

Item Nonresponse Rates and Panel Conditioning in a Longitudinal Survey among Youth

Aigul Mavletova^a, Peter Lynn^b

This is a pre-copyedited, author-produced PDF of an article accepted for publication in *Field Methods* following peer review. The version of record (published version) [Item Nonresponse Rates and Panel Conditioning in a Longitudinal Survey among Youth; Aigul Mavletova and Peter Lynn; *Field Methods* 2019; doi: 10.1177/1525822X18825103] is available online at: <https://journals.sagepub.com/doi/full/10.1177/1525822X18825103> and should be cited as follows:

Mavletova, Aigul and Lynn, Peter (2019) 'Item nonresponse rates and panel conditioning in a longitudinal survey among youth'. *Field Methods*, published online 27-03-2019.

- a) Laboratory for Comparative Social Research and Department of Sociology, National Research University Higher School of Economics, Moscow, Russian Federation, Myasnitskaya street 20, Moscow, 101000, Russia. e-mail: amavletova@hse.ru. (Corresponding Author)
- b) Institute for Social and Economic Research, University of Essex, UK

Abstract

The article examines two important aspects of data quality in self-completion surveys of young people, taking advantage of a unique data source - Understanding Society: the United Kingdom Household Longitudinal Study. Young persons aged 10-15 are asked to complete a self-administered paper questionnaire at annual intervals. The number of completed interviews varies over waves from 4,049 to 5,020. Data are also collected from parents, providing important explanatory covariates for our analysis. Stronger parent-child relationship and higher mother's involvement in education were associated with lower item nonresponse rate and lower inconsistency throughout waves. We also found some evidence for a negative panel conditioning effect with an increase of social desirability bias and measurement errors in the subsequent waves. There was a higher level of inconsistent responses and a higher probability of social desirability bias throughout waves in more sensitive items.

1. Introduction

Based on the children's intellectual development scheme proposed by Piaget (1929), researchers suggest that starting from age 10 or 11 children can answer self-administered standardized questionnaires, however even children aged 8-9 years old can complete self-administered questionnaires with a sufficient level of validity and reliability (Amato and Ochiltree, 1987; Borgers, De Leeuw, and Hox, 2000; Borgers, Hox, and Sikkel, 2003, 2004; Mavletova, 2015a, 2015b; Smith, 2008; Varni, Limbers, and Burwinkle, 2007). Studies of different indicators of data quality such as construct validity (Varni, Limbers, and Burwinkle, 2007), internal consistency (Borgers, De Leeuw, and Hox, 2000; Varni, Limbers, and Burwinkle, 2007), and test-retest reliability (Vaillancourt, 1977) have found that data quality increases with age. Similarly, item nonresponse rates (Borgers, De Leeuw, and Hox, 2000; Haunberger, 2014), number of don't know responses (Amato and Ochiltree, 1987; Vogl, 2012), and response order effect (Fuchs, 2005) have been found to decrease with age.

Since there is a growing interest in children's well-being, victimization, bullying, and risky behaviours, children are included as respondents in a number of repeated cross-sectional studies (e.g. the British Crime Survey or the National Crime Victimization Survey) and cohort panels (e.g. the Longitudinal Study of Young People in England, the National Child Development Study, or the National Longitudinal Survey of Youth). In Europe there

are also plans to start an ambitious new longitudinal survey of children and young people (Goswami et al, 2016; Pollock et al, 2018). In 1994, the British Household Panel Survey (BHPS) was the first European household panel to ask household members aged 11-15 to complete the survey. In 2010-2011 the BHPS participants joined a larger panel, *Understanding Society: the UK Household Longitudinal Study* (UKHLS, see <https://www.understandingsociety.ac.uk>). The panel also continued to collect data from youth in the sample while extending the age range to 10-15. Every household member within this age range is asked to complete a self-administered paper questionnaire.

Though there are several papers about data quality among children and adolescents in both cross-sectional surveys (see Amato and Ochilree, 1987; Beebe et al, 1998; Fuchs, 2005; König, 2011; Mavletova, 2015a; Varni, Limbers, and Burwinkle, 2007; Vogl 2012, 2013) and longitudinal surveys (Borgers, Hox, and Sikkel, 2003, 2004; Fendrich and Kim, 2001; Fendrich and Rosenbaum, 2003; Halpern-Manners, Warren, and Torche, 2014; Haunberger, 2014; Mavletova, 2015b; Smith and Platt, 2013; Vaillancourt, 1977), to our knowledge none of the studies measured data quality in household panel studies among youth. Using data from a household panel survey that includes a youth self-completion component allows us to take advantage of the rich structure of the data to better explain variations in the quality of the youth data, using parental and household-level information as covariates (such variables as household income, maternal education, parental involvement in their children's education, and parent-child relationships). Furthermore, the large sample size of UKHLS allows us to identify age effects in terms of differences between single year cohorts, while relatively high wave-on-wave response rates provides good sample sizes for studying panel conditioning effects. We develop hypotheses regarding the correlates and nature of data quality in the youth survey data and test them using two indicators of data quality: item nonresponse rates (INR) and panel conditioning.

2. Literature Review and Hypotheses

Tourangeau, Rips, and Rasinski (2004) proposed a model of the survey response process with four components: comprehension of the question, retrieval of relevant information, making judgements based on retrieved information, and mapping judgement into response category. Each of the components can be a source of measurement error. For instance, respondents can have difficulties understanding the question or some terms; difficulties in recalling some events; may not have opinion on some attitude questions, or may not be willing to respond

truthfully on some sensitive questions. This model can be applied to both adults and adolescents.

