






# Designing an Experiment on Recognition of Political Fake News by Social Media Users: Factors of Dropout

Olessia Koltsova , Yadviga Sinyavskaya  , and Maxim Terpilovskii

Laboratory for Social and Cognitive Informatics, National Research University Higher School of Economics, Saint-Petersburg, Russia

ysinyavskaya@hse.ru

**Abstract.** Although social networking sites (SNS) offer functionalities for large-scale online research, user behavior and, in particular, scale and factors of their dropout from SNS-administered research have hardly been studied. In this paper we present an SNS-based experiment and survey tool and report the results of our investigation of user dropout from a research that uses this tool. This research is a pilot stage of a cross-country comparative study of political fake news recognition. At this stage Facebook and Vkontakte users from Russia have been recruited via SNS ad managing systems, asked to evaluate the truthfulness of the displayed news items and to answer a number of questions. We find that although we had to perform thousands of ad displays, among those who clicked the ad dropout rate was 60 and 65% in Vkontakte and Facebook respectively. 1,816 complete questionnaires were collected within a few days. More educated respondents, people living in or near megalopolises and those who agreed to grant access to their Vkontakte account data were significantly more inclined to complete the survey, but the major predictor of dropout was high individual speed – an indicator of low interest. Neither device type (mobile vs desktop) nor the number of questions per screen (one vs two) affected dropout. The number of leavers declined from the first to the last screens of our tool, but transition from the experiment to the survey and demographic questions produced clear peaks in the dropout curve.

**Keywords:** Online experiment · Online survey · Factors of dropout · Social media

## 1 Introduction

Proliferating social networking sites (SNS) open great opportunities for academic surveys and experiments in social sciences. SNS ad managing services allow for control of sample parameters, while the advertised data collection applications may employ a variety of experimental designs and gamification techniques. Thus, web-based research allows combining the collection of experimental data, self-reported data, and observational data from user accounts. However, they are also not free from the typical problems faced by researchers in “paper-and-pencil” surveys and laboratory experiments.

One of the most important tasks in online research (similar to offline non-response rate reduction) is dropout prevention.

Dropout is a respondent's premature termination of a survey session. Dropout rate is the share of respondents who started the questionnaire but abandoned it on different stages of completing, among all respondents [1].

The complete (unit nonresponse) or partial (item nonresponse) loss of data due to respondent's refusal to participate seems prominent methodological issue causing the *nonresponse* errors, i.e. respondents being significantly different from nonrespondents [2]. This, in turn, poses a threat for the reliability and generalizability of results gained via Web-based surveys.

The comparison of dropout values in offline and online research reveals the higher rates among the latter: it varies from 15% to 80% depending on the recruiting method in web-based studies in contrast to 5% termination rate in face to face or telephone surveys [3].

Such striking contrast in dropout rates can be explained by the lack of communication between researchers and respondents in online settings and, therefore, limited ability to establish the rapport, provide real time feedback to respondents' actions and increase participants' loyalty by all other means available in a classical laboratory setting.

Two separate strategies (or their combination) are usually implemented by researchers for reducing the dropout in web-based studies. The first one aims at increasing the motivation of respondent to participate by proposing incentives and awards, such as money or new knowledge. The empirical evidence on the effect of monetary rewards on both dropout and data quality is inconclusive [4–7] while the non-monetary motivation, such as *interest* in research topic, seem to lead to lower incidents of respondents leaving the study [8].

The second strategy implies manipulation of the features of the web-based tool so that to increase its usability and convenience for a respondent. Interface structure and design [1, 2, 9, 10], tool's stability across different platforms and devices [11]; questionnaire features, such as its length [3], type of questions [8, 12], their order, including the request of personal information in the beginning [7]; the way the questions are displayed on the screen [3, 5, 13] – all have been reported as substantial factors of dropout in web-based research.

Decision field theory [14] provides the perspective for explaining the survey behavior by combining these two strategies. The decision to continue or to abort a survey session depends on the ratio between the strength of factors increasing or decreasing respondent's motivation.

Research on dropout factors is of special importance for politically sensitive research topics, such as our experiment on fake political news recognition which also offers a number of sensitive political questions. This goal, first, is a competing task with ethical data collection based on informed consent, second, it contradicts the goal of research thoroughness, and third, it can stumble into dropout factors that are dramatically different from what researchers know from their offline experience.

## 1.1 Objectives

In this study, we investigate respondents' behavior in a self-administered online experiment on recognition of political fake news. First, we compare the dynamics of dropout in two large SNSs – Facebook and Vkontakte – most popular in the studied country (Russia). Second, we examine what are some of the factors that may affect users' decision to leave the experiment. The first important group of factors addressed in the study is the properties of respondents, such as socio-demographic features, privacy preferences or perception of personal performance during the experiment. The other group is the effects of web interface, which includes screen type and the number of questions per screen among others. This study is a pilot stage of our future comparative cross-country research on factors of fake news recognition.

