Item Nonresponse in Face-to-Face Interviews with Children

Sigrid Haunberger[1]

This study examined item nonresponse and its respondent and interviewer correlates by means of a population-based, panel survey of children aged 8 to 11 who were surveyed using standardised, face-to-face interviews. Using multilevel, logistic analyses with cross-level interactions, this article aims to examine which effects of item nonresponse are subject to children as respondents or to the interviewers and the interview setting. Depending on the type of question, we found different effects for respondent and interviewer variables, as well as interaction effects between child age/interviewer age as well as child gender/interviewer gender. However, interviewer variance is for the most part not significant.

Key words: Panel survey; interviewer effects; interviewing children; item nonresponse; multilevel logistic analysis.

1. Introduction

1.1. Focus on Children As Respondents in Social Research

Today, children are seen as independent individuals in social survey research and no longer an ignored minority. Survey researchers interested in the growing-up, perspectives, attitudes, beliefs, and behaviour of children increasingly collect data from children themselves. Proxy reporting by parents or other caregivers is no longer seen as a suitable and satisfactory mode of data collection (Scott 1997). This is exemplified by the many child surveys where children's opinions and attitudes are collected using different modes of data collection and over a different period of time: for example, the Child Longitudinal Study (Germany), the Child Survey (Austria), the British Household Panel Study (Great Britain), the Young People's Social Attitudes Survey (Great Britain), the National Longitudinal Study of Children and Youth (Canada), the Child Development Supplement to the Panel Study of Income Dynamics (United States), and the European Longitudinal Study of Pregnancy and Childhood, to mention just a few. Large-scale assessments like PISA (Programme for International Student Assessment), PIRLS (Progress in International Reading Literacy Study) or TIMSS (Trends in International Mathematics and Science Study) are also worth noting.

Although survey methodology has been developed mainly for studies in adult populations, research has been done on adapting it for use with children and evaluating the influence of their cognitive growth on data quality (see, for example, Borgers et al. 2000,

---

[1] University of Applied Sciences Northwestern Switzerland, School for Social Work, Riggenbachstr. 16, CH-4600 Olten, Switzerland. Email: sigrid.haunberger@fhnw.ch

Borgers et al. 2003). Item nonresponse in child surveys in general, and specifically in standardised face-to-face child interviews, however, has received only limited attention to date (see, for example, Borgers and Hox 2001, Fuchs 2008). Compared with self-completion questionnaires that target children as respondents, surveying children by means of interviews is of particular interest, because the interviewer and the interview situation may affect the young respondent's behaviour. We relate the frequency of item nonresponse on particular types of questions to the characteristics of the respondents (children), the interviewers, and the interview setting. The purpose of our main research is to discover whether and how child and interviewer characteristics as well as the interview setting affect item nonresponse in standardised, face-to-face interviews with children.

2. Past Research and Theoretical Framework

2.1. Past Research on Children As Survey Participants in General

This section briefly reviews past research on children as survey participants in general. An analysis by Borgers et al. (2003, p. 91) examined the correlation of child characteristics and offering vague quantifiers and labelled response options with stability over time. They found, contrary to their expectations, that younger children did not have more difficulty with cognitively challenging questions. The older the child, however, the greater the stability was over time. Compared with younger children, older children can take greater advantage of fully labelled response options.

A methodological survey experiment on the effect of several question characteristics on the reliability of the responses conducted by Borgers et al. (2004) revealed no effects of negatively formulated questions on the reliability measures; the authors advised offering about four response options when children are respondents.

De Leeuw and Otter (1995) showed that a clear interaction existed between the age of children and the effect of ambiguous questions. Older children handled ambiguity much more easily than younger children.

Fuchs (2005, p. 701) examined several experiments on response order, question order, scale effects, and the effects of the numeric values associated with the response categories with children. His results indicated that younger and less educated children answered survey questions from a cognitively less ambitious perspective than adults did. In a later study by Fuchs (2008), the interviewer respondent interaction was videotaped, and all children underwent extensive cognitive tests. The results showed that younger children (ages 8–9) show considerably more problematic behaviours, suggesting problems in understanding and answering survey questions, than older respondents (ages 13–14).

2.2. Past Research on Item Nonresponse in Child Surveys

This section briefly reviews the state of knowledge in the field of item nonresponse in child surveys. Borgers and colleagues (1999) investigated the influence of child characteristics and cognitive growth on data quality when surveying children by means of meta-analytic techniques. They found that gender and year of education influenced item nonresponse and internal consistency in a large number of different, multi-item scales. The hypothesis that data quality increases with cognitive growth was supported.

Furthermore, Borgers and Hox (2001) investigated the effect of item and personal characteristics on item nonresponse in written questionnaires used with schoolchildren. They found that item nonresponse is relatively rare, and the predicted response differences are relatively small. They concluded that young children do not perform as well as children who have been in education longer (they produce more item nonresponse), but their response behaviour is still satisfactory.

With a more qualitative, semi-standardised approach, Vogl (2011) recently explored the question-answer process in child interviews (ages 5–11). Focussing on 'don't know' responses, the results indicated fewer 'don't know' responses as children grow older due to their cognitive state; problems with the research instrument did not result in differences in the number of 'don't know' responses.

Another analysis of an adult survey by Shoemaker et al. (2002) used question sensitivity and cognitive effort to distinguish between 'don't knows' and refusals. They found that more sensitive questions attracted more refusals, whereas questions that require more cognitive effort received more 'don't knows'. Note that cognitive effort also correlated significantly with refusals. There is also evidence of item nonresponse in the event that adult respondents do not have adequately precise answers (Juster and Smith 1997), or as Fuchs (2008) reasoned, children might answer survey questions even if they have problems processing them.

