# Mixed-Mode Experiment - Evaluation of Effects on Data Quality and Response Rates

Johan Eklund[1], Mikaela Järnbert[2], Tiina Orusild[3], Andreas Persson[4], Fredrik Scheffer[5,]

[1] *Statistics Sweden, Stockholm, Sweden; johan.eklund@scb.se*
[2] *Statistics Sweden, Stockholm, Sweden; mikaela.jarnbert@scb.se*
[3] *Statistics Sweden, Stockholm, Sweden; tiina.orusild@scb.se*
[4]*Statistics Sweden, Stockholm, Sweden; andreas.persson@scb.se*
[5] *Statistics Sweden, Stockholm, Sweden; fredrik.scheffer@scb.se*

**Abstract**
Over recent decades, response rates in interview surveys have decreased rapidly in Sweden as in many other countries. This has led to higher uncertainty in estimates as well as higher costs of data collection. One way to deal with this problem is to allow different response modes. In this paper we examine what happens with response behavior in a panel survey when web mode is introduced as an additional mode in a traditional single mode telephone interview survey. Three experiments have been carried out in 2014 and 2015 mixing telephone interviews and web mode within a political opinion poll at Statistics Sweden. The experiments have so far shown positive results on increased response rates. The effect of offering web is also positive on new panels.

**Keywords:** Mixed-mode**,** Response rates, Data quality, mode effects

## 1. Introduction

Statistics Sweden (SCB) carries out surveys on Party Preferences twice a year, in May and November, since 1972 on request of the Swedish Riksdag. The purpose of the survey is to estimate a hypothetical parliament election result and party preferences both in general and within different parts of the population. The target population is those who would have been entitled to vote if there had been a parliament election. The survey does not have any upper age limit. The sampling frame is constructed from the total population register. The Party Preference Survey is a panel survey and consists of three panels. Each panel consists of approximately 3 000 individuals and each panel takes part in the survey three times before being rotated out. In every survey one panel is participating for the first time, one for the second time and one for the last time. Consequently every survey constitutes of a total of 9 000 individuals. Data is collected through telephone interviews and the survey is announced to the sampled individuals through an advance letter the first time the sample individual is

included in the survey. The fieldwork period is usually somewhat less than four weeks. During the latest decade the nonresponse rate has increased considerably in Sweden, as in many other countries. More and more individuals decline to participate in surveys and it is also getting harder to get in touch with individuals in the sample for an interview. This is causing an increased nonresponse rate, often with unknown consequences for the estimates. Between 1984 and 2015 both the share of refusals and the share of noncontacts have increased while the share of those unable to participate has been constant during the whole time period. In order to deal with the increasing nonresponse an experiment was carried out on the Party Preference Survey when web was offered as an additional mode in the data collection. A first, small scale, experiment was carried out in September 2014 and larger experiments were then carried out in May and November 2015.

## 2. Mixing interview and self-administered modes

One concern when mixing data collection methods is the risk of mode effects. That is, that the data collection method itself has an effect on the measurement. Previous research shows that mode effects are especially important to consider when mixing self-administered methods (such as the web) and interviewer-based methods (such as telephone interview) (de Leeuw, 2008). The risks of mode effects stem from two major differences between interviews and self-administered methods (de Leeuw, 2005; de Leeuw, 2008). An interview is a social encounter whereas self-administered methods are private and the information type differs between interviews and self-administered methods. It is visual in self-administered methods but auditive in interviews. Questions generally have to be shorter in interviews than in self-administered methods and it is not given that the interviewer should read the response options. In addition, research on paper questionnaires show that question format and layout can influence the results. However, such factors are not relevant at all for telephone interviews which are based on sound and lack visuals (de Leeuw, 2005; de Leeuw, 2008). Thus, when mixing web and telephone interviews one has to consider that these methods differ in how they transfer information. The party questions are open-ended in the telephone interview. The interviewers do not read the response options and are instructed not to probe or provide