## 2 Research Design

In this section we briefly describe the design of our future research on fake news whose instrument we are testing in this paper and the approach to its testing. The goal of the future research is to experimentally find out factors influencing Facebook and Vkontakte users' ability to recognize fake news in a situation with or without international tensions. The research is supposed to be carried in three countries: Russia, Kazakhstan and Ukraine. Both in the future research and in the current pilot research we employ  $2 \times 2 \times 2$  experimental design in which each SNS user receives eight news randomly retrieved from our database with varying truthfulness (true/false), news source (from user's country/from the country covered in the news) and news frame (dominant/alternative). A user then receives two questions on whether he/she has checked any of the news or has seen any of them before, six demographic questions, three questions on news consumption, three questions on generalized trust, four questions on conspiracy thinking, and three on political interest and attitudes (5-point Likert scales). Questions are organized in 14 screens, 1–2 questions per screen. They are followed by the user's score, a half-serious comment on her/his fake-detection abilities and an offer to learn correct answers after the end of the experiment. For Vkontakte users, the instrument is presented as an application that also requests the data from user accounts, namely friend lists and group lists (not required for participation). For Facebook users, it is presented as a stand-alone website that does not collect any account data (due to restrictions of Facebook). Since Vkontakte is much more popular than FB in Russia and in Kazakhstan, but not in Ukraine, one of the goals of our instrument testing is to compare dropout in both social networks. It is also important to understand that VK terms of use allow third parties to download certain user data from accounts not protected by privacy settings without special requests. These data may include gender, age, city, views and interests and any other data from the "about" section that the user chooses not to hide. These data are also used in this research.

Instrument testing began from technical tests by the Lab members who revealed several dozens of bugs and glitches; at this stage we also refused from any screens that involved scrolling since it obviously irritated all participants, and from drop-out menus since they obviously incentivized participants to choose the first option. An earlier

offline pre-test on 100 students that compared convenience of scales and binary questions revealed no significant difference, which is why we opted for scales as more informative options.

In the present research we defined our populations as all FB and VK users from Russia, which is why we randomly targeted these groups in both networks using respective ad management systems. As we estimate our final VK sample size in Russia as 3–5 thousands, we took 400 (approx. 10%) as the target size for our pilot research, and the size of FB sample twice as small, according to its smaller penetration in Russia. After having collected half of both samples we analyzed the results and permuted the order of one of the questions in each of the instrument versions to test two additional hypotheses that emerged after the interim data analysis. All the pilot experiments were carried out in October–December 2019.

### 3 Assumptions and Hypotheses

In this study we test several statistical hypotheses, as well as a series of less rigorously formulated assumptions that are nonetheless important for calibrating our instrument. In formulating them, we based on the literature on survey non-response and dropout, on user churn in gaming and human-computer interaction. We begin with the description of two most general expectations.

First, we expect that *the pattern of dropout in our study will be similar to those detected in previous works on churn dynamic* [3]. The power-law distributed curve with only a small loyal minority of “survivors” is typical for both online game and app users [15, 16] and is expected to be unavoidable. Therefore, fluctuations can be detected correctly only against this general trend.

Second, since the online experiment as a method opens a wide range of opportunities for increasing the response rate by means of interactive stimuli, gamification of a process, presence of graphic elements in the design etc., we suppose that combination of an experimental and a more standardized survey sections in one research may provoke a dropout of the respondents. Thus, *while mixing the entertaining experimental part with the more standard survey-part (like in the present study), the highest number of dropouts will occur at the moment of transition of a respondent from experimental to survey part.*

More specifically, we focused around two lines of inquiry, the first relating to the features of respondents and the second relating to the features of data gathering process. While the second line obviously aims to reveal ways to improve our instrument, the first one is aimed to uncover potential sample biases that should be taken into account during the future research.

#### 3.1 Respondents’ Features as a Factor of Dropout

**Socio-demographic Factors.** Age of a user might have a two-way impact on a person’s behavior during online survey. On the one hand, the interaction with the online survey application requires technical skills which may be less common among so-called “digital immigrants” who encountered new digital technologies being middle-aged [17]. On the other hand, as Galesic [3] has shown, older respondents are more cooperative, with the

probability of dropout decreasing by around 1% with the each additional year of age. Taking into account these two lines, we hypothesize that:

*H1a: Dropout rate will increase with users' age.*

*H1b: Dropout rate will decrease with users' age.*

Along with the age, *gender differences* in survey behavior may occur because of differences in online behavior, such as frequency of SNS use [18], purposes of Internet usage [19] and levels of online self-disclosure [20, 21]. In addition, some scholars nominated women as more cooperative respondents [22] with higher levels of interest in survey process [3]. In consistence with these findings, we expect that:

*H2: Dropout rate will be higher among male than among female respondents.*

Finally, following the argument of Galesic [3] that lower level of education might lead to additional difficulties during surveys, we hypothesize that:

*H3: Dropout rate will be higher among respondents with lower education.*

**Individual Performance as a Factor of Dropout.** Respondent's perception of her own performance in an experimental task might be a crucial factor of attrition as it has been shown to happen in case of online learning [23, 24] and in online game contexts [16, 25]. Thus, we suppose that high individual speed of performance in the experiment may serve as a proxy for the low level of difficulty and, consequently, for high level of satisfaction that should incentivize a respondent to move further. At the same time, high speed might also indicate low quality of answers that has been found to be related to the lack of interest, while the latter, in turn, has been found to increase dropout [3]. Therefore:

*H4a: The higher the personal speed the lower the probability of dropout.*

*H4b: The lower the personal speed the lower the probability of dropout.*

**Privacy and Online Activity as a Factor of Dropout.** As functional complexity of social media and its popularity grows, privacy issues are becoming increasingly salient. Recruitment of respondents on social networking sites (SNSs) as the territories often perceive as private, thus, may be hindered by users' privacy concerns. In this context requests for personal information, inherent to survey process, are additionally increasing dropout rate [7, 26].

