Web Surveys Among Children and Adolescents: Is There a Gamification Effect?

Aigul Mavletova[1]

Abstract

Several studies have measured a gamification effect in the surveys among adults. However, no experiments have been published with a focus on younger respondents. In this article, data quality between three conditions is compared among children and adolescents 7–15 years old as follows: (1) a text-only survey, (2) a visual survey with an attractive design and images, and (3) a gamified survey. To test a gamification effect, an experiment using a volunteer online access panel in Russia was conducted among 1,050 children. The gamified survey produced completion time more than a third longer than the text-only survey. A higher overall item nonresponse rate was found in both the gamified and visual surveys. However, this was mainly due to the Flash-based questions in these conditions. Fewer respondents straight-lined and used middle responses in the gamified and visual surveys. It was also less burdensome to complete the gamified survey. Children requested help to answer survey questions less often. They found it more enjoyable and easier. Moreover, the subjective evaluation of the completion time was not different from the two other conditions. Overall, we suggest that a gamification effect in web surveys among children should be explored further.

Introduction

There has been much discussion about gamified surveys in the market research industry in the last 5 years (see Adamou, 2010, 2013; Baker, 2011; Puleston, 2011, 2013; Puleston & Sleep, 2011). Researchers argue that gamification increases the motivation and engagement of the respondents, which results in better data quality in web surveys. There is no uniform understanding of how gamification should be applied and no clear examples of good and poor gamified surveys. Gamified surveys, however, should include at least four basic elements (Adamou, 2013; McGonigal, 2011; Puleston, 2013): (1) rules and goals—it should be clear for the respondents what should be achieved and how it can be done, (2) an involving narrative—a screenplay of the gamified survey which

---

[1] National Research University Higher School of Economics, Russian Federation

Corresponding Author:
Aigul Mavletova, National Research University Higher School of Economics, Kochnovskiy proezd 3, Room 429, Moscow 155319, Russian Federation.
Email: amavletova@hse.ru

motivates respondents to achieve goals, (3) challenging tasks or quests which require some elements of skill and effort to maintain their interest, and (4) regular feedback and rewards which keep their engagement. An additional element is an attractive design. Though this is not an obligatory element, the design is considered to be one of the most important means of engagement. Puleston (2013) suggests that fun-and game-like wording of questions is also the core element of gamified surveys (see also Petit, 2014). In addition to the problems of defining gamified surveys, it can be also difficult to apply game logic in web surveys which can be contradictory to the research goals (see Poynter, 2012).

In this article, data quality between three conditions is compared among children and adolescents 7–15 years old as follows: (1) a text-only survey, (2) a visual survey with an attractive design and images (functional and illustrative), and (3) a gamified survey. An experiment using a volunteer online access panel in Russia was conducted among 1,050 children by the researcher who is not experienced in designing gamified surveys. However, we suggest that an experience in designing web surveys in general is transferable to designing gamified surveys in particular. The wording in a gamified survey has not been changed to focus on other elements of gamification and their effect on data quality than the wording itself. In the next sections, we outline the hypotheses and literature review, describe the experimental design, and present the results of the experiment, before concluding how gamification can work in web surveys among children.

Literature Review and Hypotheses

Borgers, De Leeuw, and Hox (2000) recommend designing attractive questionnaires for children to maintain their engagement. Scott (2008) argues that some images and visual stimuli are especially helpful for children under 11. This age-grouping is associated with Piaget's theory of intellectual development (Piaget, 1948). According to Piaget (1948), the child starts reason formally, think hypothetically, and make generalizations at 11–12. The children of 7–11 are at the stage of concrete operations when they have some problems with abstract concepts and hypothetical thinking (for more details of application Piaget's theory in surveys, see Borgers, De Leeuw, & Hox, 2000; De Leeuw, Borgers, & Smits, 2004).

Adamou (2012) suggests that gamified questionnaires increase engagement among children. Though she does not show any experiments, Adamou (2012) refers to the gamified survey she and her company conducted among children aged 7–10, which collected a large number of interviews within a short period of time and produced extremely positive feedback from the respondents. König (2011) compared a text-only and a visual web survey with images among children aged 8–13 and found a lower break off rate in the visual than in the text-only survey (31.5% and 28.4%, respectively, Cramer's $V = 0.033$, $p < .05$). In her second web experiment among children, König compared four conditions (a text-only survey, a survey with the image of an interviewer presented on each page, a survey with the image of a cartoon-like boy presented on each page, and an audio web interview) and found no difference in the break off rates.

Based on several experiments which compared gamified and nongamified surveys among adults, Puleston (2013) shows that gamified surveys produce lower break off rates. Findlay and Alberts (2011) found lower break off rates in the gamified web diary than in a traditional web diary. Turner, Van Zoonen, and Adamou (2013) conducted a gamified survey among adults to explore how people manage their identities and what they imagine for the future of identity management. Respondents were expected to act as spies and destroy a rival spy's identity in the survey. Though no experiment has been conducted comparing gamified and nongamified conditions, 96% of the respondents reported they would like to participate in surveys like that again. Turner and her colleagues (2013). suggest that according to "fun theory" if something is fun (e.g., gamified surveys), people are more likely to complete the task. However, Downes-Le Guin, Baker, Mechling, and Ruyle

(2012) found that a gamified survey significantly increased break off rate in their experiment among adults. They compared a text-only web survey, a "decoratively visual" version which included basic visual elements, a "functionally visual" version which included Flash type questions and images used as response categories, and a "gamified" version which included game-like design, narrative, rules, avatar, and rewards. They found that the gamified survey significantly increased break off rate compared to the other conditions. More break offs were produced mainly by the longer time requested for downloading the gamified version (up to 2 min at the beginning of the survey).

In the current experiment, the gamified version did not require any additional time for downloading. However, in both the visual and gamified questionnaires drag-and-drop Flash-based questions and Javascript slider bars were used. Some earlier experiments showed that Flash-based questions may produce higher (Laufer, Klapproth, & Noll, 2009) and lower break off rates (Puleston & Sleep, 2008). Despite the mixed results of a gamification effect in the experiments among adults, we suggest that involving narrative and challenging tasks in the current gamified survey will motivate children and adolescents to finish completing the survey.

> **Hypothesis 1:** The gamified questionnaire produces lower break off rates compared to the visual and text-only web surveys.

Younger respondents are interested and motivated to participate in web surveys. Van Hattum and De Leeuw (1999) found a significantly lower item nonresponse in a computer-assisted self-administered survey compared to a paper survey among children (5.7% and 14.1%, respectively). König (2011) found that adding images in the question texts and response categories can decrease measurement error in some of the questions in web surveys among children. She found better quality in the negatively formulated questions in the visual web survey especially among younger respondents of 8–9 compared to the text-only survey. Fuchs (2005) found stronger primacy effects in the 13–15 age-group among respondents 13–21 in self-administered paper and mail surveys. König (2011) found lower primacy effects among younger age-group 8–9 in a web survey among children aged 8–13. She also found lower primacy effects in the questionnaire with the image of a cartoon-like boy than in the text-only survey. At the same time, more socially desirable responses were received in the questionnaires with the image of an interviewer and a cartoon-like boy. Moreover, in both questionnaires data quality in the cognitively demanding questions (e.g., ranging) was lower, since the images distracted the attention of the children from the particular task (König, 2011).