Some researchers found that expressing of “don’t know” responses may be socially undesirable for children. They tend to either skip the question or give an answer rather than explicitly state they “don’t know” (Haunberger, 2014; König, 2010; Scott, 2008). Though in some waves up to 20% of the questions in the UKHLS youth questionnaire offered an explicit “don’t know” response category, only one question (in the second wave) actually produced any “don’t know” responses. As a result, we will focus on the analysis of INR. INR in surveys of children has been found to be explained more by characteristics of the children than by characteristics of the question (Borgers and Hox, 2001) and due to difficulties that children have in retrieval of relevant information or due to lack of knowledge or opinion (Vogl, 2012). Age and academic achievements explain differences in INR between children (Borgers, De Leeuw, and Hox, 2000; Haunberger, 2014; Mavletova, 2015a; Yi and Lee, 2016). In addition, children from families with a higher social status produce lower INR (Yi and Lee, 2016). A higher number of response options induce higher INR, while knowledge questions and the length of the introductory text in the question reduce INR (Borgers and Hox, 2001). A higher INR was found to be at the end than at the beginning of the questionnaire (Borgers and Hox, 2001) and in open-ended rather than closed questions (Smith and Platt, 2013).

However, little is known about how children’s INR is related to household or parental characteristics. This may be partly because surveys of children often do not collect rich data about the household context. But for surveys of children that are carried out in the home environment, and therefore typically require the co-operation and permission of a parent, knowledge of the role of household characteristics could help researchers to identify ways to improve data quality in surveys of children. Taking into consideration survey data we have from the parental questionnaires, we examine how INR is associated with relevant covariates such as household income, maternal education, maternal attitudes towards child’s education, parental involvement in their children’s education, and parent-child relationships. These covariates are particularly relevant as previous studies have shown them to be associated with the academic achievements of children, which in turn may be associated with the quality of survey data provided. The influence on children’s academic achievements has been shown for parental education (Haveman and Wolfe, 1995), household income (Dahl and Lochner, 2012; White, 1982), parental attitudes towards education (Davis-Kean, 2005; Seginer, 1983),

parental involvement in education; and the quality of parent–child relationship (Davis-Kean, 2005). The literature is not consistent on whether maternal or paternal education is more important for a different number of cognitive abilities and behavioural outcomes of a child, however maternal education and maternal expectations usually has a consistently strong effect on a wide range of child’s outcomes as mothers typically provide the main care for a child since their birth (Chevalier et al, 2013).

Another aspect of data quality which is crucial for panel surveys is panel conditioning. Panel conditioning occurs when the act of participating in the survey influences the responses given by respondents at later waves. This can be caused either because survey participation changes respondents’ attitudes and behaviours or because it changes the way they answer questions (Lynn, 2009; Sturgis, Allum, and Brunton-Smith, 2009; Warren and Halpern-Manners, 2012). A number of researchers emphasize a lack of theoretical foundation in the literature of the conditions under which panel conditioning can or can not occur (Cantor, 2008; Warren and Halpern-Manners, 2012).

Some researchers found changes in attitudes or behaviour as a result of survey participation (e.g. Bartels, 1999; Clausen, 1968; Crossley et al, 2017; Kraut and McConahay, 1973; Spangenberg and Obermiller, 1996). The changes are likely to happen when survey participation stimulate individuals to think more about the topic and provide some knowledge about the issue (Sturgis, Allum, and Brunton-Smith, 2009), or motivate to engage into the behaviour (Warren and Halpern-Manners, 2012). However, a number of researchers found no such effect (e.g. Corder and Horvitz, 1989; Mann, 2005; Smith, Gerber, and Orlich, 2003; Toepoel, Das, and van Soest, 2009). The possible explanation of having no panel conditioning effect is that either the issue is seen as unimportant for respondents or, on the contrary, the issue is seen as highly important, which means that respondents have formed crystallized attitude (Bridge et al, 1977).

The second type of panel conditioning – changes in survey response behaviour –can be either positive or negative in the sense that it may either increase or decrease measurement error in later waves. Several studies showed positive panel conditioning among adults, when respondents provide data of higher quality at later waves as indicated, for example, by lower INR, less severe rounding, and higher reliability in subsequent waves (Waterton and Lievesley, 1989; Rendtel et al, 2004; Schräpler, 2003; Sturgis, Allum, and Brunton-Smith, 2009). These changes appear to be due to increased understanding of the survey process and

response tasks. An alternative reason for respondents to provide more accurate responses across waves is that their trust in the survey organization (and possibly the interviewer) may increase (Fisher, 2016; Halpern-Manners, Warren, and Torche, 2017).

Panel conditioning can also produce an increase over waves in measurement error. This is associated with the third and fourth components of the response process, making judgements based on retrieved information, and mapping the judgement to a response category. Such negative conditioning can arise either because respondents become less willing to respond truthfully (for social desirability or other reasons) or because they become less willing to put adequate effort into responding (or learn how to reduce their effort) (e.g. Bailar, 1989; Cantor, 1989; Cohen and Burt, 1985; Duan et al, 2007; Warren and Halpern-Manners, 2012).

There is some evidence of a higher social desirability bias in responses in subsequent waves among adolescents while answering sensitive questions on illicit behaviours. Fendrich and Kim (2001) found lower level of reporting of lifetime marijuana and cocaine use in 1988, 1992 and 1994 compared to 1984 in the National Longitudinal Survey of Youth. About 40% of participants denied lifetime cocaine use in one of the following waves. Black and married respondents had higher probability of denying. Similar results were found by Fendrich and Rosenbaum (2003) as well as by Torche, Warren, and Halpern-Manners (2012).

In line with these studies we expect a negative panel conditioning effect in terms of increasing social desirability bias - a lower level of reporting of sensitive behaviours and inconsistency in reporting lifetime smoking and alcohol consumption throughout waves. We also suggest that this effect would be stronger among children who live in households with lower socio-economic status, lower parental involvement in the education, lower expectations, and a less close parent-child relationship (*Hypothesis 1*).