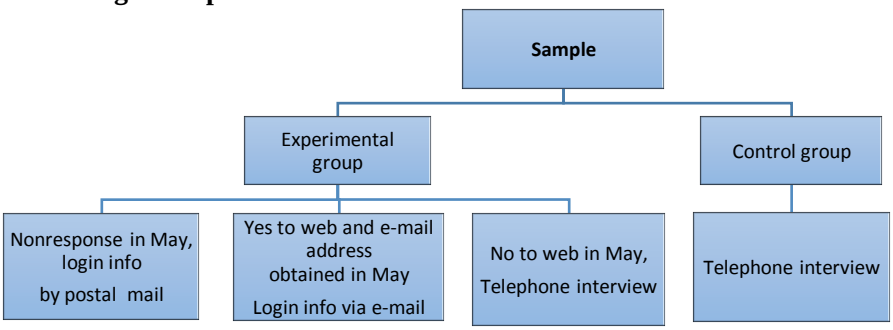
examples. How, then, should these questions be best designed for the web? Do they have to be open-ended on the web? A closed question would reduce both the respondent burden and the need for coding. However research on questions with broad topics shows that if the respondent must generate possible response options on their own, it is not certain that they will arrive at precisely the same options as if the question was closed (Schuman & Presser, 1981). Concerning political opinion polls, an open question might benefit parties that are salient in peoples mind, due to, for example, more media exposure. In a closed question, this top-of-mind effect will likely not be operating to the same extent since all the parties are listed and visible in the response options. Thus, open ended questions, when the respondents have to generate the response options themselves, might lead to a higher endorsement for parties that are more easily accessible in memory, compared to a closed question. In closed questions, however, the response options must have an order, and that order might matter. Previous research show that closed questions in visual modes are associated with a higher endorsement of early response options (Schwarz, Knäuper, Oysermann & Stich, 2008). This is due to two mechanisms: 1) the primacy effect (Glanzer, 1972), and 2) satisficing (Krosnick & Alwin, 1987). The primacy effect has to do with cognitive factors whereas satisficing primarily has to do with a lack of motivation. Thus, both the primacy effect and satisficing can cause a higher endorsement of early response options in closed questions in visual modes. Based on advantages and disadvantages mentioned we developed an open-ended question with suggestions. For example, if the respondents enter “Li” the web questionnaire suggests the party “Liberalerna” (The Liberal Party). Concerning the don’t know option on the web, we aimed to mimic the telephone interview as much as possible. The solution was not to include a don’t know option. However, if the respondent tried to progress in the questionnaire without entering any information in the text field, they were prompted to write “don’t know” if they did not know the answer to the question.

## 3. Mixed-mode experiments

Statistics Sweden has carried out three mixed-mode experiments within the Party Preference Survey using telephone interviews and web questionnaire in the data collection. To compare

the two approaches, mixed-mode and single mode, the samples for these surveys were divided randomly into an experimental group and a control group, where the sample individuals in the experimental group were offered a web questionnaire as a complement to the telephone interview. The individuals in the control group were only contacted by telephone. Since the survey is a panel survey the individuals in the randomly assigned experimental group were asked in the previous survey round whether or not they would prefer to answer the survey questions by telephone or web the next time. This resulted in three different contact strategies in the experimental group – those who said yes to answer by web, those who said no and nonrespondents in the previous round. Those who said no were contacted by telephone at the start of the field period; those who said yes did provide us with their e-mail address in the previous interview and were contacted by e-mail; to nonrespondents in the previous round information on login was sent by postal mail. In the figure below the experiment is illustrated. The experiments have been carried out using mainly the same methodology. There is however some differences between them and in this section we describe the three experiments in more detail below.

**Figure 1. Design of experiment and treatments**




Experiment 1 in September 2014: In this experiment one third of the ordinary sample size was used, i.e. only one of the three panels consisting of 3 000 individuals. The sample was randomly assigned to an experimental group (about 2 000) and a control group (about 1 000). The primary goal of the experiment was to test the web questionnaire, the data collection system and evaluate the response rates, data quality and consistency of answers between survey rounds between the groups.

Experiment 2 in May 2015: Total sample size of 12 000. The sample was randomly assigned to an experimental group (about 5 500) and a control group (about 6 500). The primary goals of the experiment was to confirm the results from the first experiment, to evaluate the effect of offering the web alternative to the new panel, to evaluate mode effects for the main variables of interest and analyze the representativity of the response set (R-indicator analysis).