Some experiments among adults showed either a positive or no gamification effect. Downes-Le Guin and his colleagues (2012) found no differences between the text-only, decoratively visual, functionally visual, and gamified questionnaires in the level of inconsistent responses, failure to follow researcher's instruction to select the "strongly agree" column in the grid question, and level of straight-lining (tendency to select the same response of a scale). Ray Poynter suggests that the gamified version developed by Downes-Le Guin and his colleagues was a poor example of gamification mainly for four reasons (personal communication at Webdatanet Conference on Mobile Research, March 31, 2014). First, the narrative and the tasks were artificially implemented in the survey, which made it more difficult for the respondents to answer questions compared to the text-only and visual conditions. Second, the graphics were poor. Third, not all research topics can be gamified, and this survey is one such example. Fourth, designing gamified surveys requires some experience in game design or gamified surveys. It may have been more effective to bring in other researchers to design the survey. We suggest that there can be different opinions about poor and good examples of gamification and who can gamify surveys well but there is a lack of research experiments in that field.

Cechanowicz, Gutwin, Brownell, and Goodfellow (2013) compared three conditions among adults: a standard web survey, a "partial game," and a "full game." The questionnaire had game-like tasks such as image selection, slogan matching, and a quiz. The "partial game" had a

more attractive design and introduced the questionnaire as the "game show." The "full game" showed if responses were correct, gave the points for correct answers, and displayed a timer which had an effect on the points respondents received (the faster a respondent answered the questions, the more points he or she received). They found the lowest overall item nonresponse rate in the fully gamified questionnaire. However, they randomized the order of the three survey tasks and found that the item nonresponse rate was lower at the beginning but not at the end of the survey. They also found shorter responses in the fully gamified survey, however, that was due to the rewards system which was correlated with the time respondents spent on each task. A number of experiments Puleston conducted (see Puleston, 2013; Puleston & Malinoff, 2011; Puleston & Sleep, 2011) showed that gamified surveys produce lower item nonresponse rates, a lower level of straight-lining, and middle-point response style, less speeding, longer responses in the open-ended questions, and a higher subjective evaluation of the survey. Puleston and Rintoul (2012) found a similar positive gamification effect across different cultures.

The gamified questionnaire in the current study is expected to produce stronger motivation and engagement among children and adolescents, which results in higher data quality: lower item non-response rate, lower primacy effects, lower level of straight-lining, extremeness (tendency to use end poles of a scale) and middle-point response styles, higher consistency, higher data quality in the partially labeled questions (see Borgers, Hox, & Sikkel, 2003), and longer responses in the open-ended questions. At the same time, the gamified questionnaire which includes personalization and avatars produces more socially desirable responses in sensitive questions and lower rate of "don't know" responses. Previous research shows the importance of having an explicit don't know option in the surveys among children, especially in the attitude questions (Scott, 2008; Vaillancourt, 1977), as younger respondents are more likely to give an answer rather than say they don't know. (König, 2011) found that they tend to agree with some statements in the surveys with the virtual interviewer rather than selecting the don't know response option compared to the text-only survey.

> **Hypothesis 2:** The gamified survey produces a lower measurement error (lower primacy effects, a lower level of straight-lining, extremeness and middle-point response styles, higher consistency, higher quality in the partially labeled questions, longer responses in the open-ended questions), and a lower nonresponse error (an item nonresponse rate). However, it also produces more socially desirable responses.

Experimental Design and Data Collection

Children aged 7–15 were randomly assigned to one of the three conditions: (1) text-only survey, (2) visual survey, (3) gamified survey. The experiment was conducted using a volunteer online access panel managed by Online Market Intelligence (see http://omirussia.ru/en) from February 17 to February 27, 2014, in Russia. Young members of the online access panel under 15 were invited to participate. Since most of the panelists are over 16, parents who reported in their profile questionnaire that they had children aged 7–15 were also asked to invite them to participate in the survey. The participation rate was 10% (which is lower compared to the usual participation rate of 30–40% in the panel among adults over 16): 13,315 invitations were sent and 1,329 respondents either completed the survey or were screened out due to their age. In total, 1,050 children were eligible for the survey and filled out the questionnaire.

Three questionnaires were programmed using Unipark online research software (see http://www.unipark.com). Though three different projects were programmed on Unipark, the panel provider used one link for the experiment, and respondents were randomly invited to one of the three questionnaires on the fly. However, the number of these assignments was different in the three conditions (for some reasons significantly lower in the second condition): 549 respondents were

assigned to the text-only condition, 472 to the visual, and 570 to the gamified. Though an increasing number of respondents complete the survey via mobile device (see Kinesis report, 2013) none of the three surveys were mobile optimized. Since both the gamified and visual questionnaires included many images and several Flash-based questions, the text-only survey was also not mobile optimized.

Questionnaires

The questionnaire included 79 items. There were several questions about school, Internet usage, and cognitive abilities, in addition to questions about whether respondents liked the survey, needed help answering questions, used the Internet to find responses, or had other parties present while completing the survey. Some of the questions were similar to the questions used in the Understanding society survey for children and adolescents (see https://www.understandingsociety.ac.uk/documentation/mainstage/questionnaires). The gamified version also allowed respondents to name themselves and select their avatars. The text-only and visual questionnaires were presented on 22 pages. The gamified questionnaire was presented on 28 pages.

The text-only survey had no images except the images in the cognitive test. No Flash or Javascript-based questions were included (see Appendix A for the introductory page), but respondents could use a Javascript-based timer in all three conditions in the cognitive test. The visual questionnaire included background color, images used as illustrations in the questions and as response categories, Flash-based drag-and-drop questions, and Javascript-based slider bars (both types of questions replaced grids used in the text-only survey). The gamified questionnaire had the same design as the visual survey and such gamification features as narrative, rules, points, rewards, feedback, and personalization. The core of the narrative was that the respondent was traveling in the Antarctic and experienced a shipwreck (see Appendix A). Some penguins saved their life and are going to repair the ship. While waiting for repair, the respondents were expected to tell the penguins about their everyday lives and earn 500 points to travel back home. To receive these points, respondents were expected to get through three levels by helping the penguins: in the first level, they were expected to bring more ice to counteract deglaciation; in the second level, they were supposed to catch fish for the penguins; and on the third level, they were expected to save the penguins from leopard seals. In the second and third levels, these tasks were in the form of short Javascript-based games (see Appendix B). The basic elements of the gamified survey can be described as follows:

- *Narrative*—traveling in the Antarctic and helping penguins
- *Rules*—time limit (in a cognitive test) and receiving points
- *Challenges*—helping and saving as many penguins as possible
- *Rewards*—points

While completing the survey, the respondents regularly received feedback, which was expected to maintain their motivation like "you have almost completed the first level, well done" or "you are about to save and feed penguins." At the beginning of the gamified survey, respondents were asked to write their name or nickname and to select an avatar. This personalization was used during the whole survey: children were addressed by nicknames, and after completing each of the three levels, their avatars were shown with the number of points they received.

The visual and gamified questionnaires included free images from the Pixabay website (http://pixabay.com). In addition, the gamified survey had the scripts for two Javascript-based games from the website http://htmlweb.ru/java/games.