We also test if there is a panel conditioning effect in terms of increasing or decreasing measurement error defined as correlations between residual variances of latent constructs throughout waves. According to the literature there can be a positive conditioning effect when respondents better understand the questions and survey process and are motivated to invest some physical and cognitive efforts to provide more accurate responses (*Hypothesis 2A*). Alternatively, there might be a negative panel conditioning effect if respondents learn to minimize their effort while answering survey questions (*Hypothesis 2B*). Since previous results showed that both increasing and decreasing measurement error can occur as a result of panel conditioning, we test two alternative hypotheses.

3. Data

The UKHLS is a multi-purpose longitudinal study based on a sample of around 100,000 individuals representing the UK population. It provides a major data resource for research in the social sciences (Buck and McFall, 2012). The sample includes a representative sample of the UK population in 2009, an ethnic minority boost, and the BHPS sample. Our analysis uses data from four waves of the study: the first wave was conducted in 2009-2010, the second in 2010-2011, the third in 2011-2012, and the fourth in 2012-2013. At each wave, every child within the age range 10-15 is asked to complete a paper self-completion questionnaire. The youth questionnaire self-completion rate was 74-75% in participating households in all four waves. There were 4,899 completed interviews in the first wave; 5,020 in the second wave; 4,427 in the third wave; and 4,049 in the fourth wave. In each wave about 50% of youth respondents were boys and the average age of the respondents was about 12.5 years old in all waves (see Table 1 in Online Appendix, unweighted statistics).

Questionnaires

The total number of items varied from 88 to 104 in different waves (see “Questionnaires” section and Table 2 in Online Appendix for more details). Some questions were included in each wave, while some were asked biennially. The questionnaires in the second and fourth waves were more cognitively demanding than the questionnaires in the first and third waves as they had more open-ended and sensitive items.

Procedures and measures

We measured data quality based on the following indicators (see “Procedures and measures” section in Online Appendix for more details):

Hypothesis 1:

- The overall item nonresponse rates (INR).

We conducted a multiple linear regression with pairwise deletion to predict the overall INR in each wave. We included such predictors as mother’s expectations, parent–child relationship reported by mother, socio-demographic variables, and some household variables.

Hypothesis 2:

- Panel conditioning effect

There are two indicators:

1) Social desirability bias:

- Inconsistency: the analysis of such items as having ever drunk alcohol and having ever smoked. We ran multiple logistic regressions to predict inconsistent responses for each of the two questions separately. We included only those respondents who completed at least two waves of the study.
- Level of reporting of sensitive behavior: the analysis of such items as having ever drunk alcohol, having ever smoked; playing truant in the last 12 months.

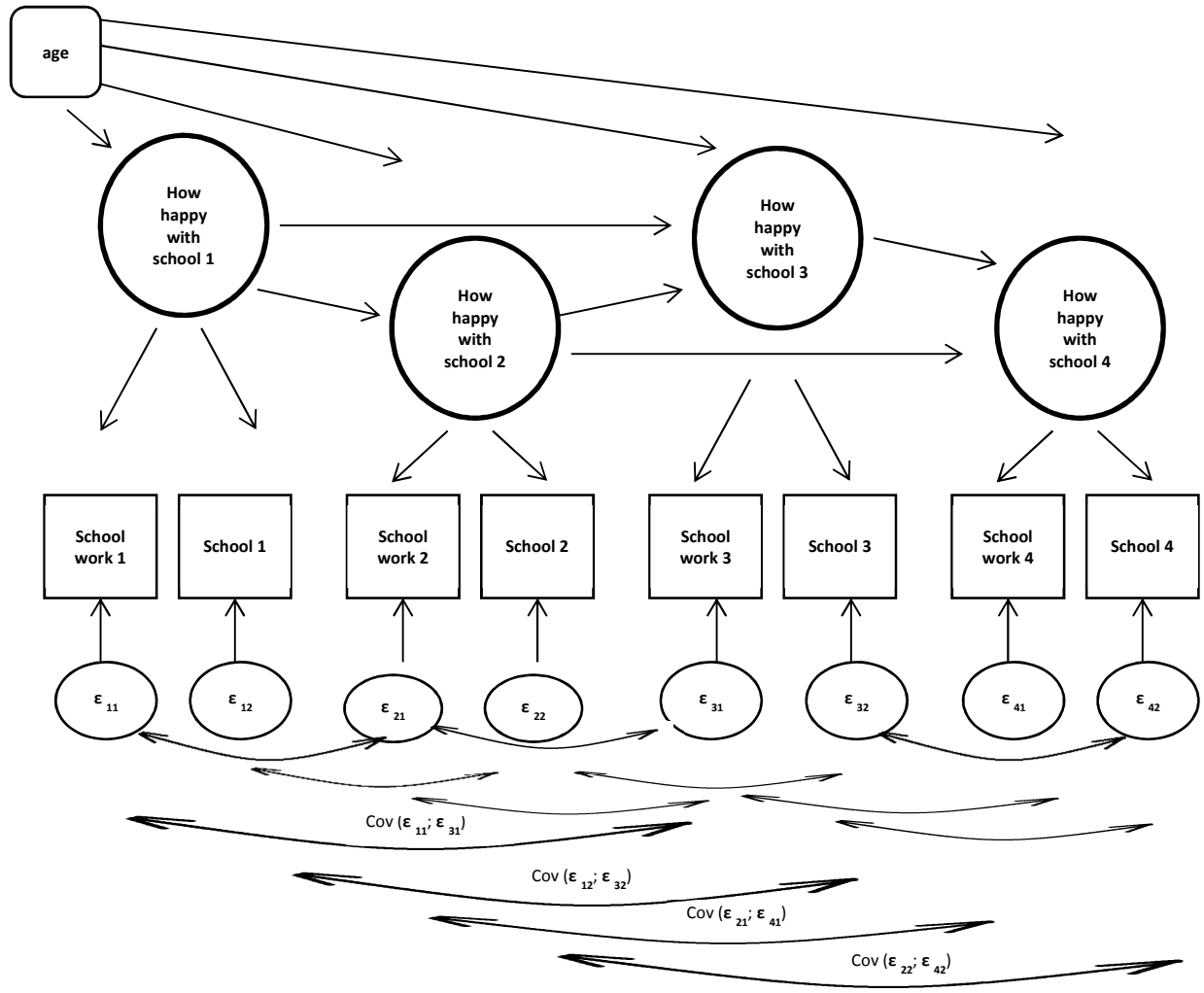
To measure an increase in social desirability bias in the subsequent waves we ran mixed-effects logit regressions. To disentangle panel conditioning effect from the attrition we included only those who completed all four waves (balanced panel analysis, N=960).