Considering the complex relationship between privacy concerns and the level of online user self-disclosure [27], we aim at testing a range of hypotheses on how privacy-related requests to the respondents along with the presence of sensitive questions in the study, might affect respondents' decisions to leave the online experiment. Therefore, we expect that:

*H5: Dropout rate will be higher among those who denied access to any of their data (Vkontakte only).*

If this is the case *and* those who have given access are substantially different from those who have not, this may result in serious sample biases. Therefore, it is important to learn whether the second condition holds:

*H6: Age (a), gender (b) and education (c) of users who have given access to their private data will be significantly different from those who have not (Vkontakte only).*

Next, we assume that respondents' privacy-related anxiety might increase when they are presented questions about their political views. In non-democratic regimes such views are often perceived as threatening individuals' security, in case their identity is disclosed, while such disclosure may, in turn, be perceived as probable due to low trust to social institutions.

*H7: The two sensitive questions (a) on political loyalty towards the government and (b) on political relations with a neighboring country will be skipped by the majority of users.*

**Political Engagement.** Since the topic of the experiment is highly politicized, politically engaged individuals might have more incentives to complete the questionnaire. In non-democratic regimes grassroots political engagement is usually of contentious or even protest nature. Since the levels of protest across Russia are very uneven, with Moscow and St. Petersburg taking the lead [28], we expect regional differences in dropout:

*H8a: Dropout rate among users from Moscow and Saint Petersburg will be significantly lower than among users from other regions.*

*H8b: Dropout rate among users from Moscow, Moscow region, Saint Petersburg, or Leningrad region will be significantly lower than users from other regions.*

### 3.2 Features of Data Collection Tool as a Factor of Dropout

The second group of the investigated factors is patterns of the visual organization which might affect the dropout rate of respondents [29].

**SNS and Interface Type.** In our research we had to employ different web interfaces for different SNSs – an app for VK and a stand-alone website for FB. This suboptimal choice was made because, on the hand, we aim at analyzing both experimental data and the data retrieved from user accounts, where possible (Vkontakte), but, on the other hand, we cannot exclude Facebook. Although we tried to make both interfaces as similar as possible, they are not graphically identical, which is why we cannot isolate the effect of the SNS from the effect of web-interface. However, we assume that one or both should matter. On the one hand, the fact that the website does not request personal data may contribute to dropout prevention. On the other hand, redirection to an unknown external website with minimalistic layout, as opposed to staying on the familiar SNS, may work in the opposite direction. Therefore, we propose two alternative hypotheses:

*H9a: Dropout rate will be higher among the users of Vkontakte as compared to Facebook.*

*H9b: Dropout rate will be lower among the users of Vkontakte as compared to Facebook.*

**Number of Questions on the Screen.** Previous research on this factor is inconclusive. On the one hand, several studies carried out on different samples have established that single-item-per-screen design leads to a slightly smaller dropout rate than scroll design (in which the entire questionnaire is presented on a single online page) [30, 31]. On the other hand, this contradicts the earlier results of Couper et al. [32] who showed that the multi-item-per-screen design requires significantly less time to complete and produces lower data loss than single-item-per-screen version. Here, we find it important to separate two factors: number of questions per screen and the necessity to scroll. Therefore, we have designed and compared only screens with one or two questions fully visible without scrolling on any device:

*H10: The probability of dropout will be higher on the pages that contain more than one question.*

**Screen Size as a Factor of Dropout.** Wenz [11] has found that smaller screens increase the probability of respondents' dropout from web-surveys. However, a major reason for that may be that he has tested this hypothesis on a web-interface that was not adapted to different devices, while our tool has been adapted. Wenz's research thus does not differentiate between screen size per se and other factors, such as font size or the necessity to zoom and scroll in different directions. With our device-adaptable instrument, we can isolate and test the screen size effect, to the extent to which we have information on it (mobile vs desktop):

*H11: The probability of dropout among mobile device users will be higher than that among desktop users.*

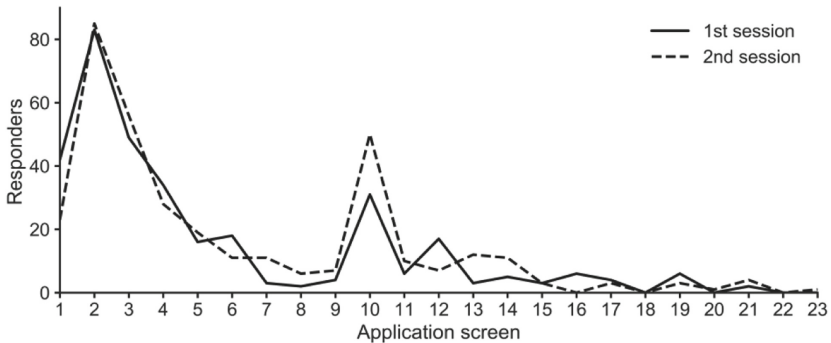
## 4 Results

### 4.1 Descriptive Analysis

**Vkontakte Campaign.** Vkontakte recruitment campaign lasted for three days (28–30 Oct 2019). The ad was displayed 198,609 times to the news feeds of 179,051 unique users (data provided by the platform ad managing service). Of them, 1,138 users clicked on the ad and 453 completed both the experimental and survey parts. The campaign was divided into two sessions, with roughly equal numbers of users in each (231 and 222 respectively). In total, 61.8% of responders were male, 18.9% were female, and 19.3% chose not to specify. The respondents aged from 18 to 55+ years (median = 19.0, IQR = 6.0). As expected, mobile device users dominated, amounting to 85.8% of all users (N = 976).