Experiment 3 in November 2015: All three panels were divided into a control and an experimental group, the experimental group with a size of about 6 000 and the control group of about 3 000. The primary goal of the experiment was to confirm the results from the previous experiments but also to test a design where all the three panels are offered mixed-mode alternatives with different contact strategies.

## 4. Response rates

### *4.1 Response rates in experimental and control group*

We received a higher response rate in the experimental group compared to the control group in all three experiments. The difference was 4.9 percentage points in September 2014, 7.0 percentage points in May 2015 and 5.6 percentage points in November 2015, see Table 1. There was no big difference between the two groups regarding shares of persons unable or declining to participate, but the shares of noncontacts were lower in the experimental groups in all three experiments.

**Table 1. Response rates. September 2014, May 2015 and November 2015. Percent**

| | **September 2014** | | **May 2015** | | **Nov 2015** | |
|---|---|---|---|---|---|---|
| | Experimental group | Control group | Experimental group | Control group | Experimental group | Control |
| Respondents | 55.4 | 50.5 | 54.0 | 47.0 | 57.0 | 51.4 |
| Unable to participate | 2.8 | 2.1 | 2.6 | 2.9 | 3.0 | 3.5 |
| Decline to participate | 12.5 | 13.4 | 15.3 | 16.0 | 15.2 | 17.0 |
| Noncontacts | 29.3 | 34.0 | 28.1 | 34.2 | 24.8 | 28.1 |
| Sum | 100 | 100 | 100 | 100 | 100 | 100 |
| Total | 2 042 | 1 020 | 5 561 | 6 530 | 6 022 | 2 999 |

### *4.2 Response rates related to previous survey round*

It is also interesting to compare the experimental and control groups by response category over two survey rounds. If we compare the response rates in the experimental group and the control

group by response category in the previous survey round we can see a higher response rate in the experimental group among those who declined and noncontacts in previous survey rounds for the experiments in both May and November 2015.

**Table 2. Response rates. May 2015 and November 2015 by response category in previous survey round. Percent**

| November 2014 | Response rate May 2015 | |
|---|---|---|
| | Experimental | Control |
| Interview | 75.5 | 65.8 |
| Decline to participate | 23.0 | 14.5 |
| Noncontacts | 32.1 | 14.2 |

| May 2015 | Response rate November 2015 | |
|---|---|---|
| | Experimental | Control |
| Interview | 85.6 | 86.1 |
| Decline to participate | 21.5 | 17.3 |
| Noncontacts | 33.9 | 28.9 |

Another way to look at the response rates related to previous survey round is to study the respondents in the previous survey round and look at the response category in the successive round. Here the two experiments do not show similar results. In May there was a higher response rate in the experimental group and a greater share of noncontacts in the control group. In November we can't see any clear differences between the two groups.

**Table 3. Response category in May and November 2015 among respondents in previous round. Percent**

| May 2015 | Experimental group | Control group |
|---|---|---|
| Interview | 75.5 | 65.8 |
| Unable to participate | 0.7 | 1.8 |
| Decline to participate | 5.3 | 5.1 |
| Noncontacts | 18.6 | 27.3 |
| Total | 100 | 100 |

| November 2015 | Experimental group | Control group |
|---|---|---|
| Interview | 85.9 | 86.2 |
| Unable to participate | 0.8 | 0.8 |
| Decline to participate | 3.7 | 4.2 |
| Noncontacts | 9.6 | 8.9 |
| Total | 100 | 100 |

### *4.4 Response rates by demographic characteristics*

As shown above all experiments rendered positive results regarding response rates by using mixed-mode compared to single mode. But does the introduction of web as an alternative to telephone interview have equally positive effect in all demographic subgroups or better yet even out the existing difference that already is produced by telephone as a single mode? To receive a better representativity in the response set we were hoping for a greater effect in groups with low response rates, young, low educated and persons that are not born in Sweden.

