Measuring the survey climate: the Flemish case

Survey Methods: Insights from the Field

Sara Barbier | Geert Loosveldt | Ann Carton

How to cite this article : Barbier S., Loosveldt G. & Carton A. (2016). Measuring the Survey Climate: The Flemish Case, Survey Methods: Insights from the Field. Retrieved from https://surveyinsights.org/?p=7430

DOI : 10.13094/SMIF-2016-00003

Copyright : © the authors 2016. This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

Abstract : Researchers in several countries have regularly reported decreasing response rates for surveys and the need for increased efforts in order to attain an acceptable response rate: two things that can be seen as signs of a worsening survey climate. At the same time, differences between countries and surveys with regard to the actual level and evolution of response rates have also been noted. Some of these differences are probably linked to differences in the survey content or design. This may hinder the study of the evolving survey climate over time, based on different surveys in different countries, because more readily comparable conditions are desirable. An optimal opportunity for describing the changing survey climate is offered by the Survey of Social-Cultural Changes in Flanders. We analyse yearly data from 1996 to 2013 to examine the evolution of several survey climate indicators. Some indicators reveal a declining survey climate, such as an increased refusal rate and a greater number of contact attempts per respondent. Other indicators reveal a stable survey climate, such as a stable response rate and respondents' positive, stable attitude towards surveys. Results show that, within the same survey, one can compensate for negative evolution by increasing the efforts made to ensure completed interviews.

Introduction

According to Wikipedia, 'climate' is defined as the long-term pattern of weather in a particular area. It is measured by assessing the patterns of variation in temperature, humidity, atmospheric pressure, wind, precipitation, atmospheric particle count and other meteorological variables in a given region over long periods of time (Wikipedia, n.d.). When we try to define 'survey climate' in a similar way, we can state it is the long-term pattern of the interplay between several indicators, for each of which it is relevant to assess how difficult or easy it is to implement a survey in a particular regional or cultural context. Typical of the survey climate is that the indicators are situated at different levels, and the survey climate can therefore be considered a multi-dimensional concept (Loosveldt & Joye, 2016). We examine three dimensions for the indicators: the individual dimension, the societal dimension and the organisational dimension. Examples of indicators for the individual dimension of the survey climate are the willingness of respondents to participate in surveys, the reasons sample units give for refusing to participate, the importance respondents attach to the results of the survey, etc. In the societal dimension, the survey climate is apparent in the number of organised surveys in a region, their content, the way the results are presented in the media, privacy legislation, etc. With regard to the organisational dimension, one can think of the way surveys are organised, the efforts made to persuade sample units to participate and to

maintain the data quality, the number of contact attempts, etc. The assessment of the survey climate is not a straightforward task, because these indicators are not easily quantifiable and they mutually interact. It should be noted that we use a broad definition of ‘survey climate’.

In relevant literature, there are mixed findings regarding the evolution of survey climate indicators. Depending on the specific indicators and surveys, we see either a negative trend or a more stable one. The differences in the evolution of the indicators suggest there is an influence from the specific content and design of a survey. The fact that indicators depend on these characteristics hampers the study of the evolution of the survey climate over an extended period based on different surveys. To address this problem and describe the survey climate appropriately, we need a survey that has been carried out repeatedly over an extended time within the same population. In this way, the essential survey characteristics are nearly constant, which offers the possibility to check if certain indicators of the survey climate, even in similar circumstances, have changed. One suitable survey is the Survey of Social-Cultural Changes, run by the Research Centre of the Flemish Government (SVR). The survey was organised for the first time in 1996 and repeated each year, with the twentieth iteration in 2015. Paradata is available for many survey years (1996–2013) to evaluate several aspects of the survey climate. The same research design, fieldwork procedures and concept definitions have been used each year, which increases the internal comparability of the data. In the next section, we take a closer look at the Survey of Social-Cultural Changes.

Survey of Social-Cultural Changes

The Survey of Social-Cultural Changes (SSC) is a survey of a representative sample of the Dutch-speaking population in the Flemish Region and Brussels-Capital Region of Belgium. It aims to assess the values, opinions and beliefs of respondents with regard to social and policy topics. The random sampling is based on the National Register, and 1500 interviews need to be completed every year. The results of the survey are an important source of information for policy preparation and evaluation, and for scientific research concerning social change. The survey is organised with a great deal of attention and care paid to the quality of the fieldwork and collected data. An overview of these efforts can be found in the yearly process evaluation reports written by the SVR (e.g. Carton, Vander Molen, & Pickery, 2013).