2) Correlations between residual variances

We use attitude questions with a Likert-type 5-point or 7-point scale throughout all waves. Several items measured how happy adolescents are with their appearance, family, friends, school, school work and life overall on a 7-point scale. We measured panel conditioning based on the analysis of latent construct of happiness with school measured by two indicators: how happy children are with school work and how happy they are with school overall. We ran a structural equation model and focus on the correlations between residual variances. The model assumed autoregressive change (see Alwin, 2007; Cernat, 2015), namely that the true score for happiness with school at time (i) is influenced by the true score at time (i-1) and at time (i-2) (see Fig. 1, for more details see “Procedures and measures” section in Online Appendix). The correlations between residual variances at time (i) and at time (i-1), as well as at time (i) and at time (i-2) were estimated. An increase from wave to wave in the correlations between residual variances would indicate an increasing measurement error which means a negative panel conditioning effect. Alternatively, a decreasing measurement error from wave to wave would indicate a positive conditioning effect. Since age has an effect on responses we controlled for age in each wave. We used lavaan package in R software environment for estimating the model (see <http://lavaan.ugent.be>). We included only those who completed all four waves in the analysis (N=960).

Fig.1. Model: autoregressive change, correlations between residual variances at time (i), (i-1), and (i-2)

Goodness-of-fit of the model: CFI=0.992, TLI=0.967, RMSEA=0.044, SRMR= 0.019



4. Results

4.1 Item nonresponse rates

Mean INR was 2.1% in the first wave (SD=4.9), 4.7% in the second wave (SD=6.5), 3.4% in the third wave (SD=4.4), and 5.4% in the fourth wave (SD=8.6). The second and fourth waves had higher INR. The highest INRs were in the cognitively demanding open-ended questions in the third wave: total amount of received money to spend on oneself last week (82.4%) and earned money if respondent had a paid job (52.7%). However, most questions with high INR were in the second and fourth waves. For questions about the respondent's height and weight the INR varied from 37% to 42%. Other questions with high INR were sensitive items in which respondents were expected to evaluate the ease of obtaining cannabis (33-35%), perceived risk of trying amphetamine (21%-26%) and ecstasy (19%-22%). Respondents might not have enough knowledge to evaluate the risk of taking some of these drugs, so this INR may be associated with the comprehension stage of the response process.

We ran a multiple linear regression with pairwise deletion to predict the overall INR in each wave. The R squared was quite small and varied from 0.007 to 0.063 depending on the wave. Overall, it was higher in the more demanding second and fourth waves (0.041-0.063) than in the first and third waves (0.007-0.034). As expected, older adolescents had lower INRs. Each additional year of age reduced the INR on average by 0.09-0.17 percentage points in the first and third waves and by 0.67-0.78 percentage points in the second and fourth waves (see Table 1). Boys tend to produce higher INR than girls ($\beta=0.57-0.89$). Mother's race had also a significant effect on INR. Children who have Caucasian mother had lower INR by 0.91-1.54 percentage points. The indicator of the parent-child relationships which showed a consistent effect was discussing books at home with children, giving them books as presents, or taking children to museums and theatres: the more often parents do that, the lower INR children produced ($\beta=-0.16-0.20$). No other variables produced a consistent effect throughout waves, though helping with homework, maternal education, mother's employment status, and living in a house owned by the household decreased INRs among children in some of the waves (Table 1).

Table 1: Predicting Item Nonresponse Rates: OLS Regression Coefficients

	Wave 1	Wave 3	Wave 2	Wave 4
Intercept	5.28*** (0.73)	5.24*** (0.72)	18.21*** (1.04)	20.31*** (1.63)
<i>Respondent's characteristics</i>				
Male	0.57*** (0.15)	0.24 (0.14)	0.19 (0.19)	0.89** (0.28)
Age	-0.17*** (0.04)	-0.09* (0.05)	-0.67*** (0.06)	-0.78*** (0.10)
<i>Socio-economic status</i>				
Gross household income (month before interview)	-0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	0.00* (0.00)
Own home	-0.37* (0.18)	-0.28 (0.18)	-0.45 (0.24)	-0.49 (0.35)
Mother has diploma in higher education	-0.05 (0.17)	-0.28 (0.21)	-0.64* (0.27)	-0.08 (0.48)
<i>Mother expectations (waves 1 and 3)</i>				
Importance for your child to complete A level exams: very important	0.02 (0.17)	-0.27 (0.16)		
<i>Parental involvement in education</i>				
My parents are interested in how I do at school: always or nearly always	-0.35 (0.21)	-0.27 (0.19)		
My parents come to school parent evenings: always or nearly always	-0.35 (0.21)	0.03 (0.21)		
Mother helps with homework: once a week or more often	-0.49** (0.17)	0.02 (0.16)		
Someone at home helps with homework			-0.13 (0.30)	-0.43 (0.43)
<i>Parent-child relationship</i>				
Spending time together on leisure activities with mother: several times a week or almost every day	0.14 (0.18)	0.18 (0.13)		
Quarrel with child: less than once a week	0.15 (0.15)	-0.19 (0.14)		
The child talks with mother about things that matter to him/her: most days	0.30 (0.16)	-0.19 (0.15)		