To summarise past research, we can state that younger children are able to answer survey questions in an appropriate way if survey instruments are tailored to them. Nevertheless, as children grow older, their ability to answer survey questions and to handle ambiguity increases. This is also evident from the fact that item nonresponse declines with increasing age and/or year of education in all of the studies mentioned above.

2.3. The Influence of Interviewers on Item Nonresponse in Child Surveys

In the special case of face-to-face interviews, the interviewer plays an important role in the question-answer process, even with children as respondents. Regarding item nonresponse in standardised, face-to-face interviews with adults, there is empirical evidence that interviewers are not neutral collectors of data but can influence the answers obtained (Pickery and Loosveldt 1998; Pickery and Loosveldt 2001). The interviewer can have a positive influence in reducing item nonresponse but may also induce item nonresponse (De Leeuw et al. 2003, p. 165). The results of comparisons of interviewer effects on factual and attitudinal questions in several studies are heterogeneous, with some of them finding that attitudinal measures are subject to higher interviewer variance. Greater effects for attitudinal questions have been found especially for questions with open-ended responses, emotionally charged questions, questions with difficult items (such as income or occupation), and questions that lack specification of an interviewing procedure (Groves 2004, p. 374). Findings on interviewer effects in adult surveys show that younger and less-educated interviewers have a higher level of item nonresponse (Huddy et al. 1997).

Item nonresponse is often the result of interaction between two sources of survey errors (Groves 2004), for instance the interaction between an interviewer and a respondent. Not much is known about how children react and behave face-to-face with a strange interviewer. There could be a huge social distance between young children and adult

interviewers. Therefore we assume that young children adapt their responses to the suspected expectations of adult interviewers and might have a tendency towards social desirability (De Leeuw et al. 2004).

2.4. The Interview Setting: Presence and Intervention of Third Parties During the Interview

The influence of third parties during the interview, especially parents, may bias responses from children in a positive or negative way: positively, if children are trying to answer the questions honestly and truthfully in the presence of their parents; negatively, if – especially as regards sensitive issues – the presence of parents or other persons leads to untruthful statements (Scott et al. 1995, p. 261; Reuband 1987). In general, the presence of third parties during standardised, face-to-face interviews is often undesirable, since researchers suspect there may be negative consequences for the question-answer process. Reuband (1984) reported a proportion of third parties during an average of about one third of adult interviews; similarly, Haunberger (2005) reported a high number of third parties present during standardised, face-to-face interviews with children, for the most part the children's parents (see Table 2 for details). Nevertheless, third parties may not necessarily act as a disrupting factor, but rather can exert a social control function and therefore contribute, especially in the case of factual questions, to more truthful answers (Reuband 1984).

2.5. Asking and Answering Survey Questions: Cognitive and Communicative Processes

The respondents' answers comprise a cognitive and communicative process (Schwarz and Sudman 1995). When answering survey questions, respondents must perform several tasks. First, they must interpret the question in order to understand what is meant, and second, they must retrieve relevant information. Third, they must integrate that information into a private opinion to finally formulate and edit a response (see Tourangeau et al. 2000 for details). This cognitive approach to the answering process shows that it is necessary to distinguish between different types of item nonresponse, which can have different causes and different meanings: Item nonresponse can easily occur when questions about events in the past are asked, or sensitive topics are probed, or when the questions are too difficult, uninteresting, too embarrassing, or too threatening.

   Middle childhood (ages 8–11) has been referred to as a pathway to future (cognitive) development. In the early middle years of childhood, children gradually increase their logical thinking, memory, and learning strategies, and consolidate important academic skills such as reading and writing. In the later middle years of childhood, children gradually expand their ability to apply learned concepts to new tasks and are increasingly interested in learning life skills from adults (Kail 2011). Therefore, answering a survey question in middle childhood might be a particular challenge, because children's cognitive, memory, communicative, and social faculties are still developing.

2.6. Towards a Theory of Item Nonresponse

The model of the response process posited by Beatty and Hermann (2002; also see Groves et al. 2009) distinguishes between four levels of cognitive states regarding information

required by survey questions: *available*, *accessible*, *generatable*, and *inestimable*. The four states are ordered by the level of retrieved knowledge suitable for a question response. If the required information can be retrieved with minimal effort, the substantive response is *available* or *accessible*. If the required information is not known exactly, the substantive response is barely *generatable* or completely *inestimable*, resulting in item nonresponse. Therefore a hypothetical question should be inestimable as it is based on assumptions rather than facts.

2.7. Research Question and Hypotheses

Given the background of the relevant research and theoretical assumptions, we want to investigate whether, and if so, how child and interviewer characteristics and the interview setting affect item nonresponse. For this purpose, several hypotheses have been developed.

Empirical evidence points to the fact that younger and less educated children produce more item nonresponses, leading to our first hypothesis.

Hypothesis 1: With increasing cognitive functioning (measured by age and educational achievement), item nonresponse in standardised, face-to-face interviews with children will be reduced.

In our next hypothesis, we specify a nondirectional premise, as research on interviewer characteristics influencing item nonresponse in standardised, face-to-face interviews with children is still lacking.

Hypothesis 2: Interviewer characteristics will influence the impact of item nonresponse in standardised, face-to-face interviews with children in different ways.