Vkontakte dropout rate constitutes 60.2% (685 users): 9.5% users launched the application by clicking on the ad and left without proceeding further (screen 1), 63.1% users left at the experimental stage (screens 2–9), and 27.4% users left at the survey stage (screens 10–23). As it can be seen from Fig. 1, the dropout decreases from the second to

the last screen, but peaks on screen 10. This was observed already after the first session when we stopped recruitment for express data analysis, however, it was not clear whether this peak was attributable to the shift between the experiment and the survey or to the content of the question displayed on screen 10. To isolate the former effect from the latter, we swapped the questions displayed on screens 10 and 15 (picking up those that would not substantially break the logic of the questionnaire). As can be seen in Fig. 1, user behavior did not change fundamentally after that.



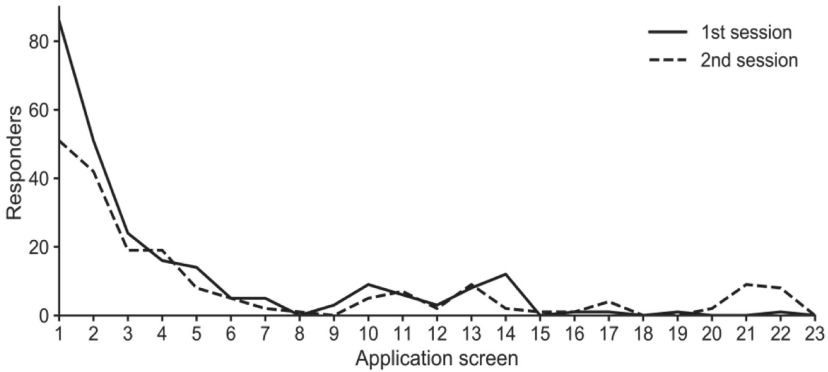
**Fig. 1.** Number of Vkontakte users who left the application on different screens (user dropout).

**Facebook Campaign.** Facebook recruitment campaign lasted for three days as well (2–4 Dec 2019) resulting in 6,644 displays of the ad. 678 participants clicked on it of whom 235 completed both the experimental and the survey parts. In total, 34.8% of respondents were male, 52.5% were female, and 12.7% chose not to specify. Mobile device users constituted 63.9% of all users ( $N = 433$ ). The respondents aged from 18 to 55+ years, although the distribution of respondents’ ages dramatically differed from what we observed in Vkontakte campaign (median = 52.0, IQR = 16.0). Furthermore, judging from the data on age distributions of overall populations of both VK and FB provided by the respective ad managing services, both samples also diverge from the populations, but in the opposite directions. While our VK sample age distribution was skewed in favor of younger groups, as compared to the overall VK audience age distribution, our FB sample was, on the contrary, older than the overall FB audience (see Table. 1). This indicates highly non-random character of ad displays across the selected audiences either on one of on both SNSs.

Facebook dropout rate was equal to 65.3% (443 users): 30.9% users ran the application but did not participate (left on screen 1), 48.3% users left the application at the experimental stage (screens 2–9), 20.8% users left the application at the survey stage (screens 10–23). Maximum dropout rate was observed at the starting screen (screen 1), while on VK it was on screen 2 (first news item).



After the first session—with 114 users completing both the experiment and the survey part—we analyzed the obtained results and found out that the FB dropout curve, just like VK curve, in fact had a second smaller peak, although not around the politically sensitive questions, as expected, but around the two screens with four socio-demographic questions. When moved to the end of the questionnaire in the second session, these questions still produced a peak (screens 21–22) (Fig. 2).



**Fig. 2.** Number of Facebook users who left our application on different screens (user dropout).

**Overall Data.** To analyze the overall data, we combined Vkontakte and Facebook datasets. A total of 1,816 users participated in our pilot trials. Out of 767 responses with specified gender data, there were 52.7% male respondents (n = 404), 30.2% female (n = 232), and 17.1% chose not to specify (n = 131). The participants had a median age of 23.0 (IQR = 33.0) and came from 76 regions of Russia. They mostly used mobile devices (77.6%, n = 1409). The respondents reported the following education levels: 21% finished secondary school or college, 15.8% were studying at the university, 31.2% already received a bachelor or master’s degree, 4.1% had a doctor’s degree, 9.2% chose not to specify. Interestingly, 18.7% of respondents claimed to be studying at the secondary school, although we had limited the age of users to 18+ in ad campaigns.

The overall dropout rate was equal to 62.1% (n = 1128): 17.9% left at the starting screen (screen 1) and did not enter the experimental stage, 57.3% left at the experimental stage (screens 2–9), 24.8% left at the survey stage (screens 10–23). Approximately 37.9% respondents finished both experimental and survey stages (n = 688).