**Table 4. Response rates by demographic characteristics. September 2014, May 2015 and November 2015. Percent**

| | September 2014 | | | May 2015 | | | November 2015 | | |
|---|---|---|---|---|---|---|---|---|---|
| | % Exp | % Contr | Diff in % points Exp -Contr | % Exp | % Contr | Diff in % points Exp -Contr | % Exp | % Contr | Diff in % points Exp -Contr |
| **Total** | **55.4** | **50.5** | **4.9** | **54.0** | **47.0** | **7.0** | **57.0** | **51.4** | **6.0** |
| Men | 57.2 | 52.3 | 4.8 | 54.3 | 49.7 | 4.6 | 60.3 | 54.9 | 5.5 |
| Women | 53.7 | 48.7 | 5.0 | 53.6 | 44.3 | 9.3 | 53.7 | 48.0 | 5.7 |
| 18-29 years | 42.6 | 44.1 | -1.5 | 45.4 | 39.1 | 6.3 | 46.0 | 43.0 | 3.0 |
| 30-49 years | 55.7 | 47.6 | 8.1 | 52.3 | 43.7 | 8.6 | 55.7 | 51.3 | 4.4 |
| 50-64 years | 57.4 | 48.4 | 9.0 | 54.8 | 46.7 | 8.1 | 58.5 | 51.6 | 6.9 |
| 65+ years | 63.4 | 61.2 | 2.2 | 61.2 | 55.9 | 5.3 | 64.7 | 57.0 | 7.6 |
| Big city | . | . | . | 53.4 | 44.3 | 9,1 | 56.6 | 49.6 | 7.0 |
| Other | . | . | . | 54.1 | 47.5 | 6,6 | 57.1 | 51.8 | 5.3 |
| Compulsory edu. | 45.5 | 39.9 | 5.6 | 44.9 | 43.1 | 1.8 | 49.6 | 42.8 | 6.8 |
| Upper secondary edu. | 51.6 | 48.9 | 2.7 | 49.1 | 43.2 | 5.9 | 52.3 | 48.2 | 4.1 |
| Post-secondary edu. | 67.7 | 58.8 | 8.9 | 66.9 | 54.9 | 12.0 | 68.2 | 61.8 | 6.4 |
| Born outside of Sweden | 44.0 | 40.7 | 3.3 | 41.6 | 39.1 | 2.5 | 47.4 | 43.8 | 3.6 |
| Born in Sweden | 56.9 | 51.8 | 5.1 | 55.6 | 48.1 | 7.5 | 58.3 | 52.4 | 5.9 |
| **Number** | 2 042 | 1 020 | | 5 561 | 6 530 | | 6 022 | 2 999 | |

Unfortunately, mixed-mode does not seem to be a remedy. In the surveys in September 2014 and May 2015 we can see a bigger effect on high educated persons compared to persons with only compulsory school, as well as among persons born in Sweden compared to persons born in another country. In November 2015 the difference regarding response by education is smaller, which is positive. In November there are also quite small differences in the effect of offering mixed-mode regarding sex and country of birth. Regarding age the effect seems, rather surprising, largest in the upper age groups. In total, the introducing of mixed-mode seems to give us a higher response rate, but not a more representative response set.

*4.5 R-indicator analysis*

In this section we present some results on indicators of representativity (R-indicators) see e.g. Shlomo et al. (2012). R-indicators were estimated for the total sample, for the experimental group and the control group. Data from all three experiments were used. As a comparison, R-

indicators were also estimated for the surveys in May 2014 and November 2014 when only one data collection mode was used i.e. telephone interview. To estimate R-indicators we need estimates of the response propensities. In Table 5 below, the response propensities were estimated using logistic regression. In the table the sample sizes, response rates, estimated R-indicators and approximate 95 percent confidence intervals (CI) for the R-indicator are shown. The confidence intervals were calculated using the variance estimator proposed by Shlomo et al. (2012). See also papers and tools for computation of R-indicators on the RISQ-project website (www.risq-project.eu/).

**Table 5. R-indicators for the experiment conducted in September 2014, May and November 2015**

| | September 2014 | | | | May 2015 | | | | November 2015 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Sample size | Response rate (%) | R-ind (%) | CI | Sample size | Response rate (%) | R-ind (%) | CI | Sample size | Response rate (%) | R-ind (%) | CI |
| **Total** | 2 986 | 53.8 | 76.3 | (72.9, 79.6) | 12 091 | 50.2 | 78.5 | (76.8, 80.2) | 9 021 | 55.1 | 76.9 | (75.0, 78.8) |
| **Experimental** | 1 999 | 55.4 | 75.0 | (70.9, 79.1) | 5 561 | 54.0 | 75.8 | (73.4, 78.3) | 6 022 | 57.0 | 76.1 | (73.7, 78.4) |
| **Control** | 987 | 50.5 | 79.5 | (73.4, 85.6) | 6 530 | 47.0 | 80.5 | (78.1, 82.8) | 2 999 | 51.4 | 79.0 | (75.6, 82.5) |