As we pointed out, we assume that it is primarily young children who adapt their responses to the supposed expectations of adult interviewers because of the huge social distance between them.

Hypothesis 3: Cross-level effects between child and interviewer characteristics (especially age) will influence the impact of item nonresponse in standardised face-to-face interviews with children in different ways.

Furthermore, we suppose that third parties during the interview act as mediators, especially for children in the presence of their parents trying to answer the questions honestly and truthfully, leading to our last hypothesis.

Hypothesis 4: Third parties during the interview will influence item nonresponses in standardised, face-to-face interviews with children.

3. Method

In this next section, the data, variables, and multilevel logistic analysis are introduced.

3.1. Data Set

The data used in the analyses come from the Child Longitudinal Study conducted by the German Youth Institute. They are based upon a prospective longitudinal survey with two national, representative group samples in the following age groups: children in the last year of kindergarten (five-year-olds) and second-year primary school children. Children in the older cohort (and their parents) were interviewed in three survey stages at intervals of approximately 18 months.

Table 1.   Child Longitudinal Study, sample size

	1st wave (2002)	2nd wave (2004)	3rd wave (2005)
Age group children	8–9	9–10	11–13
Interviewer (n)	96	54	51
Sample size (n)	1,042	722	620
Response rate	50.58%	35.05%	30.09%

Note: Gross sample $N = 2,060$

The first wave was conducted in the autumn of 2002, the second wave in the spring of 2004, and the third wave in the spring of 2005. As the study was not conducted for methodological purposes, a major drawback is that it was not possible to obtain measures of the interviewers' beliefs, expectations, and psychological characteristics or even to arrange an experimental setting. The sample size, the response rates and the number of interviewers for each wave are presented in Table 1.

3.2.   Variables

3.2.1.   Selection of Variables

In a first step, we calculated descriptive analyses for the whole data set to obtain a first impression of the distribution of item nonresponse, and subsequently made a preselection of variables. The following topic areas are addressed in the questionnaires: personality traits: 2 scales, the child's interests and activities: 3 scales, behaviour in conflict situations (with mother): 2 scales, school: well-being in school: 1 scale, parents' interest in school: 1 scale, achievement motivation: 1 scale, victims of violence: 1 scale, bullying: 1 scale, friends: child's network of friends: 1 scale, happiness with friends: 1 scale, behaviour in conflict situations (with friends): 1 scale, family climate: 1 scale, satisfaction with neighbourhood: 1 scale. All scales were asked over the three panel waves.

We found that the percentage of item nonresponse in this child survey is generally low, which creates two limitations for the selection of our dependent variables. On the one hand, we had to exclude questions with item nonresponse equal or less than 2 percent from the outset, on the other hand we were unable to follow the suggestion in the literature and distinguish between 'don't know' answers and refusals (Shoemaker et al. 2002). A separation between 'don't know' answers and refusals would have left too few cases for the analysis. Nevertheless, it was possible to select one scale from almost every topic area. This corresponds to a share of 40 percent of all scales in the questionnaire, which were used in the item nonresponse analyses.

3.2.2.   Linking Variables to the Model of the Response Process

In a second step, we linked the remaining variables with the Beatty-Hermann (2002) model of the response process. For self-description and leisure activities, we assumed that children were able to retrieve information with minimal effort (information available). The children's own achievement motivation, family climate, and behaviour in conflicts with friends could be retrieved with effort (information accessible) and represents a sensitive topic. Children might not have much knowledge of their parents' interest in school,

so information had to be estimated, resulting in a higher rate of item nonresponse (information generatable). We classified the question about the children's behaviour in hypothetical situations as inestimable.

In summary, our selected, dependent variables included questions about different topics. (See Appendix, Table A1 for question wording, response scale and percentage of item nonresponse per wave.)

3.2.3. Recoding the Dependent Variables

In a third step, all dependent response variables were dichotomised, resulting in scales with the categories adequate responses (0) and item nonresponse (1). Remember that our category for item nonresponse includes 'don't know' answers as well as refusals.

The dependent variables vary considerably in question length, sensitivity, and response scales. Obviously, the highest item nonresponse rate was found for questions offering an explicit 'don't know' category (child's rating of parental interest in school, child's achievement motivation). In any case, the main purpose of this article is to clarify whether and how child and interviewer characteristics and the interview setting affect item nonresponse in standardised, face-to-face interviews with children, and not to explain the amount of item nonresponse due to different response scales.

3.3. *Independent Variables*

The selection of the independent variables on the respondent and interviewer level was restricted due to the variables available in the existing data file and is largely based on the empirical evidence reported in Section 2.

3.3.1. Respondent Variables (Children)

On the respondent level, we included the following variables in the multilevel logistic analysis (see Table 2 for details):
Gender (girl: 0, boy: 1), *age* (metrical, centred around the grand mean), *educational achievement* (mean: marks in mathematics, language, and reading, running from very good: 4 to fail: 1, centred around the grand mean), social and cognitive open-mindedness, self-efficacy (strongly disagree: 1 to strongly agree: 4) (both mothers' estimations).

Interviewer rating: children's willingness to respond (low: 0, high: 1), *open-mindedness* (low: 0, high: 1), *concentration skills* (low: 0, high: 1) and *language skills* (poor: 0, good: 1). Interviewers rated children's abilities after completion of the interview on a 6-point scale (very good: 1 to very poor: 6), which was dummy coded by the author afterwards.

