors change over time and how they compare cross-culturally (embargo until 04/2017).

Sleep Characteristics. The instrument compiled from several validated questionnaires (e.g., Stang et al., 2012) allows the assessment of the following characteristics: sleep quality, sleep disturbances at night, chronotype, sleep onset latency, daytime sleepiness, sleep efficiency, sleep duration, sleep-related breathing disorders, sleep hygiene, sleep environment, daytime naps, reasons for daytime naps, and start of regular naps.

Socio-Economic Effects of Physical Activity. To overcome the lack of reliable measures of physical activity, this SOEP-IS module consists of detailed questions on the type of physical activity, the frequency and intensity of the activity, the organizational context in which the physical activity is practiced, and the possible persistence of activity patterns across the life course (cf. Lechner, 2009; embargo until 04/2016).

The Big Two Psychological Content Dimensions: Agency and Communion. Agency (competent, clever, influential, able) and communion (warm, helpful, empathic, caring) constitute the fundamental content dimensions of personality, group perception, and values (cf. Gebauer/Sedikides/Lüdtke/Neberich, 2014). Accordingly, this module includes self-reports of agency and communion as personality traits, personal values, and societal values. It also includes an implicit association test, measuring agency vs. communion via reaction times (embargo ended 04/2015).

6. Outlook: SOEP-IS as Basis for SOEP Related Studies

Household panels in general, and the German Socio-Economic Panel (SOEP) in particular, are useful as reference data for researchers whose primary datasets do not represent the full diversity of the population of interest (e.g., datasets obtained from clinical trials, intervention studies for special target groups, laboratory and behavioral experiments, and cohort studies; Siedler et al., 2009). The yearly core questionnaire of SOEP-IS is ideally suited as a basic instrument for SOEP-related studies. Questions on households, respondents, and children are integrated into a single instrument, and the focus on the key SOEP questions leaves enough survey time for the study-specific questions and experiments.

References

- Arni, P./Goette, L./Ziebarth, N. R.* (2013): Healthy Overconfidence? The impact of overconfidence on health-related behaviors. University of Lausanne.
- Auspurg, K./Abraham, M.* (2007): Die Umzugsentscheidung von Paaren als Verhandlungsproblem. Eine quasiexperimentelle Überprüfung des Bargaining-Modells. [Household moves as a bargaining problem in family relationships: A quasi-experimental test of the family bargaining model], *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59, 271–293.
- Back, M. D./Küfner, A. C. P./Dufner, M./Gerlach, T. M./Rauthmann, J. F./Denissen, J. J. A.* (2013): Narcissistic admiration and rivalry: Disentangling the bright and dark sides of narcissism, *Journal of Personality and Social Psychology* 105, 1013–1037.
- Beckert, J./Lutter, M.* (2013): Why the poor play the lottery: Sociological approaches to explaining class-based lottery play. *Sociology*, 47, 1152–1170.
- Bruder, M./Haffke, P./Neave, N./Nouripannah, N./Imhoff, R.* (2013): Measuring individual differences in generic beliefs in conspiracy theories across cultures: The Conspiracy Mentality Questionnaire (CMQ), *Frontiers in Psychology* 4, 225.
- Chabris, C. F./Simons, D. J.* (2010): *The invisible gorilla, and other ways our intuitions deceive us.* New York: Crown.
- Dietrich, J./Schnabel, K./Ortner, T. M./Eagly, A./Retamero-Garcia, R./Kröger, L./Holst, E.* (2013): Internalized gender stereotypes vary across socioeconomic indicators, SOEPpaper No. 558.
- Ehrich, K. R./Irwin, J. R.* (2005): Willful ignorance in the request of product attribute information, *Journal of Marketing Research* 42, 266–277.
- Fong, C.* (2001): Social preferences, self-interest, and the demand for redistribution, *Journal of Public Economics* 82, 225–246.
- Gebauer, J. E./Sedikides, C./Lüdtke, O./Neberich, W.* (2014): Agency-communion and interest in prosocial behavior: Social motives for assimilation and contrast explain sociocultural inconsistencies, *Journal of Personality* 82, 452–466.
- Hennecke, M./Bleidorn, W./Denissen, J. J. A./Wood, D.* (2014): A three-part framework for self-regulated personality development across adulthood, *European Journal of Personality* 28, 289–299.
- Hertwig, R./Barron, G./Weber, E. U./Erev, I.* (2004): Decisions from experience and the effect of rare events in risky choice, *Psychological Science* 15, 534–539.
- Hitokoto, H./Uchida, Y.* (2015): Interdependent happiness: Theoretical importance and measurement validity, *Journal of Happiness Studies* 16, 211–239.
- Kahneman, D./Krueger, A. B./Schkade, D. A./Schwarz, N./Stone, A. A.* (2004): A survey method for characterizing daily life experience: The day reconstruction method, *Science* 306, 1776–1780.
- Kaiser, F. G./Byrka, K./Hartig, T.* (2010): Reviving Campbell's paradigm for attitude research, *Personality and Social Psychology Review* 14, 351–367.

