Wagner, Gert G.; Bruemmer, Martin; Glemser, Axel; Rohrer, Julia; Schupp, Jürgen

Working Paper

Dimensions of Quality of Life in Germany: Measured by Plain Text Responses in a Representative Survey (SOEP)

IZA Discussion Papers, No. 10521

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Wagner, Gert G.; Bruemmer, Martin; Glemser, Axel; Rohrer, Julia; Schupp, Jürgen (2017) : Dimensions of Quality of Life in Germany: Measured by Plain Text Responses in a Representative Survey (SOEP), IZA Discussion Papers, No. 10521, Institute of Labor Economics (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/161144

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10521

Dimensions of Quality of Life in Germany: Measured by Plain Text Responses in a Representative Survey (SOEP)

Gert G. Wagner
Martin Bruemmer
Axel Glemser
Julia Rohrer
Jürgen Schupp

JANUARY 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10521

Dimensions of Quality of Life in Germany: Measured by Plain Text Responses in a Representative Survey (SOEP)

Gert G. Wagner
DIW Berlin, MPI for Human Development and Berlin University of Technology (TUB)

Martin Bruemmer
University of Leipzig

Axel Glemser
Kantar Public

Julia Rohrer
MPI for Human Development and Freie Universität Berlin (FU)

Jürgen Schupp
DIW Berlin, Freie Universität Berlin (FU) and IZA

JANUARY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Dimensions of Quality of Life in Germany: Measured by Plain Text Responses in a Representative Survey (SOEP)

This paper demonstrates how quality of life can be measured by plain text in a representative survey, the German Socio Economic Panel Study (SOEP). Furthermore, the paper shows that problems that are difficult to monitor, especially problems like the state of the European Union, long-term climate change but also the national debt or problems with the quality of consumer goods (like food) and services (like medical treatment), are not issues of particular importance to the majority of people. Developments and risks that are difficult to monitor and only have long-term effects should be left primarily to the discourse conducted by experts and the politically-minded "elites", the avant garde. And in representative democracies it is ultimately the parliamentarians who must decide. Parliamentarians are likely able to make somewhat better decisions using modern representative surveys and national dialogues than they would be without these instruments of civic participation. Nevertheless, improved civic participation cannot replace parliaments.

JEL Classification: B41, C81, C83, C88, D69, I31, Z13

Keywords: quality of life, dimensions, open-ended questions, automated text analysis, German Socio Economic Panel Study, SOEP

Corresponding author:
Gert G. Wagner
DIW Berlin
Mohrenstr. 58
10117 Berlin
Germany
E-mail: gwagner@diw.de

The desire of politicians, scientists and journalists to illustrate the realities of life in a differentiated manner, beyond a single measure such as gross domestic product (GDP), is easily understandable. However, it is not an easy task to determine the (social) indicators useful for such a differentiated picture of reality.[1]

It is self-evident that specialist knowledge and the professional judgement of experts are useful in choosing statistical measurements and implementing these concepts as part of statistical surveys. It would make little sense to forego the expertise of economists, sociologists and statisticians when operationalising the concept of "income", for example. Depending on the concept to be measured, the specialist disciplines involved in that concept should also be brought in along with statisticians. For example, statistical measurements of "life satisfaction" or "pollution" benefit from input from the fields of psychology and environmental science. However, whether expert judgement is just as useful or even indispensable when it comes to the selection of spheres of life and other areas which are described by using indicators is by no means clear or even apparent.

After all, why should experts choose and decide which dimensions of life, society and nature are important and which should be measured using indicators, and not the country's citizens themselves? In a democracy, it is citizens as represented by their elected parliamentarians who decide what is important, although experts might not always be particularly happy about this circumstance. Even though the "will of the people" might not always be based on the full body of facts which are, in theory, produced and managed by and available to experts, putting the needs and objectives of experts above those of citizens is not a solution: Experts by no means represent the plurality of needs and objectives in a society.