3.3.2. Interviewer Variables

On the interviewer level, we included the following variables in the multilevel logistic analyses, which were divided into two main categories (see Table 2 for details):
Interviewer characteristics: We applied a code indicating more than just one interviewer throughout the three waves to each response in each of the waves: *same or different interviewer* (different interviewer: 0, same interviewer in at least two panel waves: 1), *gender* (female: 0, male: 1), *age* (metrical, centred around the grand mean),

Table 2.   Child Longitudinal Study, independent variables

	1st wave (2002)	2nd wave (2004)	3rd wave (2005)
Respondent variables			
Age (mean/sd)[1]	8.5 (0.51)	9.5 (0.51)	10.5 (0.51)
Gender (boys)	51%	50%	48%
Educational achievement (mean/sd)[1]	1.73 (0.51)	1.87 (0.54)	1.95 (0.60)
Personality traits			
Self-efficacy (mean/sd)	1.86 (0.47)	1.86 (0.47)	1.86 (0.47)
Cognitive + social open-mindedness (mean/sd)	2.34 (0.45)	2.34 (0.45)	2.34 (0.45)
Interviewer rating			
Willingness to respond (good)	81%	87%	90%
Open-mindedness (good)	64%	70%	76%
Concentration (high)	50%	56%	65%
Language skills (good)	83%	86%	91%
Interviewer variables			
Age (mean/sd)[1]	41.5 (11.5)	48.7 (8.9)	50.5 (8.3)
Gender (male)	57%	51%	52%
Same interviewer (at least in 2 waves)	–	50%	50%
Presence of third: yes	84%	69%	49%
Intervention of third: yes	28%	12%	7%
Difficulties: yes	8%	5%	4%
Interviewer length (mean/sd)	42.7 (13.3)	39.6 (15.0)	45.8 (10.1)
Interviewer (n)	96	54	51
Sample size (n)	1042	722	620

Note: Educational achievement in the original version (4 = fail, 1 = very good), [1]for multilevel logistic analyses centred around the grand mean

Interview setting: presence of third parties during the interview (no: 0, yes: 1), *intervention of third party during the interview* (no: 0, yes: 1), *difficulties during the interview due to a third party being present* (no: 0, yes: 1)*, interview length* (metrical, in minutes).

3.4.   Multilevel Logistic Analysis

Multilevel analyses offer the best prospects to inspect interviewer effects on survey data because of the clustering of respondents by interviewers (Hox 2010).

In our case the use of a standard, two-level model would be inapplicable, since our dependent variables have binary outcomes: $Y = 1$ for item nonresponse, $Y = 0$ for response. With the software HLM 7.0 we specified a nonlinear analysis for binary outcomes. Therefore the binary outcome model uses a binomial sampling model and a logit link function (see Bryk and Raudenbush 2004 for details).

Before performing the multilevel analysis with panel data in HLM we reshaped the wide data files into long form, resulting in a pooled data set with 2,384 cases on each level. Level 1 missing data was automatically deleted when running the analyses. We controlled whether a correlation existed between the amount of item nonresponse in one panel wave per case and unit nonresponse in the following panel wave. We found no correlation,

which is not surprising given that the participation of the child is highly dependent on the participation of the parents.

After running the analyses, HLM offers different outputs (unit-specific model versus population-average model with robust standard errors). We present estimates of the population-average model with robust standard errors, since it is more appropriate for estimating the predicted population proportion and it is much less susceptible to misspecifications and distributional assumptions since it is based on generalised least squares estimation with robust standard errors (Zeger et al. 1988).

We present an example of model specification using the binary dependent variable 'family climate'. All respondent characteristics are included in Level 1 (see Equation 1). Subscripts *i* belong to the respondents and subscripts *j* to the interviewers.

Level 1 Model

$$
\text{Prob}(family\ climate_{ij} = 1|\beta_j) = \phi_{ij} \quad \log[\phi_{ij}/(1 - \phi_{ij})] = \eta_{ij}
$$

$$
\begin{aligned}
&= \beta_{0j} + \beta_{1j}^*(GENDER_{ij}) + \beta_{2j}^*(AGE_{ij}) + \beta_{3j}^*(EDUACHIEVEMENT_{ij}) \\
&+ \beta_{4j}^*(OPENMIND_{ij}) + \beta_{5j}^*(SELFEFFICACY_{ij}) + \beta_{6j}^*(WILLINGNESS_{ij}) \\
&+ \beta_{7j}^*(OPENMIND_{ij}) + \beta_{8j}^*(CONCENTR_{ij}) + \beta_{9j}^*(LANGUAGE_{ij})
\end{aligned} \tag{1}
$$

Interviewer characteristics and characteristics of the interview setting are included in Level 2. We specified a random intercept model, since only the parameters associated with the constant vary across interviewers. The residual at the interviewer level can be denoted as *u*.

In order to better disentangle the effect of the child's gender and the effect of the interviewer's gender due to item nonresponse, we included cross-level interactions on Level 2 (for example: β_{1j} represents the interaction between Level 1 variable 'gender of the child' and Level 2 variable 'gender of the interviewer') (see Equation 2).

Note that β_{3j} to β_{9j} represents the coefficients from Equation 1, without specifying an interaction effect.