Indicators

The concept of data quality is multidimensional (see Biemer & Lyberg, 2003). In this article, data quality is compared based on several indicators of nonresponse and measurement errors in the three survey conditions. Since respondent's satisfaction and survey evaluation may have an effect on both nonresponse and measurement errors, these respondent burden indicators are also compared. Below the main indicators are outlined.

Objective and subjective evaluation of respondent burden. Respondent burden is evaluated based on both objective and subjective indicators (see Bradburn, 1977; Crawford, Couper, & Lamias, 2001; Galesic, 2006; Sharp & Frankel, 1983). *Objective indicators* include:

- *Breakoff rates*.
- *Survey completion time*.
- *Request for help to complete the survey*. Though one of the requests was filling out the questionnaire by oneself, for some of the children it was difficult and they requested help to complete the survey. The percentage of the children who asked for help is compared between the surveys.
- *Score and completion time in a cognitive test*. A short and easy cognitive test was included in the survey. It was programmed in the same way in all three surveys (see Appendix C). There were 10 questions in which respondents were expected to select correct images (see Wasserman, Cherednikova, Schelkova, Malkova, & Ananjeva, 2008). Questions with images were selected to prevent children from finding responses on the Internet. The goal of this task was to select correct responses as soon as they could and to spend less than 5 min. Children could also use a timer in each survey condition to monitor the time they spent for the task. The percentage of the correct responses and average time spent on the task are compared between the conditions.

Subjective indicators of respondent burden include:

- Subjective *evaluation of the interview length* (how many minutes it took to fill out the questionnaire).
- Self-reported *difficulty* of completing the survey (5-point scale).
- Self-reported *enjoyment* of completing the survey (5-point scale).
- Self-reported *interest in receiving new survey invitations* (4-point scale).

The overall item nonresponse rate. There were only a few basic obligatory questions in the survey (gender and age questions in all three surveys and additional avatar and name questions in the gamified survey). The remaining questions were not obligatory. Respondents were told about that opportunity at the beginning of the survey in the introductory text. The overall item nonresponse rate is compared between the conditions. Since there were Flash-based questions in the gamified and visual surveys, the item nonresponse rate is calculated with and without these items.

Socially desirable and socially undesirable responses. Social desirability bias is compared based on the following two indicators:

- *Socially undesirable responses in sensitive questions*. There were 11 sensitive questions: missing school without permission of teachers or parents, smoking cigarettes, drinking alcohol, having friends who drink alcohol at least once in a month, stealing in the shop, taking someone's belongings, finding and taking someone's wallet, being physically bullied at

school, physically bullying other children, poor grades at school, and subjective evaluation of academic performance (response category "I study worse than most of my classmates"). The number of the socially undesirable responses is compared between the conditions.

- *The rate of don't know responses*. There is some research evidence that children tend to construct a response rather than state they don't know the answer. Thus, the rate of don't know responses based on 58 items is also compared between the conditions.

Straightlining, response styles, and response consistency. Extremeness, middle-point response styles, and straight-lining (tendency to select the same response category in all or all except 1 item in at least one of the grids) are analyzed based on the three questions (Q7, Q15, and Q21) with 26 items. These questions were programmed as grids in the text-only questionnaire and as Flash-based and Javascript-based questions in the two other conditions (see Appendix D). The percentages of those respondents who straight-lined and the number of responses using middle or end poles of a scale are compared between the conditions. In addition, an experiment with the partially labeled (only end poles of a scale) and fully labeled response categories was implemented in the question which included 8 items about satisfaction with family and school life using a 7-point scale (Q15). The interaction effect between the scale labeling and survey condition on a likelihood of selecting end poles is tested.

Response consistency is calculated based on the two grid questions (Q7 and Q21) with the reversed-polarity items (e.g., "I like school" vs. "I don't like school"). Respondents could agree or disagree with the statements using a 5-point scale and the don't know response category. If participants either completely agreed or completely disagreed with both reversed-polarity items responses are classified as inconsistent.

Primacy effects. There were two multiple-response and two single-response questions in which the order of the responses was manipulated (standard and reverse order). The multiple-response questions were about the social network websites children used and technical devices respondents or their family had. Respondents could select 11 technical devices such as home theater, smartphone, tablet, digital camera, smart TV, and so on. The visual and gamified surveys presented images of the devices. The single-response questions were about the average duration of Internet usage and subjective evaluation of academic performance. The percentage of those who select the first response category depending on the response order is compared between the conditions.

Open-ended questions. There were two open-ended questions in the survey: what children liked at school and what they didn't like at school. The length of the responses, the number of reasons children reported for liking or disliking school, the rate of the respondents who did not answer or wrote nonsubstantive responses (no-meaning words and signs) are compared between the conditions.

Results

Respondent Burden

In total, 1,050 children completed the survey: 372 respondents completed the text-only survey, 324 respondents completed the visual survey, and 354 respondents completed the gamified survey. Contrary to expectations, no difference in the break off rates was found: 14.0% in the text-only survey, 18.0% in the visual survey, and 17.5% in the gamified survey (see Table 1). About 13% of the participants who started the survey used mobile devices (9% used mobile phones and 4% used tablets). The break off rate was significantly higher, $\chi^2(2) = 50.759$, $p < .001$, among mobile phone users (37.1%) than tablet (21.2%) or PC users (14.1%), with no difference between the conditions,

Table 1. Break Off Rates and Sample Composition.

	Text-Only Survey	Visual Survey	Gamified Survey	$\chi^2(df)$
Number of completed interviews	372	324	354	
Number of screened out	100	63	116	
Break off rate	14.0% (77)	18.0% (85)	17.5% (100)	3.675(2) (n.s.)
Sample composition				
Age-group				
7–9	19.9%	24.7%	27.1%	9.520(6); (n.s.)
10–11	27.7%	28.4%	24.3%	
12–13	12.4%	14.8%	12.1%	
14–15	40.1%	32.1%	36.4%	
Gender				
Boys	46.2%	46.9%	47.7%	0.165(2); (n.s.)
The frequency of Internet usage				
Using the Internet every day or almost every day	79.0%	77.5%	82.2%	2.397(2); (n.s.)
The duration of Internet usage				
7 or more hours per day	7.0%	7.4%	7.4%	0.054(2); (n.s.)

$\chi^2(4) = 5.685$, $p = .224$. Among those who were eligible for the survey, 90% filled out the questionnaire via PC, 6.3% via mobile phones, and 3.7% via tablets with no difference between the conditions, $\chi^2(4) = 0.948$, $p = .918$.

The sample composition in age, gender, the frequency and duration of Internet usage was similar in all three conditions (see Table 1). In total, 47% of the respondents are boys; 24% of the respondents are aged 7–9, 27% 10–11, 13% 12–13, and 36% 14–15, with no significant difference between the conditions. About 80% use the Internet almost every day and 7% use the Internet 7 or more hours per day.

The average completion time was different by the condition: 13.90 min ($SD = 5.06$) in the text-only survey, 15.15 min ($SD = 6.14$) in the visual survey, and 19.36 min ($SD = 7.57$) in the gamified survey, $F(2, 974) = 67.744$, $p < .001$, (the values were cut at the 95th percentile for each version to remove outliers). A multivariate linear regression showed a significant age effect, with older respondents spending less time ($b = -0.405$, $p < .001$) and a survey condition effect with the respondents spending less time in the text-only ($b = -5.033$, $p < .001$) and visual survey ($b = -3.770$, $p < .001$) compared to the gamified survey. Those who had a lower item nonresponse rate ($\beta = -9.263$, $p < 0.001$) and who completed the survey via a mobile device rather than PC ($\beta = -2.192$, $p < 0.001$) spent more time on the questionnaire. At the same time, no difference in the subjective time evaluation was found between the conditions, $M = 12.88$ min, $SD = 6.46$, $F(2, 1013) = 0.617$, $p = .540$.