The SVR is responsible for the sample design, the development of the questionnaire, the methodological support and evaluation of the quality of data collection, and the creation of a research report and documented data file. A committee of survey methodologists from several Flemish universities provides methodological and scientific support. The actual data collection is carried out by an external commercial fieldwork organisation and the questionnaire is administered through a face-to-face interview at each respondent’s home. Since 2003, the interviewers have used computer-assisted personal interviewing (CAPI). To reduce the length of the face-to-face interview, respondents receive a drop-off questionnaire at the end of it. They are asked to complete this questionnaire and send it to the research centre. Respondents who sent the drop-off questionnaire to the research centre within an acceptable time span are categorized as spontaneous senders. Respondents who only sent the drop-off questionnaire to the research centre after one to three reminders, are categorized as non-spontaneous senders.

The questionnaire contains items covering several topics, such as participation in cultural activities, sociodemographic characteristics, social networks, leisure activities, Internet use, politics and desire to have children. The length of the interview ranges from 57 to 93 minutes with a mean of 74 minutes (1996–2013). A simple linear regression (see the methodological remark in the endnote), in which the length of the interview is predicted by the survey year, shows that the length of the interview has

increased by 24.3 minutes over 18 years (DW = 2.87, p = 0.957; $F_{(1,16)}$ = 13.46, p = 0.002, b = 1.35).

The change in interview length from one hour to one and a half hours is important to keep in mind, as it can be seen as representing a change in the design of the study. Unfortunately, we also see some other changes. The minimum number of contact attempts interviewers had to make for each sample unit increased from three in 1996 to four in 2001 and to five in 2007. From 2002, the SVR requested that the number of interviewers used for the fieldwork was increased to at least 100, in order to reduce interviewer effects. Moreover, refusal conversion has been implemented since 2007. These changes show that even the observation of one single survey over an extended period does not guarantee the absence of design changes. Design changes can be implemented for several reasons and one possibility is that this is as a reaction to changes in survey climate indicators. In turn, design changes can influence the survey climate indicators. For example, response rates can be altered by the implementation of refusal conversion, which might have been implemented because of decreasing response rates. We should keep this in mind when we interpret the changes in the indicators studied below.

As mentioned, the SSC offers us a good (but not perfect) opportunity to study the multi-dimensional survey climate based on several indicators. This was carried out in the extensive report by Barbier, Loosveldt and Carton (2015). Below, we summarize part of the analyses of all the indicators in this report. First, we examine the indicators for the individual dimension. Second, we consider the indicators for the organisational dimension. No indicators concerning the societal dimension are discussed here, because they are not directly based on one specific survey but instead on legislation or the media, or on a range of surveys. Apart from this, we can say that the willingness of the Flemish government to fund an expensive CAPI survey is a clear indication of a positive societal survey climate in Flanders. We end this paper with conclusions and a discussion.

Indicators for the individual dimension of the survey climate

At the individual level, we discuss some nonresponse indicators, the contrast between respondents and non-respondents, and the attitude of respondents towards surveys.

Nonresponse indicators

Outcome rates are by far the most common indicators for the survey climate. In the SSC, they are calculated based on American Association for Public Opinion Research guidelines (AAPOR, 2015). The response rate (RR1) is the number of completed interviews divided by the number of (potentially) eligible sample units (completed interviews, refusals, non-contacts, others and unknown eligibility). The refusal rate (REF1) is the number of refusals divided by the number of (potentially) eligible sample units. The co-operation rate (COOP1) is the number of completed interviews divided by the number of eligible sample units with at least one successful contact attempt (completed interviews, refusals and others). The contact rate (CON1) is the number of eligible sample units with at least one successful contact attempt divided by the number of (potentially) eligible sample units.