Level 2 Model

$$
\begin{aligned}
\beta_{0j} =& Y_{00} + Y_{01}^*(INT\_SAME_j) + Y_{02}^*(INTGENDER) + Y_{03}^*(INTAGE_j) \\
&+ Y_{04}^*(INTLENGTH_j) + Y_{05}^*(THIRD\ PARTIES_j) + Y_{06}^*(INTERVENTION_j) \\
&+ Y_{07}^*(INTDIFFICULTIES_j) + u_{0j} \\
\beta_{1j} =& Y_{10} + Y_{11}^*(INTGENDER_j) \\
\beta_{2j} =& Y_{20} + Y_{21}^*(INTAGE_j); \\
\beta_{3j} =& Y_{30}
\end{aligned} \tag{2}
$$

$$\beta_{4j} = Y_{40}$$

$$\beta_{5j} = Y_{50}$$

$$\beta_{6j} = Y_{60}$$

$$\beta_{7j} = Y_{70}$$

$$\beta_{8j} = Y_{80}$$

$$\beta_{9j} = Y_{90}$$

4. Results

Table 3 reports the results of the multilevel logistic analyses for item nonresponse in standardised, face-to-face interviews with children. For the random part, we included values for the interviewer variance ($u0j$) in the table, which corresponds to the intercept-only model. To increase interpretability of interactions, the value zero must be meaningful and actually occur in the data. For age and educational achievement we accomplished this by centring both variables on their grand mean. For gender, females were zero-coded (Hox 2010, pp. 63–68). In each column, we reported the coefficients, *t*-ratio and asterisks as indicators of the level of significance. We explained results for all analyses separately, referring only to results reaching the $p < 0.05$ level.

We first look at the interviewer level. Item nonresponse in the question about *self-description* is only explained by the variable indicating the same interviewer in at least two waves. If the interview was conducted by the same interviewer, this increased item nonresponse in the question about self-description.

On the respondent level the child's age, concentration and language skills affected the amount of item nonresponse. With increasing age, concentration and language skills, item nonresponse decreases. In addition, a significant interaction effect appeared between child gender/interviewer gender; meaning girls and female interviewers produced less item nonresponse in the question about self-description. However, the variance at the interviewer level is not significant.

On the interviewer level, only interview length affected the number of item nonresponses to the question about *leisure activities*. Increasing length of the interview correlates positively with more item nonresponse. We are not able to specify a cause and effect relationship, since we cannot clearly determine whether increased interview length led to more item nonresponses or whether more item nonresponses led to an increased interview length. On the respondent level, we found two significant effects: With increasing age, item nonresponse decreases. Children with good concentration skills produced more item nonresponses if they were asked about their hobbies. Taking a look at the interaction effect, the coefficient of child age/interviewer age is positive and statistically significant; meaning that with an increase in the age of the child and the interviewer, more item nonresponse occurs for this question. Again, the interviewer variance is not significant.

For *achievement motivation* we found only two significant effects at the interviewer level. Item nonresponses to the question about *achievement motivation* are due to the

Table 3. Results for HLM non-linear models with the logit link function for item nonresponse in standardised, face-to-face interviews with children

	Self description			Leisure activities			Achievement motivation			Parents interest school			Family climate			Behaviour: conflict with friends			Behaviour: hypothetical situations		
	Coeff.	T-ratio	Sig.	Coeff.	T-ratio	Sig.	Coeff.	T-ratio	Sig.	Coeff.	T-ratio	Sig.	Coeff.	T-ratio	Sig.	Coeff.	T-ratio	Sig.	Coeff.	T-ratio	Sig.
Interviewer level																					
Gender: male	0.05	0.18		0.07	0.27		0.26	1.57		0.15	0.95		−0.08	−0.21		−0.01	1.48		0.09	0.38	
Age in years	0.02	1.86	†	0.01	1.01		0.01	2.30	*	0.00	−0.13		0.07	4.63	***	0.01	1.48		0.00	0.59	
Interviewer: same	0.61	2.88	**	0.22	0.98		0.08	0.65		0.00	0.04		1.71	4.54	***	0.40	1.83	†	0.23	1.07	
Interview situation																					
Presence of third parties: yes	−0.18	−0.83		−0.19	−0.80		−0.14	−1.08		−0.01	−0.05	†	0.25	0.78		0.00	0.01		−0.30	−1.60	
Intervention of third parties: yes	−0.21	−0.99		0.22	0.99		0.01	0.11		0.05	0.38		0.12	0.40		0.67	3.27	**	−0.33	−1.76	†
Difficulties: yes	−0.31	−0.83		0.04	0.11		0.59	2.75	**	0.22	1.00		0.33	0.77		0.63	1.91	*	−0.26	−0.64	
Interview length in minutes	−0.01	−1.24		0.02	3.47	***	0.00	0.13		0.01	1.70		−0.02	−2.02	*	−0.02	−1.75	†	0.01	1.25	
Respondent level																					
Gender: boy	0.15	0.54		−0.56	−1.83	†	−0.01	−0.03		−0.14	−0.84		0.30	0.79		0.09	0.35		0.03	0.11	
Age in years	−0.45	−3.69	***	−0.44	−3.35	***	−0.13	−1.96	*	−0.11	−1.72	†	−0.40	−2.37	**	0.09	0.29		0.06	0.59	
Educational	−0.30	−1.48		−0.19	−1.02		−0.34	−2.99	**	−0.37	−3.24	**	−0.30	−1.29		−0.09	−0.50		0.10	0.58	
Personality traits																					
Achievement (mean)																					
Social open-mindedness (mean)	0.21	0.84		−0.25	−1.05		−0.03	−0.21		0.17	1.22		−0.43	−1.48		0.41	2.06	*	−0.12	−0.64	
Self-efficacy (mean)	−0.06	−0.26		0.22	1.00		−0.21	−1.59		−0.44	−3.36	***	0.07	0.24		−0.35	−1.83	†	−0.09	−0.47	
Interviewer rating																					
Willingness to respond (good)	−0.39	−1.59		−0.04	−0.15		−0.08	−0.42		−0.36	−2.12	*	−0.82	−2.42	**	−0.25	−0.86		−0.56	−2.18	*
Open-mindedness (good)	−0.36	−1.60		−0.42	−1.53		−0.02	−0.11		−0.04	−0.28		−0.42	−1.27		0.11	0.47		−0.29	−1.16	
Concentration (high)	−0.46	−2.07	*	0.55	2.32	**	−0.49	−3.62	***	−0.28	−2.03	*	−0.01	−0.02		0.03	0.15		−0.22	−0.96	
Language skills (good)	−0.45	−2.03	*	−0.30	−1.03		−0.19	−1.15		−0.22	−1.38		−0.87	−2.89	**	−0.27	−1.06		0.19	0.70	
Interaction effects																					
Child age/interviewer age	−0.01	−0.78		0.02	1.89	*	−0.01	−1.81	†	0.00	−0.06		0.01	0.83		−0.01	−0.91		−0.01	−1.01	
Child gender/interviewer gender	−0.81	−2.16	*	0.42	1.08		−0.21	−0.93		−0.03	−0.14		−0.57	−1.14		0.09	0.29		−0.08	−0.23	
Interviewer gender																					
Random effects																					
Intercept γ00	−2.93		***	−3.19		***	−1.49		***	−1.33		***	−3.41		***	−2.04		***	−2.30		***
Level 2 variance u0j	0.30			0.43			0.15			0.37			0.49			0.22			0.26		