Significantly fewer children requested help in the gamified (11.4%) than in the text-only (17.1%) and visual survey (17.7%), $\chi^2(2) = 6.331$, $p < .05$. A multivariate logistic regression showed an age effect with older respondents requesting assistance less often (odds ratio [OR] $= 0.798$, $p < .001$) and a survey condition effect with a higher likelihood of request for help in the text-only (OR $= 1.995$, $p < .05$) and visual survey (OR $= 1.731$, $p < .05$) compared to the gamified survey.

Significantly more children found it easier to fill out the gamified survey, $\chi^2(8) = 38.003$, $p < .001$. For instance, the top response on the 5-point scale "very easy" was selected among 31.0% of the respondents in the text-only, 30.4% in the visual, and 44.4% in the gamified survey. A multivariate logistic regression predicting the top response showed that it was easier for older respondents (OR $= 1.130$, $p < .001$), for those who had good academic performance (OR $= 1.688$, $p < .001$), and who used the Internet almost every day (OR $= 1.501$, $p < .05$). It was more difficult to complete the text-only (OR $= 0.522$, $p < .001$) and visual survey

(OR = 0.551, $p < .001$). Those respondents who asked for help while filling out the questionnaire found the survey more difficult (OR = 0.490, $p < .01$). Predicting the top two responses ("very easy" and "rather easy") showed a significant difference between the text-only (OR = 0.457, $p < .001$) and gamified survey and no significant difference between the visual (OR = 0.860, $p = .474$) and gamified survey.

More children found it more enjoyable to fill out the gamified questionnaire, $\chi^2(8) = 42.412$, $p < .001$: 28.6% selected the top response on the 5-point scale ("I liked it very much") in the text-only survey, 33.8% in the visual survey, and 48.1% in the gamified survey. A multivariate logistic regression predicting the top response showed that it was less enjoyable to complete the text-only (OR = 0.434, $p < .001$) and visual survey (OR = 0.569, $p < .01$). Younger respondents (OR = 0.923, $p < .05$), girls (OR = 1.369, $p < .01$), those who performed well at school (OR = 1.389, $p < .05$) and who spent on average 7 or more hours on the Internet per day (OR = 2.028, $p < .01$) provided more positive evaluation. Predicting the top two responses ("liked it very much" and "rather liked it") showed a significant difference between the text-only (OR = 0.539, $p < .01$) and gamified survey and no significant difference between the visual (OR = 0.852, $p = .502$) and gamified survey.

More respondents were interested in receiving new survey invitations in the visual and gamified conditions, $\chi^2(8) = 15.386$, $p = .052$: 56.8% selected the top response on the 4-point scale ("very interested"), 62.0% in the visual, and 68.7% in the gamified survey. A multivariate logistic regression showed that fewer respondents were interested in receiving new invitations in the text-only (OR = 0.633, $p < .01$) than in the gamified condition. No significant difference was found between the visual and gamified conditions. In addition, younger respondents (OR = 0.929, $p < .05$) and girls (OR = 1.519, $p < .01$) were more interested in receiving new invitations.

A short and easy cognitive test was included in the survey. The gamified survey was expected to increase engagement and produce a higher score and a faster completion time of the test. However, no differences in the average completion time (2.7 min in all three surveys), the percentage of correct responses (78% in the text-only and 81% in the visual and gamified surveys, $F(2, 1047) = 1.278$, $p = .279$), or the item nonresponse rate in the cognitive test, $F(2, 1047) = 2.291$, $p = .102$, were found.

Item Nonresponse Rate

Contrary to expectations, the highest item nonresponse rates were found in the gamified (11.73%) and visual surveys (11.12%) than in the text-only survey (1.73%; see Table 2). A linear multivariate regression showed significantly lower item nonresponse rate in the text-only ($\beta = -8.924$, $p < .001$) compared to the gamified survey. In addition, those who filled out the questionnaire via PC rather than a mobile device produced a lower item nonresponse rate ($\beta = -18.200$, $p < .001$). Though Borgers, De Leeuw, and Hox (2000) report that item nonresponse rate is higher among younger children and boys, no support for these findings were found in the current experiment.

While excluding Flash-based questions from the analysis (implemented as grids in the text-only survey), no difference in the item nonresponse rate was found: 1.89% in the text-only, 2.06% in the visual, and 2.11% in the gamified survey (see Table 2). Javascript-based slider bars in the visual and gamified surveys (Q15) also resulted in a higher item nonresponse rate (3.3% and 2.9%, respectively) compared to the grid items in the text-only survey (0.6%), $F(2, 1047) = 5.319$, $p < .01$). The results are consistent with Couper, Singer, Tourangeau, and Conrad (2006).

Socially Desirable and Socially Undesirable Responses

Contrary to expectations, the gamified survey did not produce a higher rate of socially desirable responses. A negative binomial model predicting the number of socially undesirable practices

Table 2. Overall Item Nonresponse Rate.

	Text-Only Survey	Visual Survey	Gamified Survey	Statistics
N	372	324	354	
Overall item nonresponse rate Excluding Flash-based questions	1.73% (4.43%)	11.12% (16.33%)	11.73% (16.32%)	$F(2, 1047) = 63.421$***
Overall item nonresponse rate	1.89% (5.72%)	2.06% (7.01%)	2.11% (5.64%)	$F(2, 1047) = 0.123$; (n.s.)

***$p < .001$, standard deviation in parentheses.

reported in the sensitive questions showed no survey condition effect. As expected, younger respondents (OR = 1.171, $p < .001$) and girls (OR = 0.840, $p < .05$) had a lower rate of socially undesirable practices. No difference in the item nonresponse rate (0.44% in the text-only, 0.67% in the visual survey, and 0.72% in the gamified survey), $F(2, 1047) = 0.366$, $p = .693$) in the sensitive questions was found.

No difference in the rate of don't know responses in 58 items was revealed: 2.17% ($SD = 5.14\%$) in the text only, 2.15% ($SD = 3.45\%$) in the visual, and 2.36% ($SD = 4.62\%$) in the gamified survey, $F(2, 1047) = 0.222$, $p = .801$. A negative binomial model predicting the number of don't know responses showed a significant negative effect of academic performance (OR = 0.785, $p < .01$) and a positive effect of the presence of bystanders (OR = 1.281, $p < .05$). Those who had a lower academic performance and who filled out the questionnaire in the presence of bystanders produced a higher rate of don't know responses.