The values of the R-indicators depend on which auxiliary variables are used. In the table above we show results for the auxiliary variables sex, age group, region, educational attainment and country of birth. The categories of the variables are the same as in Table 4. For this choice of auxiliary variables, the control groups seem to perform better than the experimental groups in all three experiments. We have also used other auxiliary variables and in most cases the estimated R-indicator for the control group is larger than for the experimental group.

## 5. Mode effects

In Section 4 we have seen that adding web as an additional mode to a telephone survey seems to have a positive effect on the response rate. However, all three experiments show indication on a slightly less balanced response set for mixed-mode. The other central question for carrying out the experiments was whether or not mixing modes would introduce differences in estimates and quality of data. To compare mixed-mode (experimental group) and single mode

(control group) we estimated proportions for the variable party vote if election today. The estimates were calculated using the whole sample, the experimental group and the control group using data from the experiments in May and November 2015, i.e. the larger scale experiments. The results are presented in Section 5.1. In Section 5.2 item nonresponse and the proportion of "don't know" answers in the experimental groups and control groups are compared, and in Section 5.3 consistency in answers between survey rounds is studied.

### *5.1 Party vote if election today*

In Table 6 we present estimates of proportions of votes for different parties if there was an election today for May 2015 and November 2015. Three estimates of proportion of votes are presented, using data from the full sample (Total), from the experimental group and the control group. Also the difference between the estimates using data from the experimental and control groups with an approximately 95 % confidence intervals (CI) for the differences are shown. In May 2015 the only significant difference that can be observed is for "Other" parties. In November 2015 the differences for the Liberal Party, the Sweden Democrats and for Other parties are significant on the 5 percent level.

**Table 6. Estimates of party votes if election today. May and November 2015. Percent/percentage points**

| | May 2015 | | | | | November 2015 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Party | Total | Experi-mental group | Control group | Difference (exp – control) | CI for difference | Total | Experi-mental group | Control group | Differen ce (exp – control) | CI for difference |
| Centre Party | 6.5 | 6.1 | 6.8 | -0.7 | (-2.0, 0.6) | 6.8 | 6.9 | 6.6 | 0.3 | (-1.3, 1.8) |
| Liberal Party | 4.6 | 4.6 | 4.6 | -0.1 | (-1.2, 1.0) | 5.5 | 5.0 | 6.8 | -1.8 | (-3.5, -0.2) |
| Moderate Party | 25.7 | 25.1 | 26.3 | -1.2 | (-3.6, 1.2) | 22.7 | 22.1 | 24.1 | -2.0 | (-4.9, 0.9) |
| Christian Democratic Party | 3.9 | 4.1 | 3.6 | 0.5 | (-0.6, 1.5) | 3.7 | 3.6 | 4.1 | -0.5 | (-1.8, 0.8) |
| Social Democratic Party | 29.4 | 29.4 | 29.4 | 0.0 | (-2.6, 2.5) | 27.4 | 27.4 | 27.5 | -0.2 | (-3.2, 2.8) |
| Left Party | 6.3 | 6.6 | 6.0 | 0.6 | (-0.7, 1.9) | 5.8 | 5.6 | 6.3 | -0.7 | (-2.3, 0.9 |
| Green Party | 6.5 | 6.0 | 7.0 | -1.0 | (-2.3, 0.4) | 5.8 | 5.9 | 5.5 | 0.5 | (-1.0, 1.9) |
| Sweden Democrats | 14.7 | 15.1 | 14.4 | 0.7 | (-1.5, 2.9) | 19.9 | 20.9 | 17.7 | 3.3 | (0.4, 6.1) |
| Other parties | 2.4 | 3.0 | 1.9 | 1.1 | (0.2, 1.9) | 2.3 | 2.7 | 1.5 | 1.2 | (0.3, 2.1) |

*5.2 Item nonresponse and "don't know"*

The answer "don't know" is a valid response option for those respondents who don't have a party preference or do not know which party to vote for. In the web questionnaire we chose not to explicitly offer it to the respondent, but to have the respondents typing in "do not know". Therefore it is also important to compare the share of don't knows and item nonresponse on these questions in the different groups. The results, presented in Table 7, shows no significant differences between the two groups in either of the experiments, neither regarding the share of don't knows, nor the share of item nonresponse.