Discussing books at home, discussing TV programmes, buying books as gifts etc.			-0.16*** (0.03)	-0.20*** (0.04)
<i>Household socio-demographic variables</i>				
Parents are married and live together	-0.50** (0.17)	-0.01 (0.17)	-0.31 (0.23)	0.29 (0.34)
Number of children under 15 in the household	0.22** (0.07)	0.02 (0.07)	-0.01 (0.07)	0.03 (0.15)
Mother's employment status: employed/self-employed	-0.22 (0.18)	-0.06 (0.17)	-0.55* (0.23)	-0.22 (0.34)
Mother's race: white	-1.00*** (0.19)	-0.05 (0.20)	-1.54*** (0.26)	-0.91* (0.38)
Urban area	-0.04 (0.20)	-0.20 (0.17)	-0.43 (0.24)	-0.91** (0.35)
N	4,899	4,427	5,020	4,049
R ²	0.034***	0.007*	0.063***	0.041***

*p<0.05, **p<0.01, ***p<0.001, standard errors in parentheses

The effect of gross household income (month before interview) is too small to be displayed

4.2 Panel conditioning

4.2.1 Social desirability

- Inconsistency in sensitive questions

About 9% of respondents denied ever smoking cigarettes and 30% denied ever drinking alcohol subsequent to a previous response indicating that they had smoked cigarettes or drunk alcohol respectively. Older respondents (OR=1.33-1.57) and those who had a Caucasian mother (OR=1.69-1.81) had higher odds of providing inconsistent responses across waves (see Table 4 in Online Appendix). Greater parental involvement in education when they come to parent evenings (OR=0.63-0.78), are interested in the child's academic achievements (OR=0.73) or help to do homework (OR=0.67-0.82) as well as closer parent-child relationship (OR=0.95) decrease the odds of producing inconsistent responses (Table 4 in Online Appendix). Those who live in a home owned by the household and with both parents also have lower odds of providing inconsistent responses.

- Level of reporting of sensitive behaviors

Amongst the balanced panel, each sensitive item had a higher level of reporting in each subsequent wave except for the level of lifetime drinking: while 50.7% of the participants reported lifetime drinking in the third wave, 44.8% reported it in the fourth wave (see Table 2). These differences could of course be caused by ageing (each respondent is three years older at wave 4 than they were at wave 1), so to identify any panel conditioning effect we ran mixed effects logit models to predict the level of reporting of sensitive behaviour, controlling for a number of socio-demographic variables including age (data not shown).

We found that panellists were less willing to report lifetime drinking in the fourth wave compared to the third wave (OR=2.21, $p < 0.001$). It is consistent with our previous finding that 30% denied drinking alcohol in the subsequent waves. No social desirability bias was found in other variables. Overall, we found some evidence of increasing social desirability bias in the panel.

Table 2: Level of Reporting of Sensitive Behaviours

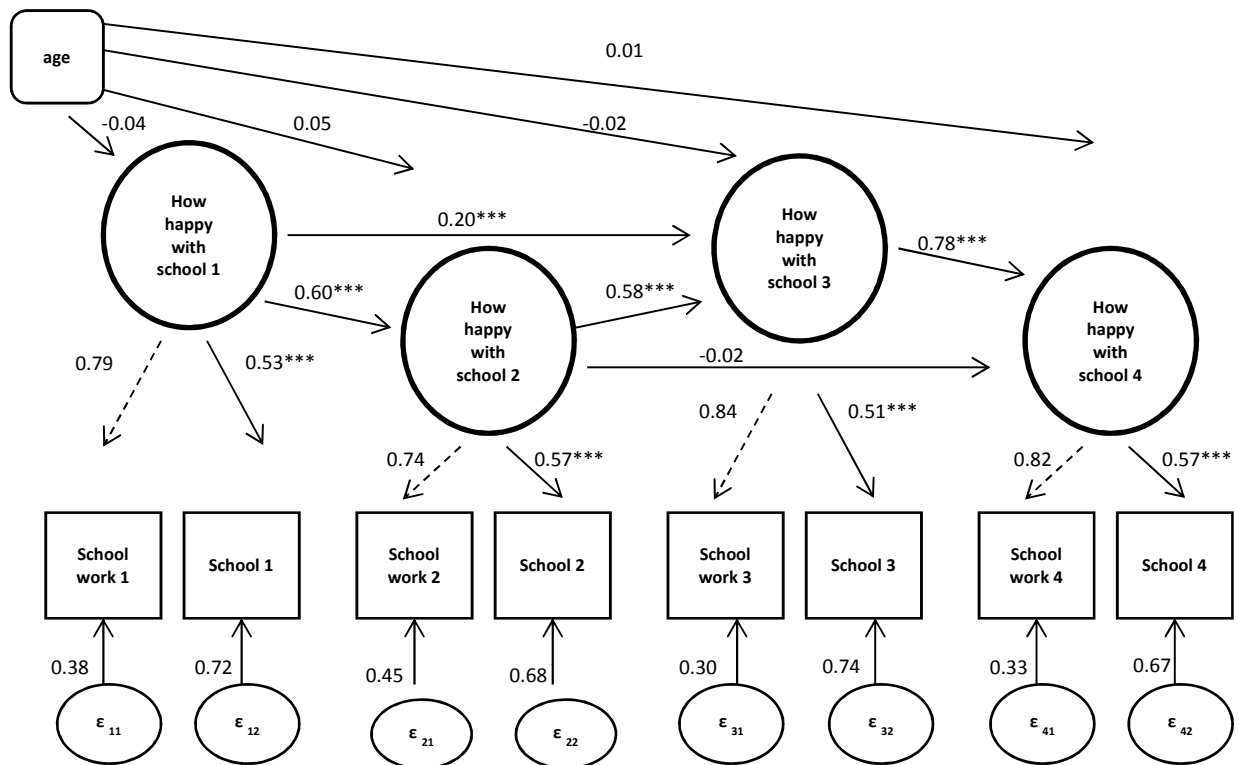
	Wave 1	Wave 2	Wave 3	Wave 4
Have ever drunk alcohol	10.4%	19.4%	50.7%	44.8%
N	948	945	954	948
Have ever smoked	1.5%	3.6%	8.1%	9.9%
N	953	946	949	946
Played truant in the last 12 months	4.9%	5.4%	6.9%	10.2%
N	952	943	944	941