Straight-Lining, Response Styles, and Consistency of the Responses

As expected, the text-only condition produced a higher level of straight-lining: 11.4% in the text-only survey, 2.8% in the visual survey, and 3.2% in the gamified survey. No difference between the visual and gamified surveys was found. This is rather the difference between the types of questions—grids and drag and drop as well as grids and slider questions—than the difference between the survey conditions (see Drolet, Butler, & Davis, 2009). However, there was also a higher item nonresponse rate in the gamified and visual surveys in both Flash and Javascript-based questions. A multivariate logistic regression showed that those children who completed the survey in the presence of bystanders (OR = 2.121, $p < .05$), who filled out the text-only survey (OR = 3.840, $p < .001$), who got a lower score in the cognitive test (OR = 0.118, $p < .001$), and who reported good performance at school (OR = 2.049, $p < .05$) had a higher likelihood of straight-lining (see Table 3).

On average, 35.5% of the responses were given using extreme scale categories. A negative binomial model predicting the number of extreme responses showed that younger respondents (OR = 0.960, $p < .05$) and those who had good academic performance (OR = 1.287, $p < .001$) produced more answers using extreme scale categories (see Table 3). No survey condition effect was found. An experiment with the partially labeled and fully labeled response categories was implemented in Q15. A negative binomial model predicting the number of extreme responses in that question showed that full labeling significantly decreased the likelihood of using extreme response categories compared to the partially labeled scale (OR = 0.818, $p < .01$). However, no interaction effect between the labeling and survey condition was found.

About 14.0% of the responses were given using the middle-point category. A negative binomial model showed that the text-only survey produced a higher rate of middle responses compared to the gamified survey (OR = 1.223, $p < .05$; see Table 3). No difference between the gamified and visual surveys was found. Younger respondents (OR = 1.079, $p < .001$), those with higher academic

Table 3. Straight-Lining, Extreme Response Style, and Middle-Point Response Style (Odds Ratio).

	Straight-Lining (A Logistic Regression)	Extreme Response Style (A Negative Binomial Model)	Middle-Point Response Style (A Negative Binomial Model)
Intercept	0.303 [0.041, 2.232]	13.075 [7.805, 21.906]***	1.541 [0.878, 2.703]
Text-only survey	3.840 [1.762, 8.370]***	1.094 [0.919, 1.303]	1.223 [1.014, 1.475]*
Visual survey	0.840 [0.300, 2.347]	0.980 [0.812, 1.182]	0.971 [0.792, 1.191]
Age	0.876 [0.760, 1.010]	0.960 [0.927, 0.993]*	1.079 [1.040, 1.120]***
Girls	1.060 [0.580, 1.934]	1.011 [0.873, 1.170]	0.987 [0.842, 1.156]
Good academic performance	2.049 [1.087, 3.860]*	1.287 [1.110, 1.494]***	0.711 [0.606, 0.834]***
Cognitive test (percentage of correct responses)	0.118 [0.048, 0.295]***	0.885 [0.656, 1.194]	1.278 [0.937, 1.741]
Using the Internet every day or almost every day	1.672 [0.771, 3.622]	1.071 [0.885, 1.297]	0.810 [0.659, 0.996]*
Using the Internet 7 or more hours per day	1.146 [0.306, 4.293]	1.077 [0.810, 1.433]	0.978 [0.722, 1.327]
Mobile web survey mode	0.211 [0.027, 1.669]	1.065 [0.741, 1.530]	0.910 [0.618, 1.340]
Presence of bystanders	2.121 [1.057, 4.256]*	0.932 [0.779, 1.116]	1.095 [0.903, 1.327]
Getting help while completing the survey	0.500 [0.209, 1.192]	0.966 [0.768, 1.215]	1.023 [0.805, 1.302]
Partially labeled scale +	0.690 [0.379, 1.256]	0.957 [0.829, 1.105]	0.983 [0.842, 1.147]
Model	$\chi^2 (11) = 57.119$*** R^2 (Cox & Snell) = .065, R^2 (Nagelkerke) =.175	Likelihood ratio $\chi^2(12) = 20.743$, $p = .054$	Likelihood ratio $\chi^2(12) = 45.946$***

Note. Reference category = gamified survey; + = the experiment with partially vs. fully labeled responses was implemented in Q15.
*$p < .05$. **$p < .01$. ***$p < .001$, 95% CI in brackets.

performance (OR = 0.711, $p < .001$), and those who used the Internet every day (OR = 0.810, $p < .05$) tended to use middle responses less often.

Based on the two questions with the reversed-polarity items (Q7 and Q21), the percentage of those respondents who gave inconsistent responses was calculated. On average, 5.7% responses were inconsistent. A negative binomial model predicting the number of inconsistent responses showed no survey condition effect (OR = 1.108, $p = .489$ in the text-only and OR = 1.045, $p = .783$ in the visual survey compared to the gamified survey).

Primacy Effects

Consistent with the results of Fuchs (2005), who reported moderate primacy effects in multiple-response questions among respondents 10–21 years old, a significant moderate primacy effect was found in the current experiment in the question about technical devices respondents had. There were two response options ("smart TV" and "home theater") which were presented either at the beginning or at the end of response list. If smart TV was presented at the beginning of response list, more children selected it (6.5 percentage points difference), $\chi^2(1) = 5.170$, $p < .05$. A logistic regression predicting the selection of the smart TV response category based on the order of the responses (standard or reverse), survey condition, and the interaction effect showed lower primacy effects in the text-only than in the gamified condition (OR = 0.508, $p < .01$), and no significant difference between the gamified and visual conditions. It seems that presenting smart TV with the image of a TV-like device produced stronger primacy effects compared to the control condition without images. At the same time, no primacy effects were found while analyzing another response option

(home theater), $\chi^2(1) = 0.046$, $p = .831$, and three other questions in which the response order was manipulated.

Open-Ended Questions

Respondents were expected to write what they liked and what they didn't like at school. The text-only condition produced a lower percentage of those respondents who missed or wrote nonsubstantive responses (no-meaning words and signs), while the visual survey produced a higher percentage of these participants. The gamified condition was closer to the text-only condition. The difference was not statistically significant in the question about the reasons for liking school (4.0% in the text-only, 7.4% in the visual, and 5.9% in the gamified survey), $\chi^2(2) = 3.709$, $p = .157$, but reached statistical significance in the question about the reasons for disliking school (5.1% in the text-only, 10.5% in the visual, and 6.8% in the gamified survey), $\chi^2(2) = 7.634$, $p < .05$.

Contrary to expectations, no difference in the length of the responses, the number of reasons children reported for liking school, and the distribution of these reasons was found between the conditions. On average, respondents wrote 49.9 characters in the text only ($SD = 58.1$), 43.6 characters in the visual ($SD = 57.6$), and 47.7 characters ($SD = 64.0$) in the gamified survey, $F(2, 987) = 0.901$, $p = .407$. Almost half of the children reported one reason they mostly liked school, 25% two reasons, and the remaining 25% three to seven reasons with no difference between the conditions, $\chi^2(12) = 12.358$, $p = .417$.

No differences in the length and the distribution of the responses were found in the question about the reasons for disliking school. On average, respondents wrote 40.3 characters in the text only ($SD = 64.5$), 35.6 characters in the visual ($SD = 52.9$), and 32.8 characters ($SD = 55.8$) in the gamified survey, $F(2, 970) = 1.451$, $p = .235$. However, a slightly higher social desirability bias was found in the gamified condition. Significantly more respondents reported at least two reasons in the text-only and visual surveys compared to the gamified survey (27.5% in the text-only, 29.7% in the visual, and 19.7% in the gamified survey), $\chi^2(6) = 12.635$, $p < .05$.