This essay will discuss and show how plain text responses in a large survey of the population (SOEP) can be used to determine the dimensions of quality of life. Ultimately, the analysis of those texts (words) will underline an old insight of political

[1] A shorter version of this paper is published – in German language – as a chapter in Gert G. Wagner, On the Role of Population-based Surveys and Citizen Dialogues in Selecting Social Indicators, https://buergerdialog.gut-leben-in-deutschland.de/SharedDocs/Downloads/DE/LB/Diskussionsbeitraege-des-wissenschaftlichen-Beirats.pdf?__blob=publicationFile.

scientist: parliamentarians serve an indispensable role at the junction between expert knowledge and the will of the people.

1 German Socio Economic Panel Study (SOEP)

The German Socio Economic panel Study (SOEP) is is a wide-ranging representative longitudinal study of private households in Germany (multi cohort approach). Every year, more than 10,000 households with about 30,000 individuals living in those households are surveyed by the fieldwork organization Kantar Public (Infratest Sozialforschung). The SOEP was started in 1984 (Wagner et al. 2007).

The data provide information on all household members, consisting of Germans living in the old and new federal states of Germany, foreigners, and recent immigrants to Germany. Some of the many topics include household composition, occupational biographies, employment, earnings, health, political attitudes, worries and satisfaction in general and in different domains of life.

The improved possibilities for automated text analysis in recent years make it feasible to pose questions in surveys that are not answered in a standardised, quantitative manner, but instead in natural language. Thus, it was methodologically justified to ask the two "key questions" of the national dialogue "Living well **in** Germany **–** what is important to us" in the individual interviews of the SOEP in 2015.[2] These questions were: "What do you think is important in life?" and "What, in your opinion, is characteristic of the quality of life in Germany?".

The national dialogue gave a lot of people the chance to make their voices heard and to contribute their ideas – either in person at about 200 events, very similar to town-hall meetings, held all over the country, online or by using a simple postcard.[3]

[2] In fact due to time restrictions of the preparation of the questionnaires thse questions were administered only for the modes CAPI and CAWI. These respondents are a virtually undistorted random selection of all respondents since the decision regarding the CAPI and CAWI survey methods was Infratest's and not respondents'.

[3] See https://buergerdialog.gut-leben-in-deutschland.de/EN/Home/home_node.html. "With the dialog the German government aimed to identify yardsticks that can be used to pinpoint the many different facets of the quality of life. This will allow the government to gear its action in future to what is important to the people in Germany. The government's strategy builds on the Chancellor's National Dialogue on Germany's Future and the Study Commission on

However, those events and channels did not guarantee the collection of answers of a representative sample of people in Germany. The SOEP took the chance and complemented the national dialogue by posing the two key questions in the 2015 SOEP survey.

All SOEP respondents participating in the survey using computers (some 71 per cent of all respondents) were asked the questions at the very end of the questionnaire. Overall, 19,352 respondents took a computer-assisted personal interview (CAPI) or computer-assisted web interview (CAWI). Additionally, a sub-sample of some 60 randomly selected SOEP respondents were able to speak with Chancellor Merkel directly at a national dialogue event June 1st 2015 in Berlin. [4] The respondents discussed their problems and wishes with the Chancellor (Figure 1). The discussion of this representative crowd was for all intents and purposes unspectacular. This was not really a surprise because the life of most people is unspectacular.

< Figure 1 about here >

As was previously already known from the "other worries", which can be expressed by respondents of SOEP since 1984, text responses as part of a survey predominantly provide answers in the form of "keywords" and do not in fact provide longer texts that place responses in any greater context (of reasoning). With an average of six words, the full answers to the two "key questions" in SOEP are comparable in length with the headings formulated by citizens in their online responses in the national dialogue. The answers collected by the national dialogue's town-hall meetings, which explicitly included discursive elements generated much more in-depth texts than the representative SOEP survey.

Growth, Wellbeing and Quality of Life. It also picked up on ideas advanced by the scientific and research community."

[4] https://buergerdialog.gut-leben-in-deutschland.de/SharedDocs/Blog/DE/05-Mai/2015-05-29-vorbericht-bkin.html?nn=1328626.