Figure 1 shows the evolution of outcome rates from 1996 to 2013. We excluded survey years 1998 and 1999 from the analyses due to notably deviating patterns, which probably originate from organisational problems. Four simple linear regressions were executed, in which the outcome rate is predicted by the survey year (RR1: DW = 2.00, p = .385; $F_{(1,14)}$ = 0.18, p = .680; COOP1: DW = 1.75, p = .207; $F_{(1,14)}$ = 10.98, p = .005, b = -0.44; REF1: DW = 1.18, p = .019; $F_{(1,14)}$ = 11.80, p = .004, b = 0.50; CON1: DW =

2.45, p = .738; $F_{(1,14)}$ = 9.86, p = .007, b = 0.46).

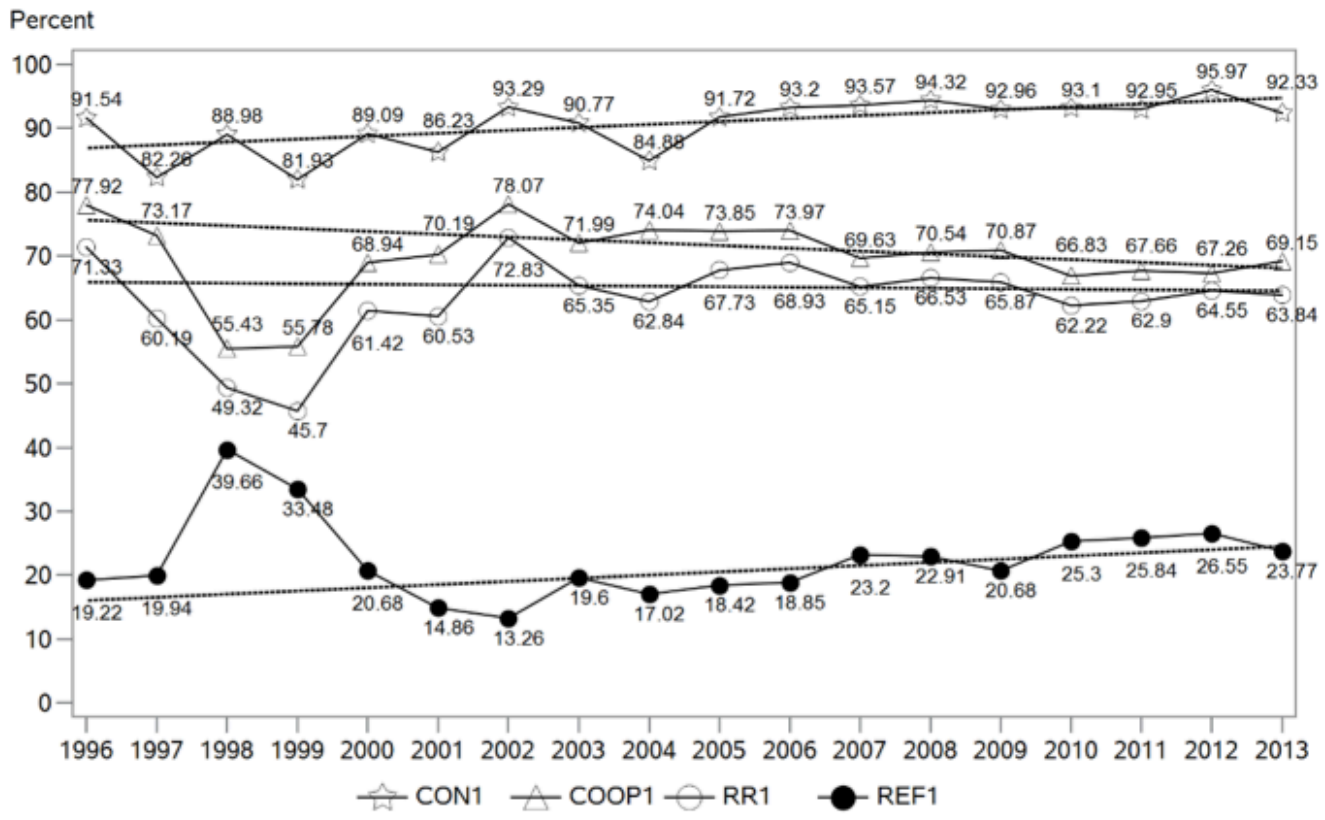


Figure 1: Outcome rates: contact rate (CON1), cooperation rate (COOP1), response rate (RR1) and refusal rate (REF1)