Note: *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, †p ≤ 0.10

interviewer's age. The older the interviewers, the more item nonresponses occurred. Difficulties during the interview led to increased item nonresponse.

Furthermore, we found three significant effects on the respondent level. Similar to the interviewer's age, older children produce more item nonresponse.

Poor academic performance produced more item nonresponse regarding the question of achievement motivation. Last but not least, with decreasing concentration skills, item nonresponse increases.

In regard to the question about parents' interest in school, we found no significant effect at the interviewer level and four significant effects at the respondent level. Poor academic performance produced more item nonresponse regarding the question of parents' interest in school. The more self-efficacy children have, the more meaningful responses will be produced. With decreasing willingness to respond and decreasing concentration skills, item nonresponse increases.

Turning to our next model, analysing *family climate*, we found three significant effects on the interviewer level: Item nonresponse increases with increasing age of the interviewer. If the interview was conducted by the same interviewer, item nonresponse increases. The shorter the interview, the more item nonresponse will be produced. We identified three significant effects at the respondent level: Item nonresponse increases with decreasing age of the children. The smaller the willingness to respond and the poorer the language skills, the more item nonresponse will be produced.

At the interviewer level, item nonresponse to the question about the *behaviour in conflicts with friends* is explained by two variables: interventions of third parties and difficulties during the interview leading to more item nonresponse.

At the respondent level we found one significant variable: The greater the social and cognitive open-mindedness, the more item nonresponse will be produced.

For the question about *behaviour in hypothetical situations*, item nonresponse was only affected by one significant variable on the respondent level: The greater the willingness to respond, the more meaningful responses were produced.

5.   Conclusions and Discussion

The main aim of this article was to answer the question whether and if so, how child and interviewer characteristics and the interview setting affect item nonresponse in standardised, face-to-face interviews with children.

For this purpose, we used data from the Child Longitudinal Study conducted by the German Youth Institute, where children (ages 8-11) were interviewed using standardized interviews in three survey waves. To analyse item nonresponse, we selected questions that met two requirements: They had to cover substantial item nonresponse and should be compatible with theoretical guidelines. We computed multilevel logistic models with the software HLM 7.0 to better disentangle interviewer from respondent effects.

In Hypothesis 1 we tested whether item nonresponse in standardized, face-to-face interview with children would be reduced with increased cognitive functioning (measured by age and educational achievement). Our results support this hypothesis for the majority of the questions analysed. This is in line with other empirical evidence (Borgers et al. 1999; Borgers and Hox 2001; Vogl 2011).

In Hypothesis 2, we tested whether interviewer characteristics would influence the impact of item nonresponse in standardized, face-to-face interviews with children in different ways. We found that interviewers in child interviews are not neutral collectors of data, but detected no systematic pattern for item nonresponse due to interviewer characteristics. A closer look at the values of the interviewer variance turns out to be somewhat disillusioning: Not a single value showed significance. This means that in all models the between-interviewer variance is acceptably mild, so it could have been ignored and we could have used simpler, single-level statistical models (Hox 2010). However, for reasons of consistency we present hierarchical models. The nonsignificant variance could indicate that there might be other, more meaningful interviewer variables which have not been taken into account.

Concerning Hypothesis 3, we found two cross-level interactions between child and interviewer characteristics. Depending on the different types of questions, it seems that the effect of child age on item nonresponse was moderated by interviewer age in one question; the effect of child gender on item nonresponse was moderated by interviewer gender in another question. Given few interaction effects, there is only little support for this hypothesis.

Our fourth and last hypothesis tested whether third parties during the interview would influence the number of item nonresponses in standardised, face-to-face interviews with children. Contrary to our assumption that third parties would act as mediators and reduce item nonresponse, we found that the intervention of third parties increases item nonresponse in one of the questions. This is good news for surveys with children, because a third presence did not influence the question-answer process concerning item nonresponse for the majority of the questions analysed.