**Table 1.** The age distribution of Vkontakte and Facebook SNSs users.

SNS	Feature	Category	Respondents		Total audience
			N	Percentage	Percentage
Vkontakte	Age	18–24	379	75.2	27.9
		25–34	34	6.7	32.0
		35–44	16	3.2	14.4
		45–54	18	3.6	6.0
		55+	57	11.3	6.7
Facebook	Age	18–24	9	3.5	6.7
		25–34	15	5.8	27.2
		35–44	55	21.2	32.3
		45–54	65	25.1	18.6
		55+	115	44.4	14.8

**4.2 Hypotheses Testing**

To test our hypotheses, we ran a series of logistic regressions and  $\chi^2$  tests. The brief summary of hypotheses testing is given in Table 2.

**User Features as a Factor of Dropout.** Here, we start with demographic feature of our users.

*H1: Dropout rate will (a) or (b) increase with users’ age.*

The analysis based on logistic regression model did not reveal any association between dropout rate and users’ age on Facebook and Vkontakte datasets ( $\beta = 0.0037$ ,  $p = 0.456$ ). Thus, the hypotheses *H1a* and *H1b* were rejected. Additional analysis showed that the distribution of dropout rate by age resembled a quadratic function, however, this resemblance was not statistically significant either.

*H2: Dropout rate will be higher among male than among female respondents.*

In addition to gender data provided at the survey stage, we had previously obtained a random sample of 1.2 million open Vkontakte accounts including gender information (specified in user profiles). Gender data agreement between survey and user profile data subsets was estimated at 77.0%. We tested the hypothesis that dropout would be higher among women than men on both self-reported and observed data from respondents’ profiles in VK. In both cases no statistically significant difference was revealed (self-reported data:  $\chi^2 = 0.09$ ,  $p\text{-value} = 0.38$ ; VK data:  $\chi^2 = 0.3$ ,  $p\text{-value} = 0.7$ ). In this relation, the hypothesis H2 was rejected.

*H3: Dropout rate will be higher among respondents with lower education.*

The hypotheses did not reveal any association between dropout and user's level of education when tested separately on Facebook and Vkontakte datasets. At the same time, the overall data support the hypothesis *H3*, suggesting dropout dependency on user education. We used logistic regression to gain insight into this dependency pattern. Model fitting showed that education level was negatively associated with user dropout ( $\beta = -0.22$ ,  $p = 0.014$ ): a lower level of a user's education is associated with a higher probability of dropout. Thus, the hypothesis *H3* was supported.

*H4a: The higher the personal speed the lower the probability of dropout.*

*H4b: The lower the personal speed the lower the probability of dropout.*

Our tool recorded the time at which each user entered and left the experimental stage (screens 2–9). Personal speed was calculated as the mean time a respondent spent per experimental item, and this time was 6.3 times smaller among those who left the survey. We applied logistic regressions for each SNS sample and for the joint dataset and confirmed that the more time user spends at the experimental stage, the less is the probability of user dropout ( $\beta = -0.015$ ,  $p = 0.033$  for Facebook;  $\beta = -0.049$  and  $-0.036$ ,  $p = 5.5e-12$  and  $3.2e-11$  for Vkontakte and overall data). Thus, we accepted hypothesis 4b. Additional exploratory analysis showed that the distributions of speeds among those who left and those who stayed were a similarly left-skewed bell-shaped curve which means that there was no latent group of slow performers among the leavers.

*H5: Dropout rate will be higher among those who denied access to any of their data (Vkontakte only).*

As our application requested access to the private user data in Vkontakte, but not in Facebook, hypotheses 6 and 7 were tested on VK subsamples only. Dropout rate among users who provided full access to their private data was equal to 51.8% ( $n = 397$  of 582), but it was higher by 16.4% (68.2%,  $n = 288$  of 556) among users who denied access to the data. Based on the results of a chi-squared test, the hypothesis *H5* was accepted ( $\chi^2 = 31.29$ ,  $p = 2.2e-08$ ). Thus, dropout rate was significantly higher among the users who did not provide their private data.

*H6: Age (a), gender (b) and education (c) of users who have given access to their private data will be significantly different from those who did not (Vkontakte only).*

Access to the requested private data (friend list and group list), was provided by approximately a half of Vkontakte users (49.5%,  $n = 556$ ). However, all VK respondents who completed the survey indicated their age, gender, education and region, which allows us to compare access “deniers” and “providers” by these features. Statistically significant difference was obtained for user age as we compared the age distributions of both user groups with the two-sample Kolmogorov-Smirnov test ( $D = 0.26$ ,  $p = 1.6e-07$ ). In addition, logistic regression model revealed the negative relation between user age and user willingness to provide full access to their private data in Vkontakte ( $\beta = -0.013$ ,

$p = 0.00086$ ). Gender and education were not significantly associated with user intent to provide the private data ( $\chi^2 = 2.73$ ,  $p = 0.26$  and  $\chi^2 = 11$ ,  $p = 0.08$  respectively). Thus, hypothesis 6 was partly supported.

*H7: The two sensitive questions (a) on political loyalty towards the government and (b) on political relations with a neighboring country will be skipped by the majority of users.*

Hypothesis 7 was rejected as only 118 users of 705 skipped the question on the political loyalty ( $\chi^2 = 310.67$ ,  $p = 1.0$ ) and 123 users of 708 chose not to answer the question about the political relations ( $\chi^2 = 300.17$ ,  $p = 1.0$ ).