4.2.2. Correlations between residual variances

We measured panel conditioning with respect to two items: how happy children are with school work and how happy they are with school overall. The standardized factor loadings for the variable of school work varied from 0.74 to 0.84, while for being happy with school the loadings were lower and varied from 0.51 to 0.57 (see Fig. 2, correlations between residual variances at time (i), (i-1), and (i-2) are excluded from the diagram). We expected changes in the correlations between residual variances from wave to wave as an indicator of a panel conditioning effect. We found no pattern in the correlations between residual variances in the item on school work, but increasing correlation over waves for the item on being happy with school: the correlation between the first and second waves was 0.28, between the second

and third waves 0.46, and between the third and fourth waves 0.64 (see Table 3). This suggests a negative panel conditioning effect when respondents provide less accurate responses throughout waves.

Fig.2. Path diagram for how happy children are with school (correlations between residual variances at time (i), (i-1), and (i-2) are excluded from the diagram): standardized coefficients



***p<0.001

Table 3: Correlations between residual variances

<i>School work</i>	Wave 1	Wave 2	Wave 3	Wave 4
Wave 1		0.19** (0.07)	0.01 (0.06)	-
Wave 2			-0.02 (0.09)	0.04 (0.05)
Wave 3				-0.02 (0.10)
<i>School</i>	Wave 1	Wave 2	Wave 3	Wave 4
Wave 1		0.28*** (0.06)	0.09 (0.05)	-
Wave 2			0.46*** (0.07)	0.44*** (0.06)
Wave 3				0.64*** (0.08)

p<0.01, *p<0.001, standard errors in parentheses

5. Discussion

Our findings revealed that the drivers and nature of INR are to some extent different amongst children, compared to adult survey respondents. Overall levels of INR are fairly low, but for questions on topics of which many children may have little knowledge (e.g. their own height and weight, amounts of money earned or received, the ease of obtaining drugs, risks of trying drugs), INR rates can be very high indeed. In combination with the observation that children hardly ever select a “don’t know” option, this appears to suggest that children tend to simply skip a question rather than admit explicitly that they do not know the answer. This raises the question of whether there is any value in offering explicit “don’t know” options in questionnaires for children. Moreover, researchers might consider filtering, or at least preceding, questions on topics that require some prior knowledge on questions that ask explicitly about knowledge levels.

Cognitive ability is also implicated as a driver of INR by some of our findings: INR was higher in the waves with more cognitively demanding questions and declined with increasing age of the children. This should serve as a reminder to researchers to be aware of the cognitive limits of children and to design age-appropriate survey questions.

While previous research (Borgers and Hox, 2001) found that both children’s characteristics and the characteristics of the questionnaire have an effect on data quality, we were able to identify an interaction effect between them. Younger respondents were heavily affected in

more cognitively demanding waves and for questions about cannabis, ecstasy, and amphetamine. Such questions are quite difficult for this age group and may need to include some age or knowledge filters. We can conclude that INR amongst children is likely more associated with difficulties in retrieving relevant information or making judgements than with difficulties in understanding the item or a tendency to avoid providing the answer.

A particular contribution of this study was to identify a strong effect on INR of the strength of the parent-child relationship, even after controlling for maternal education, maternal employment, parental involvement in education, household income, and age and gender of the child. This interesting finding certainly seems to warrant further investigation to establish the mechanisms through which the effect operates. It could be a rather direct effect, indicating that parents with a stronger parent-child relationship are more likely to be present while the child fills the questionnaire, and even to assist with the completion of the questionnaire. Or the mechanisms could be more subtle, indicating that stronger parent-child relationships are more likely to provide the child with skills and abilities that are relevant to the task of completing a social survey questionnaire. We found that inconsistency between responses to sensitive questions given at different waves depends on the question topic. A higher level of inconsistency (denial at a later wave of a behaviour admitted at an earlier wave) was found for drinking alcohol than for smoking. This is in line with the results of Fendrich and Kim (2001). Inconsistency was lower for children with a stronger parent-child relationship and higher parental involvement in education. Researchers should therefore be cautious in interpreting findings from questions about sensitive behaviours, particularly regarding the association of such behaviours with any factors likely to be associated with parent-child relationships or parental involvement: there could be correlated measurement error at play.

We found partial evidence for a negative panel conditioning effect. The effect was found for one of three sensitive measures and one of two subjective happiness measures. These findings suggest that both social desirability bias and satisficing can increase over waves of survey participation by children. While researchers should clearly be aware of this possibility, it remains unclear in which circumstances such panel conditioning effects should be expected.