Discussion and Conclusion

Is it worth designing gamified surveys for children? Jon Puleston wrote that there is almost no need to do that (April 7, 2014, e-mail communication with Jon Puleston, including Betty Adamou and Ray Poynter). They are more responsible participants than adults: they rarely fill out the questionnaires and if they do, their motivation is high. If the survey is not interesting for them and the questions are not relevant, gamification will not help to improve data quality.

The current experiment showed that the gamified survey with no changes in the question wording had no effect on the break off rates, the rate of don't know and socially undesirable responses in the closed-ended questions, the level of extreme and inconsistent responses, and the length of the answers in the open-ended questions. A higher overall item nonresponse rate, a lower level of straight-lining, and a lower level of middle-point response style were found in both gamified and visual surveys, which were mainly due to the differences in the types of questions (grids vs. Flash-based drag-and-drop questions and Javascript slider bars). The gamified survey produced completion time more than a third longer than the text-only survey. On average, it took 13.9 min to complete the text-only survey and 19.4 min (with 1.6 min for playing two games) to complete the gamified survey. The difference is large, taking into account the overall item nonresponse rate, which was significantly higher in the gamified survey. However, no difference in the subjective time evaluation was found. Most importantly, our experiment does show that designing gamified surveys can be helpful for children and adolescents. Despite the longer survey length in the gamified condition, it was less burdensome to complete it. Children requested help less often. The questionnaire was easier and more enjoyable especially for younger children.

This study has several limitations, and it should be considered as preliminary investigation of a gamification effect in the surveys among children and adolescents. First, it is hard to disentangle the effect of different types of questions (drag and drop and sliders vs. grids) from a gamification effect itself. In future studies, it is worth designing the same types of questions for both text-only and gamified surveys. Second, the author is not experienced in designing gamified surveys, which could have resulted in an example of poor gamification. Third, the experiment may be an example of topics, which are difficult to gamify. Research topics are different: some are easier and some of them are more difficult to gamify. It is easier to gamify such topics in which respondents are expected to imagine the future (e.g., the future of technologies and the future of identity management, see Turner, Van Zoonen, & Adamou, 2013) rather than describe their everyday school life. Finally, the experiment is based on a nonrandom sample of children who are active Internet users and who use the Internet almost every day. These results might be different from less active Internet users.

Despite the study limitations the current findings provide some evidence for a positive gamification effect in the surveys among children. The experiment shows that we need more research in that field. In future studies, it might be worth exploring which gamification narratives yield a more positive experience and decrease respondent burden, what research topics and questions can be well and not well gamified as well as what are the poor and good examples of gamification. It is worth exploring the effect of gamified surveys designed by different researchers with more and less experience in designing gamified surveys. We also encourage others to measure the effect of each gamification element (narratives, rewards, challenging tasks, and rules) on data quality. At the same time, one of the challenges would be designing gamified surveys that do not increase completion time as it might increase respondent burden and interview costs. Overall, we suggest that a gamification effect in web surveys among children should be explored further.

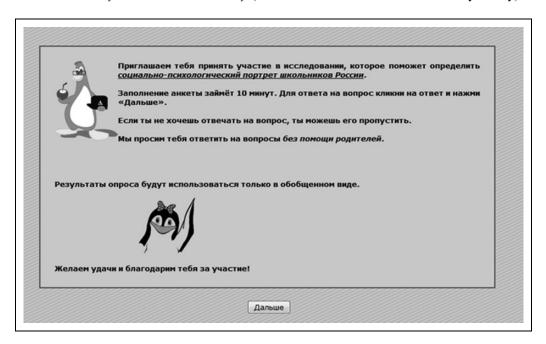## Appendix A

1. *Introductory text in the text-only survey.*

We invite you to participate in a study that will help to draw a sociopsychological portrait of schoolchildren in Russia.

Completing the survey will take 10 min. To answer questions select a response and click on the "Next" button.

If you do not want to answer a question, you can skip it. We kindly ask you to complete the survey without the help of your parents. The survey data will be reported only as a group and not individually. We wish you good luck and thank-you for your participation!

2. *Introductory text in the visual survey* (the same introduction as in the text-only survey).

Приглашаем тебя принять участие в исследовании, которое поможет определить *социально-психологический портрет школьников России*.

Заполнение анкеты займёт 10 минут. Для ответа на вопрос кликни на ответ и нажми «Дальше».

Если ты не хочешь отвечать на вопрос, ты можешь его пропустить.

Мы просим тебя ответить на вопросы *без помощи родителей*.

Результаты опроса будут использоваться только в обобщенном виде.
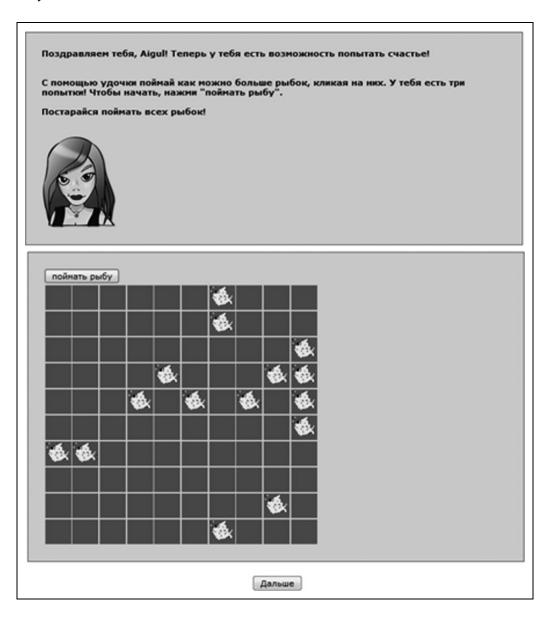
Желаем удачи и благодарим тебя за участие!

Дальше

3. *Introductory text in the gamified survey.*

Hi! This is a game like a survey. Imagine that you have been traveling in the Antarctic on a ship. One night the ship collided with an iceberg and you have experienced a shipwreck.

However, you have been saved by some penguins.

They are willing to help you with the ship, but they expect you to tell them about yourself. You can earn 500 points and pass the following three levels:

Level 1: Help penguins with ice (there is deglaciation in the Antarctic)
Level 2: Catch fish for them
Level 3: Save them from leopard seals!

After passing Level 3, the ship will be repaired.
You'll be able to continue your journey!
Instructions:
Completing the survey will take 10 min. To answer questions select a response and click on the "Next" button. If you do not want to answer a question, you can skip it. We kindly ask you to complete the survey without the help of your parents. The survey data will be reported only as a group and not individually.

We wish you good luck and thank-you for your participation!

Привет! Это опрос в виде игры.

Представь, что ты совершал на корабле путешествие в Антарктиду.

Ночью корабль столкнулся с айсбергом, и ты потерпел кораблекрушение.

Но тебя спасли пингвины!

Они готовы помочь починить корабль, но взамен ты должен рассказать им о себе, заработать 500 очков и пройти три уровня:

Уровень 1 - помочь льдом (в Антарктиде происходит таяние ледников)

Уровень 2 - поймать для них рыбу

Уровень 3 - спасти от морских леопардов!

Через 3 уровня корабль будет починен.

Ты сможешь возобновить свое путешествие!