**Political Engagement.** The regional group of Moscow and Saint Petersburg among our respondents was found to be significantly underrepresented as compared with general Vkontakte population, as determined by our random sample of 1.2 million users ( $\chi^2 = 230.64$ ,  $p < 2.2e-16$ ). This actually contradicted our expectations as we had assumed that inhabitants of both cities, as more politicized, would be more interested in the topic of the experiment which had also lead us to hypothesis 8.

*H8a: Dropout rate among users from Moscow and Saint Petersburg will be significantly lower than among users from other regions.*

*H8b: Dropout rate among users from Moscow, Moscow region, Saint Petersburg, or Leningrad region will be significantly lower than users from other regions.*

Logistic regression analysis has shown that, indeed, users of all SNSs coming from Moscow, Saint Petersburg, and the corresponding surrounding regions finish both stages of the experiment with higher probability than users from other regions (Facebook:  $\beta = 5.39$ ,  $p = 9.3e-08$ ; Vkontakte:  $\beta = 2.88$ ,  $p = 2.0e-06$ ; overall data:  $\beta = 3.99$ ,  $p = 5.9e-15$ ). At the same time, this hypothesis was rejected for users coming from just Moscow and Saint Petersburg (Facebook:  $\beta = 18.49$ ,  $p = 0.97$ ; Vkontakte:  $\beta = 16.02$ ,  $p = 0.96$ ; overall data:  $\beta = 17.18$ ,  $p = 0.95$ ). Thus, only the hypothesis H8b was supported.

**Features of Data Collection Tool as a Factor of Dropout.** In this section we report the factors that theoretically can be manipulated.

*H9: Dropout rate will be (a) higher or (b) lower among the users of Vkontakte as compared to Facebook*

The tests of hypotheses H9 revealed the association between user dropout and SNS, with Vkontakte dropout rate being significantly lower than that of Facebook ( $\chi^2 = 4.47$ ,  $p = 0.017$ ). Thus, the hypothesis H9b was supported.

*H10: The probability of dropout will be higher on the pages that contain more than one question.*

We further explored how the number of questions on the screen might be related to the users’ dropout. Logistic regression model failed to establish any significant association between these variables ( $p = 0.987$ ). Thus, the hypothesis H10 was rejected.

*H11: The probability of dropout among mobile device users will be higher than that among desktop users.*

Our instrument automatically collected user device data (operational system, browser, device type) that allowed us to estimate the mobile device percentage. Based

**Table 2.** Hypotheses testing results

Hypotheses: Predictor; association with dropout rate or probability (except H6, H7)	Status		
	Facebook	Vkontakte	Overall
H1a: age; positive. H1b: age; negative	Rejected Rejected	Rejected Rejected	Rejected Rejected
H2: gender (male), positive.	-	Rejected	-
H3: education level; negative	Rejected	Rejected	<b>Accepted</b>
H4a: Speed; negative H4b: Speed, positive	Rejected <b>Accepted</b>	Rejected <b>Accepted</b>	Rejected <b>Accepted</b>
H5: Access denial; positive	—	<b>Accepted</b>	—
H6: Users who have given access to their private data will be significantly different from those who did not, by: <ul style="list-style-type: none"> <li>- Age (young)</li> <li>- Gender (any)</li> <li>- Education (any)</li> </ul>	—	<b>Accepted</b> Rejected Rejected	—
H7: The two sensitive questions will be skipped by the most users.	Rejected	Rejected	Rejected
H8a: region (Moscow & St. Petersburg); negative H8b: region (Moscow, St. Petersburg & surroundings); negative	Rejected <b>Accepted</b>	Rejected <b>Accepted</b>	Rejected <b>Accepted</b>
H9a: SNS (Vkontakte); positive H9b: SNS (Vkontakte); negative	-	-	Rejected <b>Accepted</b>
H10: N of Qs per page; positive	Rejected	Rejected	Rejected
H11: Device (mobile); positive	Rejected	Rejected	Rejected

on all SNS data and the overall dataset, mobile devices dominated over desktop (Facebook:  $\chi^2 = 51.58$ ,  $p < 3.4e-13$ , Vkontakte:  $\chi^2 = 580.82$ ,  $p < 2.2e-16$ , overall data:  $\chi^2 = 551.76$ ,  $p < 2.2e-16$ ). The highest percentage of mobile devices was observed among Vkontakte users (85.8%). However, the analysis did not reveal any association between dropout and the type of device which user used for completing the survey on either Facebook or Vkontakte datasets (Facebook:  $\chi^2 = 0.005$ ,  $p = 0.528$ , Vkontakte:  $\chi^2 = 5.08e-30$ ,  $p = 0.5$ ). Thus, the hypothesis H11 was rejected.

## 5 Discussion

The results of this research have several important implications for dropout reduction and for a broader methodology of online experiments and surveys carried out via social networking sites.

First, although we targeted the entire Russian populations of both VK and FB, the resulting samples had significant differences from those populations in a number of demographic characteristics, the first of which is age. If age bias were the same in both networks, one could assume more interest in the experiment topic among certain age groups. However, the biases were both very strong and opposite in the two SNSs which suggests that ad managing systems employ both highly non-random and non-transparent targeting strategies. Therefore, it is highly recommended that researchers control demographic proportions of their samples while administrating SNS-based surveys. Under-representation of Moscow and St. Petersburg in our VK sample most probably has a different cause. Here we used our own data on population distribution (the only available) dating back to 2015. Since then the penetration of VK beyond big cities might have grown. Also, these two cities are likely to host the largest share of business and spam accounts that might have naturally self-deselected from our research.