The first notable result is that the response rate, RR1, remained stable over a period of 18 years with a mean of 65.14%. By contrast, the co-operation rate, COOP1 (mean: 71.51%), decreased by 7.92 percentage points over 18 years and the refusal rate, REF1 (mean: 20.63%), increased by 9 percentage points over the same period. This apparent contradiction can be explained by the increase in the contact rate, CON1 (mean: 91.14%), of 8.28 percentage points over the 18 years. Through the years, sample units more frequently refused to participate in the interview. However, because interviewers were able to contact more sample units, this is not visible in the overall response rates. The increase in contact rates can be linked to changes in the design of the SSC. As mentioned above, the contact procedure was extended over the years: the minimum number of contact attempts per sample unit increased and refusal conversion was implemented.

In addition to the traditional outcome rates, we describe three other survey climate indicators at the individual level that are related to participation. The first is the extent to which a first contact attempt resulted in (an appointment for) an interview. This tells us something about how easy it was for the interviewers to complete an interview. Relevant data is available for 2000 to 2013. The percentage of first contact attempts that resulted in (an appointment for) an interview ranges from 27.32% to 47.34% with a mean of 35.96%. The result of a simple linear regression, where this percentage is predicted by the survey year, is significant (DW = 2.14, p = .481; $F_{(1,12)}$ = 11.57, p = .005, b = -0.79). The percentage decreased by 11.06 percentage points over 14 years. In those years, it became more difficult for the interviewers to establish (an appointment for) an interview at the first contact attempt.

The second indicator is the extent to which respondents say they would participate again or refuse to participate again in a survey such as the SSC. A decline in the willingness of respondents to participate again in a comparable survey can be seen as a sign of a deteriorating survey climate. Data is available for 2002 to 2013. The percentage of respondents who stated they would participate again ranges from 81.59% to 87.45% with a mean of 84.18%. The result of a simple linear regression, where this percentage is predicted by the survey year, is not significant (DW = 1.86, p = .273; $F_{(1,12)}$ = 0.74, p = .405). The percentage of respondents who said they would participate again in a comparable survey was stable over time.

The third indicator is the (non)response to the drop-off questionnaire. A decline in the number of respondents sending the drop-off questionnaire can be seen as a sign of a deteriorating survey climate. Relevant data is available for 2002 and for 2004 to 2012. The percentage of senders (spontaneous and non-spontaneous) ranges from 80.63% to 92.08%, with a mean of 87.78%. This response rate is high. The result of a simple linear regression, where this percentage is predicted by the survey year, is significant (DW = 1.50, p = .099; $F_{(1,8)}$ = 12.25, p = .008, b = -0.86). The percentage decreased by 8.6 percentage points over 10 years. The decline is even worse for the percentage of spontaneous senders. This percentage ranges from 53.75% to 78.40% with a mean of 67.16%. The result of a simple linear regression, where this percentage is predicted by the survey year, is significant (DW = 1.94, p = .303; $F_{(1,8)}$ = 23.93, p = .001, b = -2.20). The percentage decreased by 22 percentage points over 10 years. This means the willingness of respondents to (spontaneously) send the questionnaire clearly declined over time.

If we only look at the response rate (RR1) and the extent to which respondents stated they would participate again in a comparable survey, the survey climate appears stable. However, a thorough examination of other indicators at the individual level undermines this positive conclusion. Over the years, the interviewers' first contact attempts have become less successful and respondents have been less willing to send the drop-off questionnaire (spontaneously). Similarly, co-operation rates have decreased and refusal rates have increased. Based on these indicators, the survey climate seems to have deteriorated. The increasing contact rate should accordingly be seen as a consequence of design changes in the contact procedure and increased interviewers' efforts, rather than as a sign of an improving survey climate.

Contrast between respondents and non-respondents

Another way to learn something about the survey climate is to look at the evolution of the contrast between respondents and non-respondents. In comparison with the response rate, it can also indicate changes over time in the representation of particular groups of respondents. There is always a degree of bias in a realized sample due to over or under representation of some groups (such as people with a low education level). This bias can be addressed statistically by weighting the data and the standard deviation of the weighting coefficients can be seen as an indicator of the survey climate. If the standard deviation increases, the contrast between respondents and non-respondents and the related bias increases. This means that over time, some groups are less well represented than others in the realized sample.