Инструкция:

Заполнение анкеты займёт 10 минут. Для ответа на вопрос кликни на ответ и нажми «Дальше». Если ты не хочешь отвечать на вопрос, ты можешь его пропустить. Мы просим тебя ответить на вопросы *без помощи родителей*. Результаты будут использоваться только в обобщенном виде.

Желаем удачи и благодарим за участие!

Дальше

Appendix B

1) *Gamified survey*: game on Level 2.

Congratulations, Aigul! You have a chance to try for your fortune!

Using a fishing tackle catch as many fish as you can by clicking on them. You have three attempts. Click on ''catch fish'' to start the game.

Try to catch all of the fish!

2) *Gamified survey*: game on Level 3.

Congratulations, Aigul! You have a chance to save the penguins!

   Try to liquidate as many leopard seals as you can by clicking on them. Click on "save penguins" to start the game.

Appendix C

An example of a cognitive test (1 out of 10 questions).

Q20. Try to spend not more than 5 min on this test. You can use a timer.

In each row you will find four squares. Three squares have some figures. You should decide what should be in the fourth square.

Examples (right answers are marked red):

Try to spend not more than 5 min. But the faster, the better!
Tasks:

1. What should be in the fourth square? Select the response.

Appendix D.

1. Q7.

Text-only survey.

Q7. Please rate how much you agree/disagree with the statements: 1 = *you strongly disagree*, 5 = *strongly agree*.

Give your answers on the basis of how things have been for you over the last 6 months.

I like going to school
I feel myself happy
Usually other children/teenagers bully me
I have very good relationships with my classmates
Almost in all situations I lose confidence
Usually it's difficult to get me angry
Usually I like doing my homework
Usually I cheat off of my classmates

1 = *strongly disagree*
2 = *rather disagree*
3 = *neither disagree nor agree*
4 = *rather agree*
5 = *strongly agree*

Don't know

Visual survey.

Q7. Please rate how much you agree/disagree with the statements: $1 = you\ strongly\ disagree$, $5 = strongly\ agree$.

Give your answers on the basis of how things have been for you over the last 6 months.

Drag the answer to boxes. You can drag as many answers to the same box as you want.

Gamified survey.

Q7. Well done! You are almost there to get ice for penguins!

Please rate how much you agree/disagree with the statements: $1 = you\ strongly\ disagree$, $5 = strongly\ agree$.

Give your answers on the basis of how things have been for you over the last 6 months.

Drag the answer to boxes. You can drag as many answers to the same box as you want.

Превосходно! Ты в одном шаге от добычи льда для пингвинов!

7. Оцени, насколько ты согласен/согласна с утверждениями по шкале от 1 до 5, где 1 – ты совсем НЕ согласен / НЕ согласна, 5 – полностью согласен/согласна.

Оцени утверждения на основании того, что происходило за последние 6 месяцев.

Для ответа перетащи утверждения в коробочки. В каждую коробочку можно перетащить столько утверждений, сколько ты считаешь нужным.

[9] / [9]

Мне нравится ходить в школу

1 – совсем не согласен	2 – скорее не согласен	3 – ни согласен, ни не согласен	4 – скорее согласен

5 – полностью согласен	Не знаю		

Дальше

3. Q15.

Text-only survey

Q15. The next few questions are about how you feel about different aspects of your life. Please rate how much you are satisfied/ not satisfied: 1 = *completely not satisfied*, 7 = *completely satisfied*.
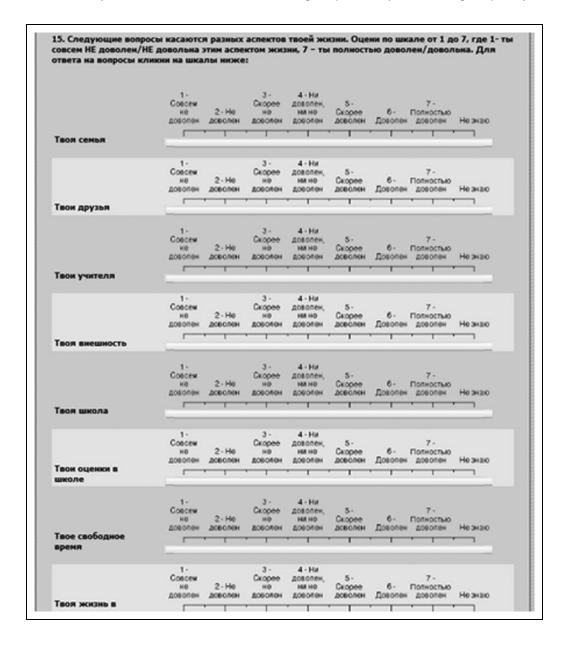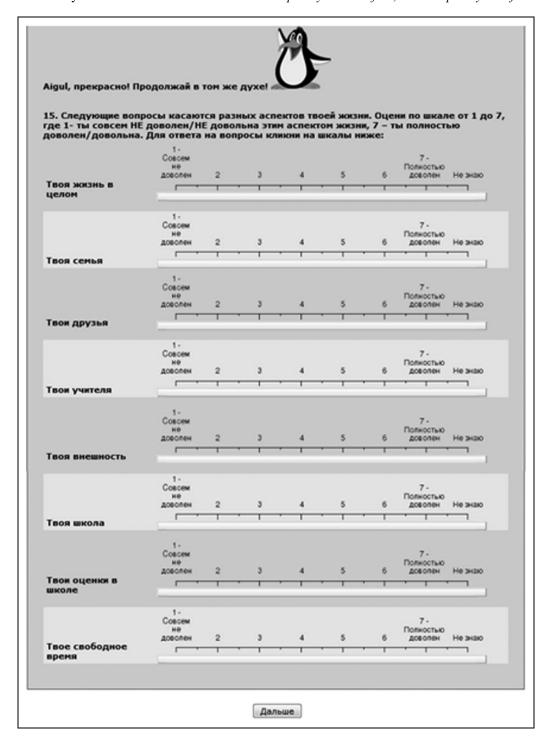
Scale:

- 1 = *completely not satisfied*
- 2 = *not satisfied*
- 3 = *rather not satisfied*
- 4 = *neither not satisfied, nor satisfied*
- 5 = *rather satisfied*
- 6 = *satisfied*
- 7 = *completely satisfied*

Don't know

Visual survey

Q15. The next few questions are about how you feel about different aspects of your life. Please rate how much you are satisfied/not satisfied: 1 = *completely not satisfied*, 7 = *completely satisfied*.

Gamified survey
Q15. Well done Aigul! Go on!

The next few questions are about how you feel about different aspects of your life. Please rate how much you are satisfied/ not satisfied: 1 = *completely not satisfied*, 7 = *completely satisfied*.

Declaration of Conflicting Interests

Funding

References

Adamou, B. (2010). *The future of research through gaming*. Presented at CASRO Online Research Conference. March 3–4, 2010. Retrieved May 11, 2014, from http://c.ymcdn.com/sites/www.casro.org/resource/collec-tion/27A136CC-A18D-4BB5-979F-B2658A3F95C9/Presentation_-_Betty_Adamou_-_Nebu.pdf

Adamou, B. (2012). *Magazine Question Hunt: Using Gamification in Surveys for Children.* Retrieved May 11, 2014, from http://www.gamification.co/2012/07/09/magazine-question-hunt-using-gamification-in-surveys-for-children

Adamou, B. (2013). *Introduction to gamification*. Presented at 2013 NewMR Lecture Series. Retrieved May 11, 2014, from http://newmr.org/play-again/5f9aa2dd

Baker, R. P. (2011). *Survey gamification: Old wine in new bottles?* Presented at The New MR Virtual Festival. November 3, 2011. Retrieved May 11, 2014, from http://www.websm.org/uploadi/editor/1372841699Baker_2011_Survey_Gamification_Old_Wine_In_New_Bottles.pdf

Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to survey quality*. Hoboken, NJ: John Wiley.