5.1. *What Do These Results Mean for Surveys With Children?*

Overall, the amount of item nonresponse in the child survey was considerably low.

The highest item-nonresponse rate was found for questions offering an explicit 'don't know' category, though not necessarily for sensitive questions. This might mean that children aged 8–11 by and large perform well in face-to-face surveys.

Respondents' characteristics that correlate with item nonresponse are age and education. This may be an indication that interviewer training should focus more on how to deal with young and less-educated children.

The interviewers' rating of the child's ability to manage the interview points to concentration skills as an important factor. Item nonresponse increases with decreasing concentration skills, independent of age and education.

To improve the children's concentration, the survey researcher could vary the structure of the questionnaire by using a range of different question forms.

In the future, third parties will continue to be present during interviews with children. But this is good news for data quality, as their presence does not influence item nonresponse.

5.2. *What Do These Results Mean for Survey Research on Item Nonresponse?*

In order to explain and predict item nonresponse, it is important to completely understand what happens during the question-answer process. Although a number of approaches exist (Krosnick, 1991; Tourangeau et al. 2000), we still lack a comprehensive theory explaining item nonresponse in surveys. Even Borgers and Hox (2001) conclude that they were not able to unequivocally confirm or reject Krosnick's satisficing theory. Furthermore, it is not

clear whether these approaches can be adapted to child surveys (first attempts by Vogl 2011). This also applies to the Beatty-Hermann (2002) model of the response process. We regard it as more of a heuristic than a verifiable theory. Therefore the present study did not aim to test the model in a strict empirical sense, but uses it as a helpful framework to rearrange our dependent variables.

The analysis of secondary data material has considerable disadvantages. First, there were only a limited number of interviewer characteristics available. Second, because of the small proportion of item nonresponse in general, we were unable to separate 'don't know' responses from refusals. Against the background of empirical evidence (see Section 2), we assume differences in the influence of interviewer and respondent characteristics on item nonresponse, broken down by these two categories.

More elegant ways to shed light on the question-and-answer process in standardised, face-to-face interviews with children would be experimental designs (first attempts by Fuchs 2008) or collecting reasons for item nonresponse and viable interviewer characteristics from the outset (De Leeuw et al. 2003).

Appendix

Table A1.   Child Longitudinal Study, dependent variables

Question topic	Percentage of item nonresponse per wave			Question wording	Response scale
Beatty-Hermann model of response process for item nonresponse					
Self-description available	(1) 7.0	(2) 4.5	(3) 2.4	. . .15 items with which one can describe oneself; for example: love to laugh. I'm sometimes sad. I like to scuffle. et cetera	Four-point scale without 'don't know' option: yes, rather yes, rather no, no
Leisure activities available	6.3	2.6	3.2	6 items about things one can do alone or with others, for example: playing game consoles, make music, go to the cinema et cetera	Dichotomous scale without 'don't know' option: yes, no
Parents' interest in school generatable	25.6	17.8	18.3	6 items, for example: Do your parents note school certificates and ratings? Are your parents satisfied with your academic performance in general? et cetera	Dichotomous scale offering 'don't know' option: yes, no, don't know
Achievement motivation accessible	17.9	20.3	20.7	6 items, for example: Do you often have problems getting along at school? Do you enjoy learning? et cetera	Dichotomous scale offering 'don't know' option: yes, no, don't know
Family climate accessible	4.2	2.8	3.1	5 items about how one feels about the family: I'm happy when my family is together. We have got many conflicts in our family. et cetera	Four-point scale without 'don't know' option: always, often, seldom, never

Table A1.   Continued

Question topic	Percentage of item nonresponse per wave			Question wording	Response scale
Beatty-Hermann model of response process for item nonresponse					
Behaviour: conflict with friends accessible	10.1	9.0	16.2	13 items on reaction, if child has conflict with friends: I roar with anger at him/her. I leave so as not to be annoyed anymore. I hustle, kick or beat him/her. et cetera	Four-point scale without 'don't know' option: always, often, seldom, never
Behaviour: hypothetical situations inestimable	–	–	9.7	9 items about evaluation of different hypothetical situations: How good or bad are you at telling a child that he/she has done something that has annoyed you? How good or bad are you at calling a new child to make an appointment with him/her? et cetera	Five-point scale offering 'don't know' option: very bad, rather bad, ok, rather good, very good, don't know

Note: Detailed questionnaires are available (in German only) on: www.dji.de/kinderpanel

6. References

Beatty, P. and D. Hermann. 2002. "To Answer or Not to Answer: Decision Processes Related to Survey Item Nonresponse." In *Survey Nonresponse*, edited by R. Groves, D. Dillman, R.J.A. Little, and T.L. Eltinge, 71–86. New York: Wiley.

Borgers, N., E. de Leeuw, and J. Hox. 1999. "Surveying Children: Cognitive Development and Response Quality in Questionnaire Research." In *Official Statistics in a Changing World*, edited by A. Christianson, 133–140. Stockholm: SCB.

Borgers, N., E. de Leeuw, and J. Hox. 2000. "Children as Respondents in Survey Research: Cognitive Development and Response Quality." *Bulletin de Méthodologie Sociologique* 66: 60–75.

Borgers, N. and J. Hox. 2001. "Item Nonresponse in Questionnaire Research with Children." *Journal of Official Statistics* 17: 321–335.

Borgers, N., J. Hox, and D. Sikkel. 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.

Borgers, N., D. Sikkel, and J. Hox. 2004. "Response Effects in Surveys on Children and Adolescents: The Number of Response Options, Negative Wording, and Neutral Mid-Point." *Quality & Quantity* 38: 17–33.