Second, we observed a slightly lower level of dropout in VK than in FB (60.2% vs 65.3%), although the latter SNS, being more politicized, could be expected to contain more users interested in the topic of our experiment. A possible explanation might be derived from the fact that the largest number of dropouts among FB users happened on the very first page of the website, while among VK users this maximum happened on the next page containing the first news item. This suggests that additional losses from FB might be explained by the website design or by the fact of it being stand-alone; however, this hypothesis should be tested by isolation of this effect from that of SNS.

Third, the dropout curves were similar in SNSs declining from the beginning to the end and produced two peaks: at the transition from the experimental part to the survey and on the screens with socio-demographic questions. While the first peak cannot be avoided, the sensitive questions can be shifted to the end. This may not increase the share of users willing to answer them, but it will increase the share of surveys complete in all other respects. It is also interesting that political questions were not perceived as sensitive by respondents, even if they were asked after socio-demographic data had been obtained.

Fourth, although we were afraid that mobile device users would be either distracted from the experiment by their environments or would have difficulties reading from the smaller screens, in fact, dropout among them was the same as among desktop users. The

overall high prevalence of mobiles over desktops may suggest that by now users have been already well adapted to small screens and to noisy and changing environments, and that mobiles can be used for research along with desktops. Likewise, once all the content of each page could fit the screen and no scrolling was needed, the number of questions per screen did not influence dropout.

Fifth, we have found out that individual speed or time per item is by no means an indicator of perceived difficulty; instead, high speed reliably indicates lack of interest and motivation. It is by far the best predictor of dropout.

Sixth, demographic features of respondents had different effects on their dropout. Gender was definitely unimportant, while people with higher education were visibly more inclined to complete our survey. This might be explained by its topic (politics), as well as by the fact that it appeared as an intellectual challenge. Influence of age demands further research. Although we find that dropout was higher among both the oldest and the youngest groups, this trend has not only been statistically confirmed, and, furthermore, may ultimately be some artifact of targeting which has resulted in overrepresentation of these two age groups in our FB and VK samples respectively.

Finally, as those users who denied access to their data were more likely to drop out, and simultaneously were more likely to be older (as judged from the age of those “deniers” who reached the stage of indicating their age), older age groups might be systematically underrepresented in our VK sample. This is consistent with the direction of age bias of our sample as compared to the overall VK population and may serve as an additional explanation for it, along with the features of VK ad managing service algorithm.

Overall, we can conclude that SNS-administered online experiments are quite efficient, fast and relatively inexpensive instruments, although they are not without limitations. Despite low rates of conversion of ad displays into clicks and fairly high dropout rates, such instruments can quickly collect large amounts of completed surveys. The highest caution should be exercised in relation to sample construction and recruitment. The biggest challenge for social scientists and psychologists is, however, task formulation for software developers and usability testing.

**Acknowledgements.** The research was implemented in the framework of the Russian Scientific Fund Grant № 19-18-00206 (2019–2021) at the National Research University Higher School of Economics.

## References

1. Vicente, P., Reis, E.: Using questionnaire design to fight nonresponse bias in web surveys. *Soc Sci Comput Rev* **28**(2), 251–267 (2010). <https://doi.org/10.1177/0894439309340751>
2. Dillman, D., Tortora, R.L., Conradt, J., Bowker, D.: Influence of plain vs. fancy design on response rates for Web surveys. In: *Proceedings of the Joint Statistical Meetings*. American Statistical Association, Alexandria (1998)
3. Galesic, M.: Dropouts on the web: effects of interest and burden experienced during an online survey. *J. Off. Stat.* **22**(2), 313 (2006)
4. Deci, E.L.: *Intrinsic Motivation*. Plenum Press, New York (1975)