Some limitations of our study should be noted. First, as with many panel conditioning studies, we note that the panel conditioning analysis is based on a non-experimental design and has a number of limitations as a result. Results would be more reliable if an experimental

design were to be embedded in the panel with an experimental group not receiving some survey items in the previous waves and a control group receiving survey items in all waves. Second, the panel conditioning effect in terms of correlations between residual variances was measured only for two items. It would be useful to replicate the kind of analysis we have proposed and implemented here for a number of other variables. Overall, we point to the need for further research in two areas. First, we lack understanding of the mechanisms that have led to some of our findings: for example, the finding of a positive effect of the strength of parent-child relationship on the quality of the data provided by the child. Studies designed specifically to identify the mechanisms would be helpful. Second, though our findings provide pointers to aspects of survey practice that could be improved, as discussed above, knowledge of how best to make these improvements is only partial. Methodological studies should be designed with a specific focus on identifying how panel conditioning can be reduced and how data quality can be increased especially in more socially vulnerable households with lower parental involvement in education and less close parent-child relationships.

Funding

The contribution of the first author is funded by The Russian Science Foundation, Grant № 17-78-20172.

The contribution of the second author is funded by the UK Economic and Social Research Council award for Understanding Society: the UK Household Longitudinal Study.

Note

Aigul Mavletova's visit to Essex University was supported within the framework of a subsidy by the Russian Academic Excellence Project '5-100' within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE).

References

Alwin, D. F. (2007). *The margins of error: A study of reliability in survey measurement*. Wiley-Blackwell.

Amato, P.R, and Ochiltree, G. (1997). Interviewing children about their families: A note on data quality. *Journal of Marriage and the Family*, 49, 669-675.

- Bailar, B. A. (1989). Information needs, surveys, and measurement errors. In *Panel Surveys*, Edited by D. Kasprzyk, G. Duncan, G. Kalton and M. P. Singh, pp. 1-24. New York: Wiley.
- Bartels, L.M. (1999). Panel effects in the American National Election Studies. *Political Analysis*, 8(1), 1-20.
- Beebe, T. J., Harrison, P.A., Mcrae, Jr., J. A., Anderson, R. E., and Fulkerson, J.A. (1998). An evaluation of computer-assisted self-interviews in a school setting. *Public Opinion Quarterly*, 62, 623–632
- Borgers, N., Hox J., and 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(1), 83-94.
- Borgers, N., Hox, J., and Sikkel, D. (2004). Response effects in surveys on children and adolescents: The effect of number of response options, negative wording, and neutral mid-point. *Quality & Quantity*, 38, 17–33.
- Borgers, N., and Hox, J. (2001). Item nonresponse in questionnaire research with children. *Journal of Official Statistics*, 17(2), 321-335.
- Borgers, N., De Leeuw, E., and Hox, J.P. (2000). Children as respondents in survey research: Cognitive development and response quality. *Bulletin of Sociological Methodology*, 66, 60–75.
- Bridge, R. G., L. G. Reeder, D. Kanouse, D. R. Kinder, V. T. Nagy, and Judd, C. M. (1977). Interviewing changes attitudes - sometimes. *Public Opinion Quarterly*, 41(1), 56-64.
- Buck, N. and McFall, S. (2012). Understanding Society: design overview. *Longitudinal and Life Course Studies* 3(1), 5-17.
- Cantor, D. (2008). A review and summary of studies on panel conditioning. In *Handbook of Longitudinal Research: Design, Measurement and Analysis*, Edited by S. Menard. Burlington, Academic Press, pp.123-138.
- Cantor, D. (1989). Substantive implications of selected operational longitudinal design features: The National Crime Survey as a case study. In *Panel Surveys*, Edited by D. Kasprzyk, G. Duncan, G. Kalton and M. P. Singh, pp. 25–51. New York: Wiley.

Cernat, A. (2015). The Impact of mixing modes on reliability in longitudinal studies. *Sociological Methods & Research*, 44(3): 427-457

Chevalier, A., Harmon, C., O' Sullivan, V., and Walker, I. (2013) The impact of parental income and education on the schooling of their children. *IZA Journal of Labor Economics*, 2: 8. <https://doi.org/10.1186/2193-8997-2-8>.

Clausen, A. (1968). Response validity: Vote report. *Public Opinion Quarterly*, 41, 56-64

Cohen, S. B., and Burt, V. L. (1985). Data Collection frequency effect in the National Medical Care Expenditure Survey. *Journal of Economic and Social Measurement*, 13(2), 125-151.

Corder, L. S., and Horvitz, D. G. (1989). Panel effects in the National Medical Care Utilization and Expenditure Survey. In *Panel Surveys*, Edited by D. Kasprzyk, G. Duncan, G. Kalton and M. P. Singh, pp. 304-318. New York: Wiley.

Crossley, T.F., de Bresser, J., Delaney, L. and Winter, J., (2017). Can survey participation alter household saving behaviour? *The Economic Journal*. 127 (606), 2332–2357.

Dahl, G. B., and Lochner, L. (2012). The impact of family income on child achievement: Evidence from the earned income tax credit. *American Economic Review*, 102(5), 1927-56.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19(2), 294–304.

Duan, D., Alegria, M., Canino, G., McGuire, T., and Takeuchi, D. (2007). Survey conditioning in self-reported mental health service use: Randomized comparison of alternative instrument formats. *Health Services Research*, 42(2), 890-907.

Fendrich, M., and Kim, J. Y.S. (2001). Multiwave analysis of retest artifact in the National Longitudinal Survey of Youth drug use. *Drug and Alcohol Dependence*, 62, 239–253.