Bryk, A.S. and S.W. Raudenbush. 2004. *Hierarchical Linear Models for Social and Behavioural Research: Applications and Data Analysis Methods*. Newbury Park, CA: Sage Publications.

De Leeuw, E., N. Borgers, and A. Smits. 2004. "Pretesting Questionnaires for Children and Adolescents." In *Methods for Testing and Evaluating Survey Questionnaires*, edited by S. Presser, J.M. Rothgeb, M.P. Couper, J.T. Lessler, E. Martin, J. Martin, and E. Singer, 409–429. New York: Wiley.

De Leeuw, E., J. Hox, and M. Huisman. 2003. "Prevention and Treatment of Item Nonresponse." *Journal of Official Statistics* 19: 153–176.

De Leeuw, E. and M. Otter. 1995. "The Reliability of Children's Responses to Questionnaire Items: Question Effects in Children's Questionnaire Data." In *Advances in Family Research*, edited by J.P. Hox, B.F. van der Meulen, J.M. Kanssens, J.J. ter Laak, and L.W.C. Tavecchio, 251–257. Amsterdam: Thesis Publisher.

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.

Fuchs, M. 2008. "The Reliability of Children's Survey Responses: The Impact of Cognitive Functioning in Respondent Behavior." In *Proceedings of Statistics Canada Symposium*, Accessed March 15, 2013, http://www.statcan.gc.ca/pub/11-522-x/2008000/article/10961-eng.pdf.

Groves, R. 2004. *Survey Errors and Survey Costs*. Hoboken, NJ: John Wiley & Sons.

Groves, R.M., F.J. Fowler Jr,, M.P. Couper, J.M. Lepkowski, E. Singer, and R. Tourangeau. 2009. *Survey Methodology*. Hoboken, NJ: John Wiley & Sons.

Haunberger, S. 2005. "Interviewer und Befragte im Kinderpanel. Interviewdauer und Panelbereitschaft." In *Kinderleben – Aufwachsen Zwischen Institutionen, Familie und*

Freunden. Band 2, edited by C. Alt, 285–316. Wiesbaden: VS Verlag für Sozialwissenschaften.

Hox, J.J. 2010. *Multilevel Analysis. Techniques and Applications*. Second Edition. New York and Hove: Routledge.

Huddy, L., J. Billig, J. Bracciodieta, L. Hoeffler, P. Moynihan, and P. Pugliani. 1997. "The Effects of Interviewer Gender on the Survey Response." *Political Behavior* 19: 197–220. DOI: http://dx.doi.org/10.1023/A:1024882714254.

Juster, F.T. and J.P. Smith. 1997. "Improving the Quality of Economic Data: Lessons from the HRS and AHEAD." *Journal of the American Statistical Association* 92: 1268–1278.

Kail, R.V. 2011. *Children and Their Development*. Cambridge: Pearson Education.

Krosnick, J.A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5: 213–236. DOI: http://dx.doi.org/10.1002/acp. 2350050305.

Pickery, J. and G. Loosveldt. 1998. "The Impact of Respondent and Interviewer Characteristics on the Number of 'No Opinion' Answers." *Quality & Quantity* 32: 31–45. DOI: http://dx.doi.org/10.1023/A:1004268427793.

Pickery, J. and G. Loosveldt. 2001. "An Exploration of Question Characteristics that Mediate Interviewer Effects on Item Nonresponse." *Journal of Official Statistics* 17: 337–350.

Reuband, K.-H. 1984. "Dritte Personen beim Interview – Zuhörer. Adressaten oder Katalysatoren der Kommunikation?" In *Soziale Realität im Interview. Empirische Analysen und Methodische Probleme*, edited by H. Meulemann and K.-H. Reuband, 117–156. Campus: Frankfurt am Main.

Reuband, K.-H. 1987. "Unerwünschte Dritte beim Interview. Erscheinungsformen und Folgen." *Zeitschrift für Soziologie* 16: 303–308.

Schwarz, N. and S. Sudman. 1995. *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco: Jossey-Bass Publishers.

Scott, J. 1997. "Children as Respondents: Methods for Improving Data Quality." In *Survey Measurement and Process Quality*, edited by L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, and D. Trewin, 331–350. New York: Wiley.

Scott, J., M. Brynin, and R. Smith. 1995. "Interviewing Children in the British Household Panel Survey." In *Advances in Family Research*, edited by J. Hox, B.F. van der Meulen, M.A.M. Janssens, J.J.F. ter Laak, and L.W.C. Tavecchio, 259–266. Amsterdam: Thesis Publisher.

Shoemaker, P.J., M. Eichholz, and E.A. Skewes. 2002. "Item Nonresponse: Distinguishing Between Don't Know and Refuse." *International Journal of Public Opinion Research* 14: 193–201.

Tourangeau, R., L.J. Rips, and K. Rasinski. 2000. *The Psychology of Survey Response*. New York: Cambridge University Press.

Vogl, S. 2011. "Children Between the Age of 5 and 11: What 'Don't Know' Answers Tell Us." *Quality & Quantity* 46: 993–1011. DOI: http://dx.doi.org/10.1007/s11135-011-9438-9.

Zeger, S.L., K.-Y. Liang, and P.S. Albert. 1988. "Models for Longitudinal Data: A Generalized Estimating Equation Approach." *Biometrics* 44: 1049–1060. DOI: http://dx.doi.org/10.2307/2531734.