Borgers, N., De Leeuw, E., & Hox, J.P. (2000). Children as respondents in survey research: Cognitive development and response quality. *Bulletin of Sociological Methodology*, *66*, 60–75.

Borgers, N., Hox, J., & Sikkel, D. (2003). Response quality in survey research with children and adolescents: The effect of labeled response options and vague quantifiers. *International Journal of Public Opinion Research*, *15*, 83–94.

Bradburn, N. (1977). *Respondent burden*. Proceedings of the Health Survey Research Methods Second Biennial Conference. Washington, DC: U.S. Department of Health, Education and Welfare.

Cechanowicz, J., Gutwin, C., Brownell, B., & Goodfellow, L. (2013). Effects of gamification on participation and data quality in a real-world market research domain. In *Proceedings of the first international conference on gameful design, research, and applications*. Retrieved May 11, 2014, from http://hci.usask.ca/uploads/324-135_Camera_Ready_Effects_of_Gamification_on_Participation_and_Data_Quality_in_a_Real-World_Market_Research_Domain.pdf

Couper, M. P., Singer, E., Tourangeau, R., & Conrad, F. G. (2006). Evaluating the effectiveness of visual analog scales: A web experiment. *Social Science Computer Review*, *24*, 2, 227–245.

Crawford, S. D., Couper, M. P., & Lamias, M. J. (2001). Web surveys: Perceptions of burden. *Social Science Computer Review*, *19*, 146–162.

De Leeuw, E., Borgers, N., & Smits, A. (2004). Pretesting questionnaires for children and adolescents. In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 409–429). New York, NY: John Wiley.

Downes-Le Guin, T., Baker, R., Mechling, J., & Ruyle, E. (2012). Myths and realities of respondent engagement in online surveys. *International Journal of Market Research*, *54*, 613–633.

Drolet, J., Butler, A., & Davis, S. (2009). *The survey "burden factor"– how important is respondent's perception of survey length?* Presented at CASRO Panel Conference. February 2-3, 2009. New Orleans, LA.

Findlay, K., & Alberts, K. (2011). *Gamification: What it is and what it is not*. Presented at ESOMAR Congress, September 18–21, 2011, Amsterdam.

Fuchs, M. (2005). Children and adolescents as respondents. Experiments on question order, response order, scale effects and the effect of numeric values associated with response options. *Journal of Official Statistics*, *21*, 701–725.

Galesic, M. (2006). Dropouts on the web: Effects of interest and burden experienced during an online survey. *Journal of Official Statistics*, *22*, 313–328.

Kinesis. Online survey statistics from the mobile future. (2013). *Kinesis whitepaper*. Retrieved May 11, 2014, from http://www.kinesissurvey.com/wp-content/uploads/2013/10/UPDATED-with-Q3-2013-Data-Mobile-whitepaper.pdf

König, S. (2011). *Online-Forschung mit Kindern*. Farmington Hills: Budrich UniPress.

Laufer, S., Klapproth, U., & Noll, S. (2009). *Using flash type questions—Stroke of luck or curse for data quality?* Presented at General Online Research Conference April 6–8, 2009, Vienna, Austria. Retrieved May 11, 2014, from http://www.slideshare.net/Globalpark/using-flash-type-questions-stroke-of-luck-or-curse-for-data-quality

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. London, England: The Penguin Press.

Piaget, J. (1948). *The moral judgment of the child*. Glencoe, IL: The Free Press.

Petit, A. (2014). Forget gamification; try writing a humanized survey. *Quirk's Marketing Research Review*, Retrieved June 14, 2014, from http://www.quirks.com/articles/2014/20140225-1.aspx?searchID=1031428 200&sort=5&pg=1

Poynter, R. (2012). *Understanding gamification*. Retrieved May 11, 2014, from http://www.visioncritical.com/blog/understanding-gamification

Puleston, J. (2011). Improving online surveys. *International Journal of Market Research*, *53*, 557–560.

Puleston, J. (2013). *Online research, game on!* Presented at NCRM conference "Web Surveys for the General Population: How, Why and When?", February 25–26, 2013, London. Retrieved May 11, 2014, from http://www.websm.org/uploadi/editor/1364212617Puleston_2013_Online_Research_Game_On.pdf

Puleston, J, & Malinoff, B. (2011). *How far is too far? Traditional, flash, and gamification interfaces, and implications for the future of market research online survey design*. Presented at ESOMAR 3D Digital Dimensions (Online + Social Media + Mobile) Research, October 26–28, 2011, Miami.

Puleston, J, & Rintoul, D. (2012). *Can survey gaming techniques cross continents? Examining cross cultural reactions to creative questioning techniques*. Presented at ESOMAR Asia Pacific Conference, April 15–17, 2012, Shanghai.

Puleston, J., & Sleep, D. (2008). *Measuring the value of respondent engagement. Innovative techniques to improve panel quality*. Presented at ESOMAR Panel Research Conference, October 13–14, 2008, Dublin.

Puleston, J., & Sleep, D. (2011). *The game experiments: Researching how gaming techniques can be used to improve the quality of feedback from online research*. Presented at ESOMAR Congress, September 18–21, 2011, Amsterdam.

Scott, J. (2008). Children as respondents: The challenge for quantitative methods. In P. Christensen & A. James (Eds.), *Research with children: Perspectives and practices* (pp. 87–108). New York, NY: Routledge.

Sharp, L. M., & Frankel, J. (1983). Respondent burden: A test of some common assumptions. *Public Opinion Quarterly*, *47*, 36–53.

Turner, G., Van Zoonen, L., & Adamou, B. (2013). *Research through gaming: Public perceptions of (the future of) identity management*. Sage Case Studies in Research Methods.

Vaillancourt, P. M. (1977). Stability of children's survey responses. *Public Opinion Quarterly*, *37*, 373–387.

Van Hattum, M. J. C., & De Leeuw, E. (1999). A disk by mail survey of pupils in primary schools: Data quality and logistics. *Journal of Official Statistics*, *15*, 413–429.

Wasserman, L., Cherednikova, T., Schelkova, O., Malkova, E., & Ananjeva, E. (2008). *Testovaya metodika dlya opredeleniya intellektualnogo potenciala u detey i podrostkov: Standartizaciya i prakticheskaya*

aprobaciya (in Russian) [Method for testing intellectual capacity among children and adolescents: Standardization and approbation]. Saint-Petersburg, The Bekhterev St. Petersburg Psychoneurological Research Institute. Retrieved January 05, 2014, from http://z203712.infobox.ru/content/42/2008_08.pdf

Author Biography

Aigul Mavletova, PhD, is a senior lecturer at the National Research University Higher School of Economics. Her current research interests focus on data quality in web surveys and new technologies used in data collection process. E-mail: amavletova@hse.ru.