In the SSC, weights are based on gender, age and education level. The weights are transformed with a $\log_{10}$-transformation to create more-symmetric data. The minimum value ranges from -1.72 to -0.40 with a mean of -0.91, the maximum value ranges from 0.89 to 2.14 with a mean of 1.37 and the standard deviation ranges from 0.21 to 0.33 with a mean of 0.29. Three simple linear regressions were executed, in

which a characteristic of the transformed weight was predicted by the survey year (1996 – 2013). None of the results are significant (minimum: $DW = 1.55$, $p = .100$; $F_{(1,16)} = 2.74$, $p = .117$; maximum: $DW = 1.57$, $p = .109$; $F_{(1,16)} = 0.32$, $p = .578$; standard deviation: $DW = 1.25$, $p = .023$; $F_{(1,16)} = 2.65$, $p = .123$). Overall, the contrast between respondents and non-respondents for age, gender and education is stable over the years, as the standard deviation of the weights is stable. This means that no specific groups are systematically over or under represented over time in the realized sample, which can be considered as a positive aspect of the survey climate.

Attitude of respondents towards surveys

As our last individual indicator for the survey climate, we examine the attitude of respondents towards surveys such as the SSC, which contains some questions to measure this. Based on an analysis of these questions, we can conclude that in general, respondents have a positive attitude towards surveys. Respondents describe the SSC as a pleasant experience with easy questions. They see the results of surveys as credible and useful for policy, because the government gets an idea of the opinion of the population. Surveys are not seen as an invasion of privacy or as a waste of time. Respondents do not want to be paid for their co-operation and feel that all people have a responsibility to participate in surveys. In addition to being positive, the attitude of the respondents is stable over time. There are no significant changes in the way respondents answered these questions over the years (2000–2013), which indicates a stable 'subjective' survey climate. Nevertheless, some remarks need to be made regarding this type of indicator. First, most of the questions were asked in the face-to-face interview, but in some years, a number were included in the drop-off questionnaire. The comparison of these answers reveals that the face-to-face answered questions are clearly positively biased due to social desirability. Second, this stable 'subjective' survey climate is based on respondents only. Unfortunately, nothing is known about the attitude of the non-respondents.

Indicators for the organisational dimension of the survey climate

Individual indicators tell us something about changes in the survey climate. However, using these indicators alone would result in a narrow perspective. The survey climate is a broad concept and many different indicators are required in order to obtain an accurate idea of its evolution. The climate is, for example, also observable in the way a survey is implemented. We examine several organisational indicators: the number of interviewers used to complete the total number of interviews compared with the length of the fieldwork, the number of contact attempts and the financial cost of a completed interview. These are also (indirect) indicators of how easy or difficult it is to conduct surveys.

Many interviewers were employed in order to reach the amount of 1500 interviews each year, and the number of interviewers increased over the years (See Figure 2). In the first three years, there were around 53 interviewers. In the next three years, there were around 87. Since 2002, the number of interviewers has been approximately 100, as requested by the SVR to reduce interviewer effects. The fieldwork organisation needs a certain amount of time to complete the 1500 interviews, expressed by the number of days of the fieldwork period. We expected the length of the fieldwork period to decrease, as the number of interviewers to reach the same number of completed interviews increased and all these interviewers could work simultaneously. A simple linear regression where the length of the fieldwork period was predicted by the survey year produces a result that is not significant ($DW = 1.93$, $p = .330$; $F_{(1,15)} = 0.04$, $p = .841$). The length of the fieldwork period was stable throughout the years. This is a remarkable finding: although there were almost twice as many interviewers in the later years of the survey than in its first years, the fieldwork period remained unchanged. The increase in the number of

interviewers only stabilised the fieldwork period, but did not decrease it (See Figure 2). We see this as a sign of a deteriorating survey climate, as a larger group of interviewers needed the same amount of time as a smaller group of interviewers to achieve the same number of completed interviews. This can be linked to the increased workload of the interviewers (longer interviews, intensified contact procedure and fewer successful first contact attempts).