**Table 7. Item nonresponse and "don´t know"-answers on question about party to vote for. September 2014, May 2015 and November 2015. Percent**

| | September 2014 | | May 2015 | | November 2015 | |
|---|---|---|---|---|---|---|
| | Experiment | Control | Experiment | Control | Experiment | Control |
| Don´t know | 18.0 | 18.5 | 10.8 | 13.7 | 12.7 | 17.1 |
| Item nonresponse | 1.9 | 1.8 | 1.2 | 1.6 | 0,9 | 1,4 |

*5.3 Consistency in answers between survey rounds*

In earlier sections we have compared estimates between the experimental group and control group. Since the experiments were carried out on a panel survey we also have the possibility to study the behavior of the same individuals in two consecutive survey rounds.

In Table 8 the share of respondents who answered the same party to the question about party vote in the most recent parliamentary election is presented. There are no significant differences between the two groups in either experiment.

**Table 8. Same answers to question about party vote in the most recent parliamentary election in September 2014, May 2015 and November 2015. Percent**

| | September 2014 | | | May 2015 | | | November 2015 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Exp. | Cont. | Total | Exp. | Cont. | Total | Exp. | Cont. | Total |
| Same answer in two consecutive survey rounds | 90.7 | 85.4 | 89.0 | 90.4 | 90.5 | 90.4 | 91.3 | 92,5 | 91.6 |

## 6. Conclusions

Three mixed-mode experiments have been carried out on the Swedish Party Preference Survey during the years 2014 and 2015 using mainly the same experimental design. All experiments show positive results regarding the possibility to increase response rates when adding web as an additional mode to a telephone survey. A positive result, regarding response rates, can be received regardless if information on login is sent out by e-mail or by postal mail. The result

also shows that the largest gain of respondents comes from the noncontacts in earlier survey rounds. However, the gain in response rate is unevenly distributed in the population. From the analysis of R-indicators it seems like the response set from the experimental group is less representative than the control group given the auxiliary variables that were used. Despite the fact that mixed-mode data collection result in slightly less balanced response set there are no signs that mixing the modes web with telephone introduce any severe mode effects. In designing the questionnaire for the web we chose to make it as similar as possible to the telephone version in the sense that it does not offer any response options initially to the open party question and both experiments show that the web questionnaire functions well and that there does not seem to be a decrease in data quality when using self-administrated mode and an open ended party question. Neither does mixed-mode seem to introduce any larger mode effects on the data. The analysis that has been carried out on the main variable of interest show similar party distributions in the two groups (experimental and control).

**References**

Glanzer, M. (1972). Storage mechanisms in recall. In G.H. Bower (Ed.). The psychology of learnings and motivation: Advances in research and theory (Vol. 5). New York: Academic Press.

Krosnick, J. and D. Alwin. 1987. "An Evaluation of a Cognitive Theory of Response Order Effects in Survey Measurement." Public Opinion Quarterly 52:526-38.

de Leeuw, E.D. (2008). Choosing the method of data collection. In (eds.) Hox, J.J., & Dillman, D.A. (2008). International handbook of survey methodology. NY: Lawrence Erlbaum Associates.

de Leeuw, E.D. (2005). To Mix or Not to Mix Data Collection Modes in Surveys. Journal of Official Statistics, Vol. 21, No. 2, 2005, pp. 233–255

Schuman, H., Presser, S. eds. (1981). Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording and Context. New York: Academic

Schwarz, N., Knäuper, B., Oyserman, D., & Stich, C. (2008). The psychology of asking questions. In E.D. de Leeuw, J.J. Hox, & D.A. Dillman (Eds.), International handbook of survey methodology (pp. 18–34). New York, NY: Lawrence Erlbaum Associates.

Shlomo, N., Skinner, C. and Schouten, B. (2012). Estimation of an indicator of the representativeness of survey response. Journal of Statistical Planning and Inference, 142 (1). pp. 201-211.