2 Automated text analysis[5]

Open-ended questions such as “Would you like to add anything?”, “Is there anything else you would like to tell us?”, “Please tell us anything you think is important” are commonly used as complements in surveys that otherwise rely heavily on closed-ended questions. However, to this day – decades after the collection of such textual answers began – routines have yet to be established for analyzing the unstandardized textual answers – so-called free texts – and for integrating them into quantitative social science research (e. g. Rohrer et al. 2017).

Open-ended questions on surveys typically generate a large number of short texts, in contrast to the small numbers of long and comprehensive texts that are routinely analyzed in Qualitative Content Analysis (QCA).[6] Respondents of large surveys often provide only one or a handful of words in their answers because of (1) the narrow phrasing of the questions, (2) a lack of motivation to answer exhaustively, (3) space limitations on the questionnaire, or (4) time pressure in the interview situation due to the interest of the interviewer to complete the interview quickly.

The responses to the SOEP survey were largely short sentence fragments, key points separated by commas or phrases with just a few number of words. Analyses incorporating syntactic characteristics are therefore not very promising. In addition, standard software trained using machine learning, such as software that detects parts of speech, only work at fairly low levels of precisions because texts differ greatly from the training material used. In order to facilitate computer-aided quantitative analysis, the texts were initially prepared in several stages. The aim of this process was to reduce words with similar meanings to identical strings to reduce the inherent heterogeneity of the language. This involved taking into account the peculiarities of the German language, in particular the high degree of inflection. The

[5] This chapter is a much shorter version of the respective chapter in Rohrer et al. (2017).

[6] QCA offers a range of techniques to approach the content of a text on different levels, from the gist of the text to subtle references that can be understood only in the broader context of current events and public discourses. See Mayring (2000) for a brief overview and examples of its application. Content analysis itself covers a wide range of different strategies and offers promising new possibilities (e. g. Nacos et al. 2009). For example, researchers can derive categories of interest from the data itself, from theory, or from prior research. Researchers are also able to focus on the keywords that are identified from the underlying context (Hsieh and Shannon 2005).

word processing aimed at presenting words with the same meaning as identical strings to be able to carry out thematic analyses.

As a result, a customised processing pipeline comprising the following components can be used:

- Starting with tokenising texts into individual words and removing punctuation marks and punctuation;
- using a stop word list to remove stop words like articles, pronouns and conjunctions since they have no semantic content for analytical purposes;
- automatically correcting grammar to correct faults;
- replacing frequently used abbreviations using a dictionary customised for the data;
- converting all characters to lower case to align individual spellings and increase the integrity of the reduction in the next step;
- stemming to reduce formulations of equal importance to identical strings; and
- expanding reduced formulations to the corresponding dictionary formulation to make the results more readable.

More in-depth analysis is largely dismissable, especially for rating purposes (valence of terms). Nevertheless, valence (positive, neutral, negative) was taken into account when counterintuitive terms like “war” were used in response to questions about quality of life in Germany. These kinds of counterintuitive terms were supplemented in the analysis by their rating (for example, “no_war”).

3 Empirical results

The same questions which were asked to non-representative samples within the governmental project “Living well in Germany” were asked in the representative

sample of SOEP in 2015[7]: “What do you think is important in life?” and “What, in your opinion, is characteristic of the quality of life in Germany?”

A non-marginal proportion of SOEP respondents gave no text responses when asked the two key questions. The 6.5 per cent to the question on the importance of different spheres of life may be treated as regular “no response”. In contrast, it could be the case that 14.4 per cent who gave no answer to the question of quality of life in Germany wanted to express their unhappiness with live in Germany.

3.1 Selection effects

Which variables determined whether a respondent answered or ignored the question? We ran multilevel binary logistic regressions predicting the binary outcome answer/no answer to investigate selection effects.

All analyses of selection effects were run on the sample of 16,440 respondents who were asked the two key questions and furthermore provided answers to all items included as predictors of response behavior, see Table 1 and 2 for detailed results.