Fendrich, M., and Rosenbaum, D. P. (2003). Recanting of substance use reports in a longitudinal prevention study. *Drug and Alcohol Dependence*, 70, 241-253.

Fisher, P. (2016) Does repeated measurement improve income data quality? ISER Working Paper. No. 2016-1. Colchester: University of Essex.

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(4), 701-725.

Goswami, H., Fox, C., and Pollock, G. (2016) The current evidence base and future needs in improving children's well-being across Europe: is there a case for a comparative longitudinal survey? *Child Indicators Research* 9(2), 371-388.

Halpern-Manners, A., Warren, J. R., and Torche, F. (2017). Panel conditioning in the General Social Survey. *Sociological Methods & Research*, 46(1), 103-124.

Haunberger, S. (2014). Item nonresponse in face-to-face interviews with children. *Journal of Official Statistics*, 30, 459–477.

Hox, J.J., de Leeuw, E.D., and Chang, H.-T. (2012). Nonresponse versus measurement error: Are reluctant respondents worth pursuing? *Bulletin de Methodologie Sociologique*, 113, 5-19.

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

Kraut, R. E., and McConahay, J. B. (1973). How being interviewed affects voting: An experiment. *The Public Opinion Quarterly*, 37(3), 398-406.

Lynn, P. (2009) 'Methods for longitudinal surveys', chapter 1, 1-19, in *Methodology of Longitudinal Surveys* (ed. P. Lynn), Chichester: Wiley.

Mann, C. B. (2005). Unintentional voter mobilization: Does participation in preelection surveys increase voter turnout? *Annals of the American Academy of Political and Social Science*, 601, 155–168.

Mathiowetz, N. A., and Lair, T. J. (1994). Getting better? Change or error in the measurement of functional limitations. *Journal of Economic and Social Measurement*, 20(3), 237-262.

Mavletova, A. (2015a). Web surveys among children and adolescents: Is there a gamification effect? *Social Science Computer Review*, 33(3), 372-398.

Mavletova, A. (2015b). A gamification effect in longitudinal web surveys among children and adolescents. *International Journal of Market Research*, 57(3), 413–438.

Piaget, J. (1929). Introduction to the child's conception of the world. New York: Harcourt.

- Pollock, J. Ozan, H. Goswami, G. Rees & A. Stasulane (2018). *Measuring Youth Well-being: How a Pan-European Longitudinal Survey Can Improve Policy*. Berlin: Springer.
- Rendtel, U., Nordberg, L., Jäntti, M., Hanish, J., and Basic, E. (2004). Report of quality of income data. Chintex. Working paper 21.
- Schräpler, J-P. (2003). Respondent behaviour in panel studies: a case study for income-nonresponse by means of the British Household Panel Study. ISER Working Paper Series, No. 2003-08.
- Scott, J. (2008). Children as respondents: The challenge for quantitative methods. In *Research with children: Perspectives and practices*, Edited by P. Christensen, and A. James, pp. 87-108. New York: Routledge.
- Seginer, R. (1983) Parents' educational expectations and children's academic achievements: A literature review. *Merrill-Palmer Quarterly*, 29(1), 1-23.
- Smith, J.K., Gerber A. S., and Orlich A. (2003). Self-prophecy effects and voter turnout: An Experimental replication. *Political Psychology*, 24(3), 593-604.
- Smith, K., and Platt, L. (2013). How do children answer questions about frequencies and quantities? Evidence from a large-scale field test. Centre for Longitudinal Studies.
- Spangenberg, E. R., and Obermiller, C. (1996). To cheat or not to cheat: Reducing cheating by requesting self-prophecy. *Marketing Education Review*, 6(3), 95-103.
- Sturgis, P., Allum, N., and Brunton-Smith, I. (2009). Attitudes over time: The psychology of panel conditioning. In *Methodology of Longitudinal Surveys*, Edited by P. Lynn. John Wiley & Sons, pp. 113-126.
- Toepoel, V., Das, M., and van Soest, A. (2009). Relating question type to panel conditioning: Comparing trained and fresh respondents. *Survey Research Methods* 3(2), 73-80.
- Torche, F., Warren, J.R., and Halpern-Manners, A. (2012). 'Panel conditioning in a Longitudinal Study of Chilean Adolescents' Substance Use: Evidence from an experiment. *Social Forces*, 90, 891-918.
- Tourangeau, R., L. Rips and K. Rasinski (2004). *The Psychology of Survey Response*. Cambridge University Press.

Triplet T., Blair, J., and Kang, Y.C. (1996). Initial cooperators vs. converted refusers: Are there response behavior differences? Proceedings of the Survey Research Methods Section, American Statistical Association, 1038-41.

Vaillancourt, P.M. (1977). Stability of children's survey responses. *Public Opinion Quarterly*, 37(3), 373-387.

Vogl, S. (2012). Children between the age of 5 and 11: what “don’t know” answers tell us. *Quality & Quantity*, 46(4), 993–1011.

Vogl, S. (2013). Telephone versus face-to-face interviews: A mode comparison of semi-structured interviews with children. *Sociological Methodology*, 43(1), 138-182.

Waterman, A.H., Blades, M., and Spencer, C. (2001). Interviewing children and adults: the effect of question format on the tendency to speculate. *Applied Cognitive Psychology*, 15(5), 521–531.

Waterton, J., and Lievesley, D. (1989). Evidence of conditioning effects in the British Social Attitudes Panel Survey. In *Panel Surveys*, Edited by D. Kasprzyk, G. Duncan, G. Kalton and M. P. Singh, pp. 319-339. New York: Wiley.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91(3), 461-481.

Yi, Y., and Lee, B. (2016). Adolescents' characteristics associated with item nonresponse on BMI questions. *Health Behavior and Policy Review*, 3(6), 597-606.