



Deutsches
Jugendinstitut

Allegra Lenz, Mirjam Weis, Anja Linberg, Sabine Düval, Thorsten Naab, Alexandra Langmeyer and Susanne Kuger

Web-Based Self-Reports of Young Children: A Pilot Study on Age Differences, Response Formats, Data Quality, and Survey Experience

2/2026

Research on children, young people and families at the interface of science, politics and professional practice

The German Youth Institute (DJI) is one of the largest social science research institutes in Europe. For over 60 years, it has been researching the living conditions of children, young people and families, advising the federal, state and local governments and providing important impulses for professional practice.

The DJI currently employs around 380 people at its two locations in Munich and Halle (Saale), including around 240 researchers.

The DJI is mainly funded by the Federal Ministry of Education, Family, Seniors, Women and Youth (BMBFSFJ) and the federal states. It receives additional funding for projects from the European Commission, foundations and other institutions that promote science, among others.

The authors are solely responsible for the contents which do not necessarily represent the opinion of the German Youth Institute (DJI).

Impressum

© 2026 Deutsches Jugendinstitut e.V.

Deutsches Jugendinstitut
Nockherstraße 2
81541 München

Datum der Veröffentlichung September 2026
DOI: 10.36189/DJI202619

Deutsches Jugendinstitut
Außenstelle Halle
Franckeplatz 1, Haus 12/13
06110 Halle

Ansprechpartner:
Allegra Lenz
Telefon +49 89 62306-278
E-Mail alenz@dji.de

Abstract

Children's self-reports are essential for capturing their individual perspectives but pose specific methodological challenges for survey research. Particularly among younger children, cognitive, linguistic, and motivational prerequisites may affect the feasibility of self-administered surveys. Against this background, the present pilot study examined the feasibility of an age-appropriate self-administered web survey (CAWI) among children aged 5–11 years, considering two key dimensions: data quality and survey experience. Data quality was assessed using completion time and item nonresponse, while survey experience was examined through child and parent ratings of understanding, perceived difficulty, and enjoyment. The study investigated differences related to children's cognitive abilities, approximated by age, and survey complexity, operationalized through four-point versus six-point response formats. Based on a sample of $N = 60$ parent–child dyads, findings indicated generally low item nonresponse and positive survey experiences. No statistically significant differences in data quality were found between age groups or response formats. Regarding survey experience, older children (8–11 years) showed higher levels of understanding and lower perceived difficulty, particularly from the parents' perspective. Contrary to expectations, parents rated the six-point response format as less difficult than the four-point format. Overall, the findings support the feasibility of age-appropriate CAWI for collecting self-reports from young children in terms of both data quality and survey experience, while highlighting the need for further validation in larger and more diverse samples.

1 Introduction

1.1 Children's self-reports and the need for age-appropriate survey methods

In psychology and related social and behavioral sciences, children are increasingly viewed not merely as objects of research, but as social actors with their own perspectives and experiences (Christensen & James, 2017; Facca et al., 2020; McCafferty & Mercado García, 2025). Accordingly, research on children's lives and well-being has emphasized the importance of asking children directly about their experiences in order to capture their perspectives more adequately (Ben-Arieh, 2005; Falkenberg et al., 2021). Although parent and other proxy reports provide valuable information, they cannot be assumed to fully represent children's own perceptions, evaluations, and everyday experiences (Erhart et al., 2009; Punch, 2002; Rees, 2017; Riley, 2004). Children's self-reports are therefore particularly important for psychological research on constructs concerning their subjective experiences, preferences, or evaluations (e.g., of well-being). This raises the question of whether and how self-administered web surveys can be implemented with young children while accounting for high data quality and positive survey experience.

From a methodological perspective, previous research suggests that children can provide self-reports in structured survey settings even within primary school age, provided that questionnaires are adapted to their cognitive and linguistic abilities (Borgers et al., 2000; de Leeuw, 2011; Scott, 1997). Children's ability to respond to standardized questions, however, is closely linked to developmental factors such as reading skills, language comprehension, and task comprehension (Borgers et al., 2000; de Leeuw, 2011; Niehaus et al., 2017). These factors may influence not only the quality of the resulting data, for example, in terms of item nonresponse (Haunberger, 2014), but may also be relevant for children's subjective experience of the survey task, including perceived difficulty and response burden (Borgers et al., 2000; Coombes et al., 2021; de Leeuw, 2011; Yan & Williams, 2022).