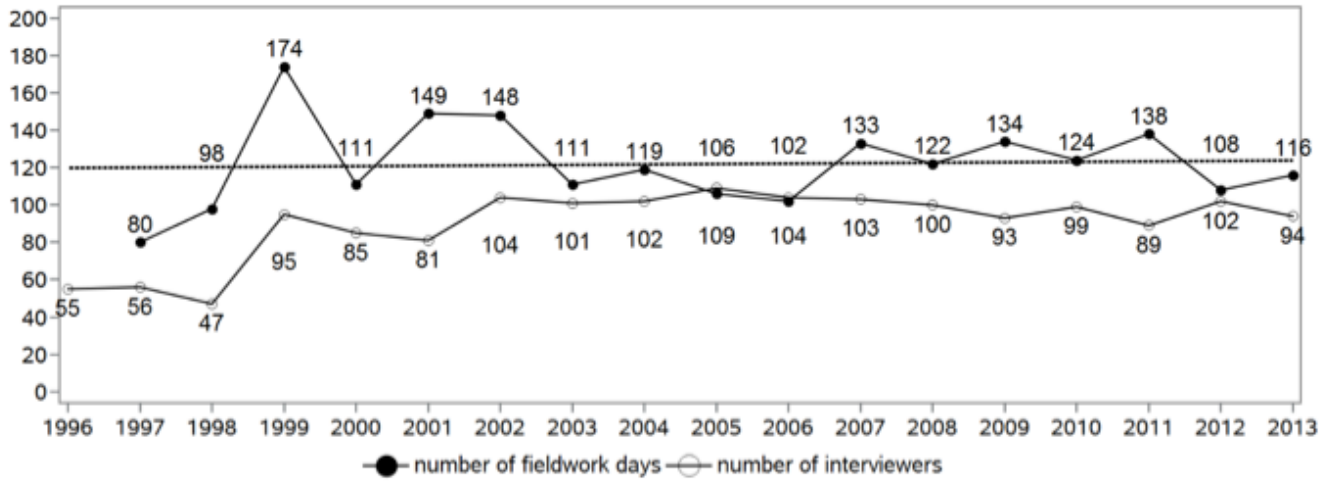


Figure 2: Organisational indicators: number of fieldwork days and number of interviewers

Another indicator of the effort required to realise an interview is the mean number of contact attempts needed until a final outcome can be assigned to a sample unit. Relevant data is available for 2000 to 2013. The mean number of contact attempts ranges from 2.27 to 2.93 with a mean of 2.62. The result of a simple linear regression that predicts this number by the survey year, is significant (DW = 1.52, p = .100; $F_{(1,12)}$ = 51.90, p < .0001, b = 0.05). The mean number of contact attempts increased by 0.7 over 14 years. This indicates, together with the decrease in the percentage of first contact attempts with an (appointment for an) interview, that the survey climate is less 'enjoyable' and that interviewers needed to make a greater effort in order to realise an interview.

In social sciences, a face-to-face survey is an expensive form of research. A large amount of the survey costs relate to the fieldwork (the cost of the fieldwork organisation and of the sampling of the National Register). The fieldwork costs are expressed by the price of a completed interview and are indexed to make a year-to-year comparison valuable. Data is available for 2004 to 2013. The cost for a completed interview ranges from €136.13 to €167.45 with a mean of €154.84. The result of a simple linear regression, which predicts this cost by the survey year, is significant (DW = 2.15, p = .436; $F_{(1,8)}$ = 7.24, p = .028, b = -2.47). The cost for a completed interview decreased by €24.7 over 10 years, which is a striking finding. We expected the fieldwork costs to increase, because the interviewers' effort per interview had increased: there were more contact attempts per respondent and longer interviews. Moreover, the fieldwork organisation also needed to direct and control more interviewers in later years. One possible explanation for the decreasing costs is that survey agencies implement fieldwork for unrealistically low prices, because they see a governmental survey as being positive for their image instead of as a source of profit.

Conclusion and discussion

In this paper, we investigate the survey climate in Flanders based on the SSC. This survey has been run annually since 1996 and it offers a unique opportunity to study the survey climate in the same region over an extended period. By using one survey, we can largely remove the influence of different design characteristics on the survey climate indicators. Although the design of the SSC was relatively stable during the whole period, some changes were implemented as described above. This is inherent to a real life survey organized for a longer period, as improvements are made over time. We used the changes as relevant context information and they were taken into account when interpreting the trends in the survey climate indicators.