5. O'Neil, K.M., Penrod, S.D., Bornstein, B.H.: Web-based research: methodological variables' effects on dropout and sample characteristics. *Behav. Res. Methods Instrum. Comput.* **35**(2), 217–226 (2003). <https://doi.org/10.3758/bf03202544>
6. Bosnjak, M., Tuten, T.L.: Prepaid and promised incentives in web surveys—an experiment. *Soc. Sci. Comput. Rev.* **21**, 208–217 (2003). <https://doi.org/10.1177/0894439303021002006>
7. Frick, A., Bächtiger, M.T., Reips, U.D.: Financial incentives, personal information and drop-out rate in online studies. In: Reips, U.D., Bosnjak, M. (eds.) *Dimensions of Internet Science*, pp 209–219. Pabst Science Publishers, Lengerich (1999). <https://doi.org/10.5167/uzh-19758>
8. Knapp, F., Heidingsfelder, M.: Drop-out analysis: effects of the survey design. In: Reips, U.D., Bosnjak, M. (eds.) *Dimensions of Internet Science*, pp. 221–230. Pabst Science Publishers, Lengerich (1999)
9. Healey, B.: Drop downs and scroll mice: the effect of response option format and input mechanism employed on data quality in web surveys. *Soc. Sci. Comput. Rev.* **25**(1), 111–128 (2007). <https://doi.org/10.1177/0894439306293888>
10. Couper, M.P., Traugott, M.W., Lamias, M.J.: Web survey design and administration. *Public Opin. Q.* **65**(2), 230–253 (2001). <https://doi.org/10.1086/322199>
11. Wenz, A.: Completing web surveys on mobile devices: does screen size affect data quality? (No. 2017-05). ISER Working Paper Series (2017)
12. Heerwegh, D., Loosveldt, G.: An evaluation of the effect of response formats on data quality in web surveys. *Soc. Sci. Comput. Rev.* **20**(4), 471–484 (2002). <https://doi.org/10.1177/089443902237323>
13. Crawford, S.D., Couper, M.P., Lamias, M.J.: Web surveys: perceptions of burden. *Soc. Sci. Comput. Rev.* **19**(2), 146–162 (2001). <https://doi.org/10.1177/089443930101900202>
14. Busemeyer, J.R., Townsend, J.T.: Decision field theory: a dynamic cognition approach to decision making. *Psychol. Rev.* **100**, 432–459 (1993). <https://doi.org/10.1037//0033-295x.100.3.432>
15. Perriñez, Á., Saas, A., Guitart, A., Magne, C.: Churn prediction in mobile social games: towards a complete assessment using survival ensembles. In: *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pp 564–573. IEEE Press (2016). <https://doi.org/10.1109/dsaa.2016.84>
16. Hadji, F., Sifa, R., Drachen, A., Thureau, C., Kersting, K., Bauchhage, C.: Predicting player churn in the wild. In: *IEEE Conference on Computational Intelligence and Games*, pp 1–8. IEEE Press (2014). <https://doi.org/10.1109/cig.2014.6932876>
17. Prensky, M.: Digital natives, digital immigrants. *On Horiz.* **9**(5), 1–6 (2001). <https://doi.org/10.1108/10748120110424816>
18. Correa, T., Hinsley, A.W., de Zúñiga, H.G.: Who interacts on the Web?: The intersection of users' personality and social media use. *Comput. Hum. Behav.* **26**(2), 247–253 (2010). <https://doi.org/10.1016/j.chb.2009.09.003>
19. Muscanell, N.L., Guadagno, R.E.: Make new friends or keep the old: gender and personality differences in social networking use. *Comput. Hum. Behav.* **28**(1), 107–112 (2012). <https://doi.org/10.1016/j.chb.2011.08.016>
20. Fogel, J., Nehmad, E.: Internet social network communities: risk taking, trust, and privacy concerns. *Comput. Hum. Behav.* **25**, 153–160 (2009). <https://doi.org/10.1016/j.chb.2008.08.006>
21. Raacke, J., Bonds-Raacke, J.: MySpace and Facebook: applying the uses and gratifications theory to exploring friend-networking sites. *Cyberpsychol. Behav.* **11**(2), 169–174 (2008). <https://doi.org/10.1089/cpb.2007.0056>
22. Groves, R.M., Couper, M.P.: *Nonresponse in Household Interview Surveys*. Wiley, New York (1998)



23. Chyung, S.Y.: Systematic and systemic approaches to reducing attrition rates in online higher education. *Am. J. Distance Educ.* **15**(3), 36–50 (2001). <https://doi.org/10.1080/08923640109527092>
24. Terrell, R.S.: A longitudinal investigation of the effect of information perception and focus on attrition in online learning environments. *Internet High. Educ.* **8**(3), 213–219 (2005). <https://doi.org/10.1016/j.iheduc.2005.06.003>
25. Borbora, Z., Srivastava, J., Hsu, K.W., Williams, D.: Churn prediction in MMORPGs using player motivation theories and an ensemble approach. In: *IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, pp 157–164. IEEE Press (2011). <https://doi.org/10.1109/passat/socialcom.2011.122>
26. Heerwegh, D., Loosveldt, G.: An evaluation of the semi-automatic login procedure. *Soc. Sci. Comput. Rev.* **21**, 223–234 (2003). <https://doi.org/10.1177/0894439303021002008>
27. Kokolakis, S.: Privacy attitudes and privacy behavior: a review of current research on the privacy paradox phenomenon. *Comput. Secur.* **64**, 122–134 (2017). <https://doi.org/10.1016/j.cose.2015.07.002>
28. Information Agency Regnum. <https://regnum.ru/news/polit/2588639.html>
29. Ganassali, S.: The influence of the design of web survey questionnaires on the quality of responses. *Surv. Res. Methods* **2**, 21–32 (2008). <https://doi.org/10.18148/srm/2008.v2i1.598>
30. Manfreda, L.K., Batagelj, Z., Vehovar, V.: Design of web survey questionnaires: three basic experiments. *J. Comput. Mediat. Commun.* **7**, 3 (2002). <https://doi.org/10.1111/j.1083-6101.2002.tb00149.x>
31. Peytchev, A., Couper, M., McCabe, S., Crawford, S.: Web survey design: paging versus scrolling. *Public Opin. Q.* **70**, 596–607 (2006). <https://doi.org/10.1093/poq/nfl028>
32. Couper, M.P., Traugott, M., Lamias, M.: Effective survey administration on the Web. In: *Midwest Association for Public Opinion Research Conference*, Chicago, Illinois (1999)