< Tables 1 and 2 about here >

Overall, there are signs that several selection effects might affect who does and who does not answer the open-ended question. Not all of these effects seem to affect the two questions to the same extent. For example, male respondents were less likely to answer the question regarding what is important in life (Table 1), but no comparable trend is visible for the question regarding quality of life in Germany. However, both questions showed an age trend: Older respondents were more likely to answer the open-ended question. Unemployed individuals seemed more likely to answer the question regarding what is important in life, but this effect again was not visible for

[7] The “Living well in Germany” project goes into far greater depth than what was possible in the context of the SOEP survey. Thus, the SOEP representative plain language survey was only considered a supplement and not a replacement for the national dialogues and written surveys of the governmental project.

the question regarding quality of life in Germany. However, for both questions, respondents with at least one parent whose native language was not German (a proxy for migration history) had much higher odds to provide an answer.

Regarding education, effects again differ for the two answers: Respondents with a higher level of education were more likely to answer the question regarding quality of life in Germany, but this effect was not visible on the question regarding what is important of life.

A positive relationship between level of education and answering open-ended questions has been observed before, both in the SOEP (Rohrer et al., 2017) and in other studies (Garcia, Evans, & Reshaw, 2004; Rich, Chojenta, & Loxton, 2013). There might be two explanations for the fact that this effect was not found for the question regarding what is important in life: First, respondents might have been motivated to answer this question regardless of level of education, but then when being asked the second question, which is admittedly somewhat similar in content, respondents with lower levels of education might have been less motivated to answer the question or might have struggled to perceive the exact difference between the questions. Second, while the first question regarding what is important in life was phrased in everyday language and asks for the individual priorities of the respondent, the question regarding quality of life in Germany is far more abstract and the phrasing alone might have given certain respondents the impression that they cannot provide an adequate answer.[8]

Regarding the emotional state of respondents, both question indicate that happier respondents were more likely to provide an answer, whereas individuals reporting sad feelings were less likely to answer. Thus, the free text responses might be slightly biased towards a more positive perspective.

Lastly, respondents who reported following a political party were significantly more likely to answer the question regarding quality of life in Germany, and this is in particular true for followers of the new right-wing party AfD.

[8] To avoid non-response behavior, it might thus be advisable to ensure that open-ended questions do not overlap in content and are phrased clear and in everyday language, avoiding academic language to not give respondents that their knowledge is being tested.

Taken together, finding indicates that even if a representative sample is asked a particular question, non-response conditional on characteristics of the interviewed person might still lead to a sample of text responses in which more vocal groups are overrepresented.

3.2 Keywords

The differences between East and West Germany are minimal. Thus the following "word clouds" demonstrate the outcome for Germany as a whole.

< Figures 2 and 3 about here >

In terms of the most important things in life, Figure 2 shows that terms relating to health and family were mentioned by far most frequently. Yet a societal term comes in at fourth place: "peace" (which was mentioned together with "social" in around one-third of mentions and in the sense of international peace in another third). Also in the top 25 are 'material' terms such as "work", "money" and "financial", and 'immaterial' terms like "satisfaction" and "harmony" are mentioned. What is striking is that terms associated with conservation and environmental protection are not mentioned very often. The term "environment" comes in 43rd place. Missing values (in the figure labelled as "no_statement") do not play a significant role.

In terms of quality of life in Germany (Figure 3), the term "secure" is the top ranked by far, followed by "social" (often co-occuring as "social security"). Missing values (in the figure labelled as "no_statement") play a significant role. Somewhat surprising may be that the terms "peace" and "freedom" (as well as "free") take third and fourth place. These top rankings are plausible considering peace was also one of the most significant terms for what is important in life. "Work" took sixth place (which in turn corresponds with the high importance of work). Words containing the string "work" (e.g., job, job opportunities [which all include the word *"Arbeit"* in German]) took third place. Once again, the term "environment" is not amongst the top ranked terms in evaluating quality of life; it was ranked 36th. The issue of the national debt was not ranked very high in either the question on importance or quality of life.