There are mixed findings for the direction in which the survey climate is changing, for both the individual and the organisational dimensions of the survey climate. At first sight, a stable survey climate is evidenced by the constant response rates, the unchanging contrast between respondents and non-respondents, the positive and stable attitude of respondents towards surveys and their stable willingness to participate again in a comparable survey. However, a deteriorating survey climate is suggested by an increasing refusal rate, a decreasing co-operation rate, fewer successful first contact attempts, more contact attempts per sample unit and a decreasing response rate for the drop-off questionnaire. Moreover, the response rate only seems to be stable because an increase in the contact rate compensates for the increase in refusal rate and decrease in co-operation rate. The increasing contact rate probably originates from the design changes that intensified the contact procedure. These results indicate that the changes in the individual measures stimulate changes in the organisational measures. The positive and stable attitude towards surveys and the stable willingness to participate again in a comparable survey can only be measured for respondents. There is no relevant data about the opinion of non-respondents.

The use of different indicators for different dimensions of the survey climate compels us to make nuanced judgments about the changes in the Flemish survey climate. However, based on the results it seems justified to conclude that the survey climate in Flanders from 1996 to 2013 has moderately deteriorated. Our results show that within the same survey, one can compensate for this moderate negative evolution by increasing efforts to ensure the desired number of completed interviews. Nevertheless, the range of improvements available to optimise the contact procedure seems to be almost exhausted. The next area where researchers and interviewers can intervene in order to counteract the decreasing survey climate is that of refusal. We can think about better training for interviewers, incentives for respondents and shorter interviews. However, if the survey climate is really worsening, these extra efforts will probably only provide a temporary solution.

Although in this paper we provide a profound analysis of the survey climate using several indicators for two dimensions, the societal dimension is missing. We do not report anything about privacy legislation and the changes in this domain in Flanders since 1996. Nor do we examine the way survey results are presented in the media and if, or how, this has changed over the last 20 years. Findings regarding the general societal level of the survey climate can enrich our knowledge about it and can help us to understand some of the changes. Further, similar analyses related to surveys repeatedly organised in other regions or countries can contribute to the generalization of the Flemish results and to a more general assessment of the survey climate.

The analyses nevertheless reveal that over the years, the job of the interviewer has become more difficult and stressful: the contact procedure has been extended, the interview length has grown, refusal rates have increased, co-operation rates have decreased, the first contact attempt has become less successful

and the average number of contact attempts per sample unit has increased. As interviewers are key players in the data collection of substantive variables and the paradata needed to assess the data quality, good training and appropriate remuneration for the interviewers is important in order to obtain high quality data in a less comfortable survey climate.

Endnote

Methodological remark: The use of simple linear regression to analyse time series is not optimal because of the risk of autocorrelation. To test for autocorrelation, we used the Durbin-Watson statistic (DW), where a test statistic close to 2 and the associated non-significant p-value indicate the absence of autocorrelation in the data. In this case, we were able to use simple linear regression. If the test statistic is significantly different from 2, autocorrelation in the data is present and care is needed for the significance tests of the simple linear regression. Due to autocorrelation, the standard error of the regression coefficients might be underestimated. In these cases, we used 0.01 as the critical p-value for significance instead of 0.05.

References

1. The American Association for Public Opinion Research. (2015). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys.* 8th edition. AAPOR.
2. Barbier. S., Loosveldt, G., & Carton, A. (2015). *Het surveyklimaat in Vlaanderen: een analyse op basis van de SCV-surveys.* Brussel: Vlaamse overheid, Studiedienst van de Vlaamse Regering, SVR-Methoden & Technieken.
3. Carton, A., Vander Molen, T., & Pickery, J. (2013). *Sociaal-culturele verschuivingen in Vlaanderen 2011. Methodologisch-technisch rapport en procesevaluatie van de dataverzameling.* Brussel: Vlaamse overheid, Studiedienst van de Vlaamse Regering, SVR-Methoden & Technieken 2013/3.
4. Loosveldt, G., & Joye, D. (2016). Defining and Assessing Survey Climate. In Wolf, C., Joye, D., Smith W., & Fu, Y. (Eds.), *The SAGE Handbook of Survey Methodology* (pp. 67-77). London, England: Sage.
5. Wikipedia (n.d). *Climate*. Retrieved October 15, 2015 from https://en.wikipedia.org/wiki/Climate.