- Kessler, E.-M./Bowen, C. E./Baer, M./Frölich, L./Wahl, H.-W. (2012): Dementia Worries: A psychological examination of an unexplored phenomenon, *European Journal of Ageing*, 9 (4), 275–284. doi: 10.1007/s10433-012-0242-8.
- Kohlmann, S./Gierk, B./Löwe, B. (2014): PHQ-4. Patient Health Questionnaire-4, in: C. J. Kemper/E. Brähler/M. Zenger (eds.), *Psychologische und sozialwissenschaftliche Kurzskalen. Standardisierte Erhebungsinstrumente für Wissenschaft und Praxis*, 242–244, Berlin.
- Lechner, M. (2009): Long-run labour market and health effects of individual sports activities, *Journal of Health Economics* 28, 839–854.
- Leßmann, O./Masson, T. (2015): Sustainable consumption in capability perspective: Operationalization and empirical illustration, *Journal of Behavioral and Experimental Economics* 57, 64–72.
- Luhmann, M./Hofmann, W./Eid, M./Lucas, R. E. (2012): Subjective well-being and adaptation to life events: A meta-analysis, *Journal of Personality and Social Psychology* 102, 592–615.
- Matheny, A. P./Wachs, T. D./Ludwig, J. L./Phillips, K. (1995): Bringing order out of chaos: Psycho-metric characteristics of the confusion, hubbub, and order scale, *Journal of Applied Developmental Psychology* 16, 429–444.
- OECD (1999): *Classifying educational programmes: Manual for ISCED-97 implementation in OECD countries*. Paris: OECD.
- Schmitt, M. (1996): Individual differences in Sensitivity to Befallen Injustice (SBI), *Personality and Individual Differences* 21, 3–20.
- Schubach, E./Zimmermann, J./Noack, P./Neyer, F. J. (in press): Me, Myself, and Mobility: The Relevance of Region for Young Adults' Identity Development. *European Journal of Personality*.
- Siedler, T./Schupp, J./Spiess, K. C./Wagner, G. G. (2009): The German Socio-Economic Panel (SOEP) as reference data set, *Schmollers Jahrbuch* 129, 367–374.
- Siedler, T./Schupp, J./Wagner, G. G. (2011): Innovative methods within the context of secondary data: Examples from household panel surveys, in: K. H. Trzesniewski/M. B. Donnellan/R. E. Lucas (eds.), *Secondary data analysis: An introduction for psychologists*, Washington, DC, 103–118.
- Stang, A./Dragano, N./Moebus, S./Möhlenkamp, S./Schmermund, A./Kälsch, H. (2012): Midday Naps and the Risk of Coronary Artery Disease: Results of the Heinz Nixdorf Recall Study, *Sleep* 35, 1705–1712.
- Wagner, G. G./Frick, J. R./Schupp, J. (2007): The German Socio-Economic Panel Study (SOEP) – Scope, evolution and Enhancement, *Schmollers Jahrbuch* 127, 139–169.