Most surprising amongst the responses regarding what was important to people personally and what made up quality of life was that environmental protection and conservation were only rarely mentioned, and innovation as a driver of economic growth and quality of life were practically not mentioned at all. This may come down to the structure of these processes. The vast majority of people who eventually enjoy the end products of innovation (for example, internet users) are not involved in the often tedious process of generating them. In terms of nature and the environment, it should be noted that the state of the environment in Germany has improved remarkedly since reunification, which is especially visible in East Germany in contrast to the GDR era. Furthermore, long-term problems like high CO_2 emissions are not noticeable in everyday life. In this respect, it is no surprise that the terms “innovation” and “environment” are barely mentioned. This also applies to “national debt”, since the costs associated with it are not directly noticeable (and the extent to which a low national debt is useful is an especially controversial topic, even amongst experts).

Unlike the problems with the environment that are not immediately visible and that might only appear threatening from a long-term perspective (such as gradually rising sea levels), war is happening all the time somewhere in the world. Even though Germany has lived in (external) peace for decades, war and the associated horrors are visible every day in the media, and it is now happening quite close to Western Europe in what was formerly the Soviet Union. Thus, war is part of people’s everyday lives and peace is valued accordingly.

In terms of the “Living well in Germany” project overall, it is striking that the terms collected by the SOEP are so similar in their ranking and importance to those terms found in the online dialogues/postcards/national dialogue events.

4 Conclusions

The empirical results presented in this paper make clear that it is entirely possible nowadays to representatively determine the “will of the people” by open-ended questions in a survey (here: SOEP). However, the informative value of the “Living

well in Germany" project went into far greater depth than was possible as part of the SOEP's plain language questions. This means that a representative plain language survey could in no way have replaced the national dialogues (*Bürgerdialoge*). As was previously already known from similar surveys, a representative survey of text responses as part of a conventional survey predominantly only provides answers in the form of "keywords" and does not in fact provide longer texts that place responses in any greater context (of reasoning). The text responses given in answer to the two questions asked in SOEP are comparable in length (an average of six words) with the headings formulated by citizens in their responses online. The national dialogues in particular, which explicitly included discursive elements, generated much more in-depth texts than could have been (and was) the case in the representative SOEP survey.

One detail of the government's project makes especially clear the difference between the cognitive value of representative surveys and that of individual national dialogue events with interest groups: The German chancellor's national dialogue event with 60 randomly selected SOEP respondents which formed a representative crowd was for all intents and purposes unspectacular (Figure 1). It did not reveal any surprising findings and has presumably not changed anything politically. The chancellor's next national dialogue event with pupils, however, led to an exchange between the chancellor and a young refugee, a girl named Reem Sawihl.[9] This not only drew attention to the story of the pupil from a Palestinian family living in Rostock, but the overall issue of "fleeing to Germany" advanced further into the public eye and likely amplified the existing positive feelings many already towards refugees.

All of the methods used for the project "Living well in Germany" make clear that problems that are difficult to monitor, especially problems like long-term climate change but also the national debt or problems with the quality of consumer goods (like food) and services (like medical treatment), are not issues of particular importance to the majority of people.

[9] See http://www.faz.net/aktuell/gesellschaft/merkel-trifft-reem-ein-zweites-mal-14339741.html and https://www.youtube.com/watch?v=iWPZuZU5t44.

For example, the results indicate that although the environment is discussed extensively in the political sphere, it does not play a major role in terms of its importance to life in Germany because readily apparent environmental problems have already been overcome (East Germany) and future problems are not particularly noticeable in everyday life. Even those who do acknowledge global warming do little for climate protection (Hornsey et al. 2016). This means that the indirect “will of the people” for issues or problems that are important in the long term but are at the same time less noticeable at present cannot be politically decisive.

In other words, developments and risks that are difficult to monitor and only have long-term effects should be left primarily to the discourse conducted by experts and the politically-minded “elites”, the avant garde. And in a parliamentary democracy it is ultimately the parliamentarians who must decide. Parliamentarians are likely able to make somewhat better decisions using modern representative surveys and national dialogues than they would be without these instruments of civic participation. Nevertheless, improved civic participation cannot replace parliaments.

References

Garcia J, Evans J, Reshaw M. (2004), "Is There Anything Else You Would Like to Tell Us" – Methodological Issues in the Use of Free-Text Comments from Postal Surveys. *Quality & Quantity*. 38(2):113–25. doi: 10.1023/B:QUQU.0000019394.78970.df.

Hsieh H, Shannon SE. (2005), Three approaches to qualitative content analysis. *Qualitative Health Research*. 15(9):1277–88. doi: 10.1177/1049732305276687.

Hornsey, Matthew J., Emily A. Harris, Paul G. Bain and Kelly S. Fielding (2016), Meta-analyses of the determinants and outcomes of belief in climate change, in: *Nature Climate Change*, doi:10.1038/nclimate2943.

Mayring P. Qualitative Content Analysis (2000), Forum: Qualitative Social Research. 1(2). Available from: http://www.qualitative-research.net/index.php/fqs/article/view/1089.

Minozzi, William, Michael A. Neblo, Kevin M. Esterling and David M. J. Lazer (2015), Field experiment evidence of substantive, attributional, and behavioral persuasion by members of Congress in online town halls, in: *PNAS.* 112(13), 3937-3942.

Nacos, Brigitte, Robert Y. Shapiro, John T. Young, David P. Fan, Thorsten Kjellstrand & Craig McCaa (2009): Comparing Human Coding and a Computer-Assisted Method. In: Kippendorff, Klaus and Mary Angela Bock (Eds.), *The Content Analysis Reader*, Thousand Oaks: Sage, 243-252.

Rich JL, Chojenta C, Loxton D. Quality (2013), Rigour and usefulness of free-text comments collected by a large population based longitudinal study - ALSWH. *PLoS one*. 8(7):e68832. doi: 10.1371/journal.pone.0068832.

Rohrer, J. M., Bruemmer, M., Schmukle, S. C., Goebel, J., & Wagner, G. (2017). "What Else Are You Worried About?" – Integrating Textual Responses into Quantitative Social Science Research. http://doi.org/10.17605/OSF.IO/KSG7D.

Wagner, Gert G., Joachim R. Frick and Jürgen Schupp (2007), The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancement. *Schmollers Jahrbuch*, 127 (1), 139-169.

Table 1

Results of binary logistic multilevel regressions predicting responses to the open-ended question regarding what is important in life, including 16,440 individuals.

Trait (unit of change)	Odds Ratio	*p*
Gender – female		
Male	0.80	.002
Age (years)	1.01	< .001
Employment status – full time		
Part time	1.02	.858
Unemployed	1.29	.003
Other	0.96	.733
Native language of parents – German		
Other	1.64	< .001
Education (years)	1.00	.581
Feeling happy (*SD*)	1.07	.039
Feeling anxious (*SD*)	0.93	.039
Feeling angry (*SD*)	1.01	.688
Feeling sad (*SD*)	0.90	.007
Party – none		
SPD	1.13	.280
CDU/CSU	1.14	.193
Grüne	0.89	.390
FDP	0.85	.600
AfD	2.39	.057
Other	0.89	.434

Source: SOEP 2015 (.v32); own calculations

Table 2

Results of binary logistic multilevel regressions predicting responses to the open-ended question regarding life quality in Germany, including 16,440 individuals.

Trait (unit of change)	Odds Ratio	*p*
Gender – female		
Male	1.02	.721
Age (years)	1.01	< .001
Employment status – full time		
Part time	0.94	.377
Unemployed	0.99	.929
Other	0.89	.126
Native language of parents – German		
Other	1.31	< .001
Education (years)	1.04	< .001
Feeling happy (*SD*)	1.11	< .001
Feeling anxious (*SD*)	0.98	.366
Feeling angry (*SD*)	1.04	.104
Feeling sad (*SD*)	0.95	.053
Party – none		
SPD	1.63	< .001
CDU/CSU	1.61	< .001
Grüne	1.57	< .001
FDP	1.95	.029
AfD	2.71	.002
Other	1.22	.086

Source: SOEP 2015 (.v32); own calculations

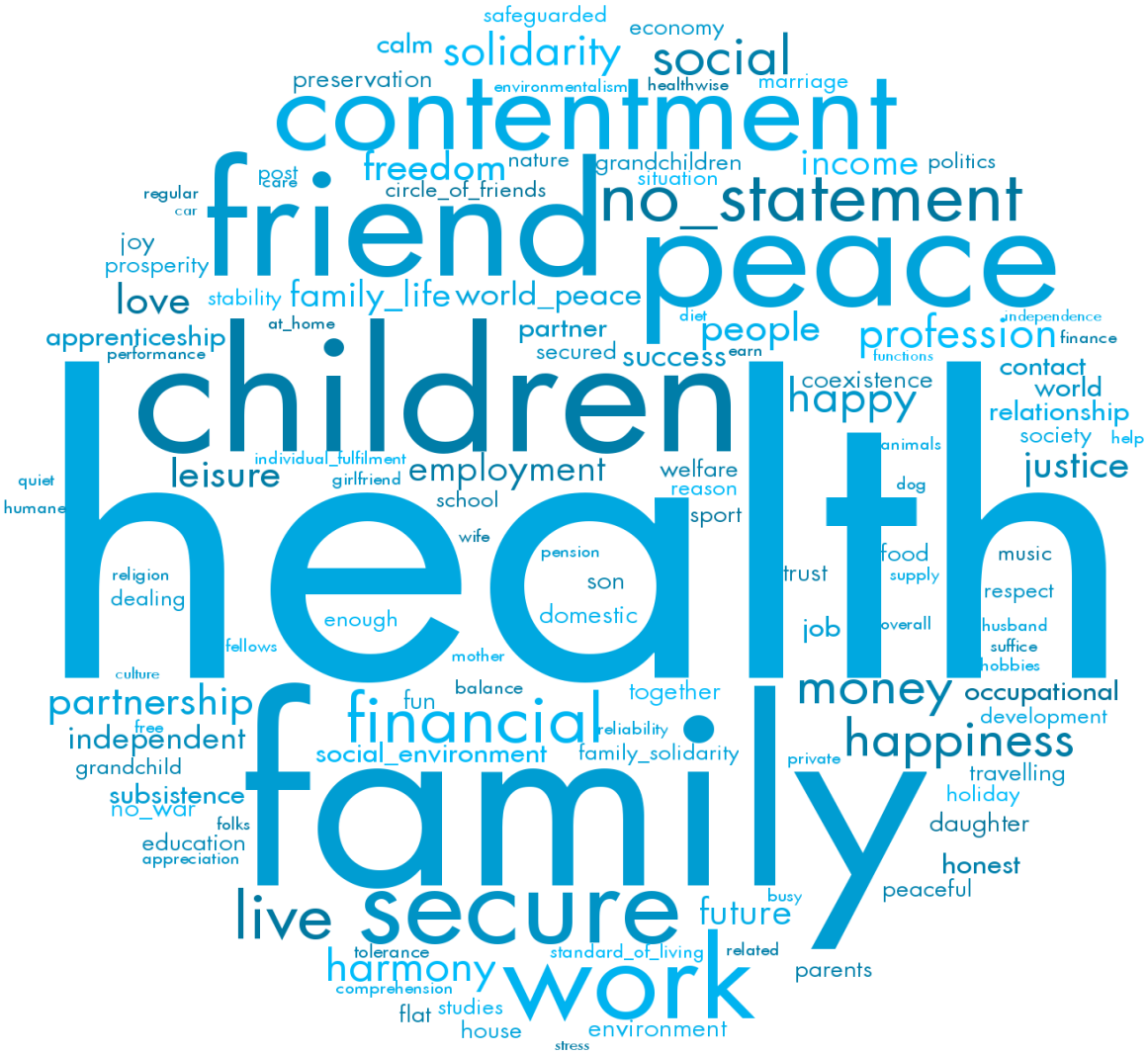
Figure 1

Town Hall Meeting: Chancellor Merkel and SOEP Respondents

Source: Federal Press Office

Figure 2

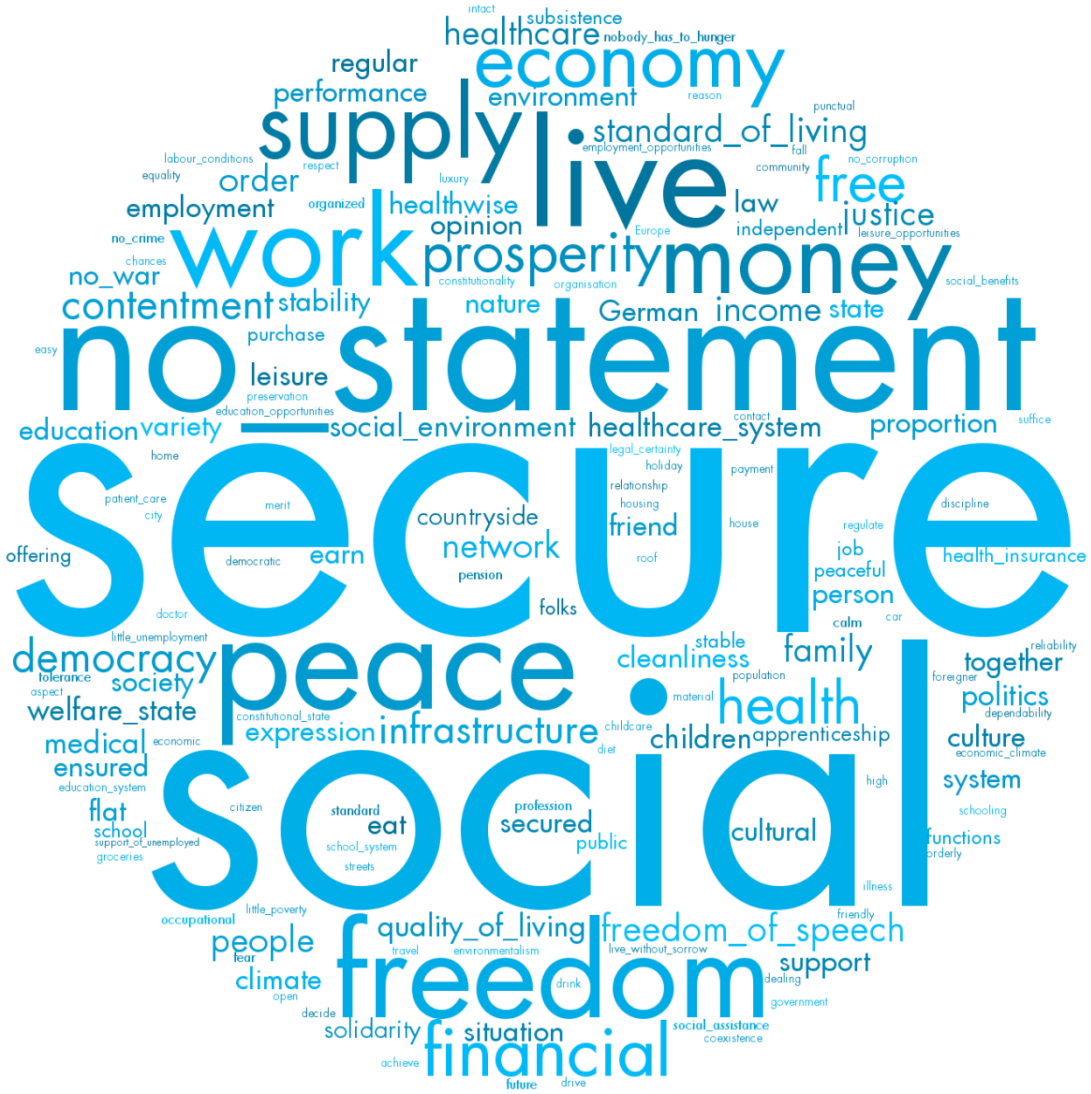
What is personally important to you?



Source: SOEP 2015 (.v32); own calculations.

Figure 3

What constitutes quality of living in Germany for you?



Source: SOEP 2015 (.v32); own calculations.