This is particularly relevant when questionnaires are used with younger children in the preschool and elementary school age. Completing a questionnaire involves several cognitive steps: Respondents need to understand the question, retrieve the relevant information from memory, integrate it into a summarized judgment, and translate the judgment to the response-scale format. Because children's cognitive and communicative skills are still developing, these response processes may be more demanding for younger than for older children (Borgers et al., 2004). Compared with older children, younger children may have more difficulty understanding abstract wording, complex sentence structures, and negations, as well as interpreting frequency categories and response scales with larger numbers of response options (Babarović et al., 2025; Borgers et al., 2000, 2003, 2004; Niehaus et al., 2017). According to satisficing theory, low motivation, high task difficulty, and limited cognitive resources can increase the likelihood that respondents provide satisfactory rather than optimal answers (Borgers et al., 2004; Krosnick, 1991). In child surveys,

such response strategies are relevant because they may affect response reliability and other data-quality indicators (Borgers et al., 2000, 2004).

A related concept in survey methodology is response burden, which refers to the subjective burden experienced by respondents when completing a survey. Response burden can vary across the course of a questionnaire and is shaped by characteristics of the instrument, such as length, task difficulty, and frequency of assessment, as well as by respondent characteristics, including ability and motivation (Yan & Williams, 2022). Empirically, response burden may become visible in participation and data-quality indicators such as breakoffs, missing data, measurement error, and response time (Yan & Williams, 2022). In research with children, these indicators are especially relevant because high task demands may affect both children's ability and motivation to complete a questionnaire and the quality of their responses (Borgers et al., 2000, 2004; de Leeuw, 2011).

Taken together, these considerations highlight the need for age-appropriate survey methods when collecting self-reports from children. Questionnaires for children should use simple, clear, and concrete wording, avoid complex sentence structures and negations, and limit the number of different scales within the same instrument (Babarović et al., 2025; Borgers et al., 2000, 2003, 2004; Niehaus et al., 2017). For younger children (i.e., in the preschool and elementary school age) in particular, response formats should be visually clear and limited in complexity. Additional design features, such as visual support, short and clearly structured questionnaire sections, practice items, and audio-based assistance, may help reduce comprehension demands and make the response task more concrete and manageable (Babarović et al., 2025; Borgers et al., 2000; Coombes et al., 2021; Naab & Abraham, 2024).

1.2 CAWI as flexible but challenging mode for younger children

Computer-assisted web interviewing (CAWI) offers several practical and methodological advantages for survey research, including flexible implementation, programmable presentation of questions and response options, automatic routing, integration of multimedia elements, and direct digital storage of responses (Bandilla, 2016; Callegaro et al., 2015). Because CAWI is self-administered and does not require interviewer-administered questionnaire completion, it may also reduce administrative demands as well as potential interviewer-related mode effects, particularly social-desirability bias (Bandilla, 2016; Liedl & Steiber, 2024; Zaal et al., 2026).

In research with children, these features may be practically useful, however, at the same time, these advantages depend on the respondents' abilities to understand, navigate, and complete the questionnaire with limited external support. This creates specific challenges when CAWI is used with younger children (i.e., in the preschool and elementary school age). In contrast to interviewer-administered surveys, questionnaire completion during self-administered web surveys requires respondents to complete the questionnaire without direct interviewer guidance

during the response (Callegaro et al., 2015; Daikeler et al., 2025). This is especially relevant in child surveys, because children’s understanding of survey tasks and questions is closely linked to their cognitive and linguistic development (de Leeuw, 2011). The absence of direct interviewer supervision in self-administered web surveys may reduce control over the response process and has been discussed as a potential risk factor for inattentive responding, misunderstandings, minimal cognitive effort, and mode-related response differences (Callegaro et al., 2015; Liedl & Steiber, 2024; Mühlböck et al., 2017). Consequently, in CAWI for children, clear instructions and age-appropriate instrument design are particularly important for supporting children’s comprehension and task completion (Borgers et al., 2000; de Leeuw, 2011).

These considerations suggest that CAWI for children cannot be treated merely as a direct transfer of adult-oriented web survey methods to a younger population. Instead, web-based questionnaires for children require careful age-appropriate design adaptations. These may include simple and concrete wording, short and clearly structured questionnaire sections, visual response formats, audio presentations of instructions and items, practice items, and clear instructions (Babarović et al., 2025; Borgers et al., 2000; Coombes et al., 2021; de Leeuw, 2011; Naab & Abraham, 2024; Niehaus et al., 2017). Such adaptations are intended to reduce reading and comprehension demands and to support children’s understanding of questions, response options, and the response task, particularly given differences in their cognitive and linguistic development (Babarović et al., 2025; Borgers et al., 2000; de Leeuw, 2011).

Taken together, these findings suggest that the feasibility of CAWI in child populations depends not only on whether children can generally provide self-reports, but also on whether the web-based instrument is adapted to their age. This is especially relevant for children at the lower end of the target age range (e.g., in the age of 5 years), whose reading and language skills, response-option understanding, and task comprehension are still developing (Borgers et al., 2000, 2003, 2004; de Leeuw, 2011). Evaluating the feasibility of child CAWI should therefore combine data-quality indicators such as item nonresponse, process-related indicators such as completion time, and child-reported indicators of subjective survey experience (Borgers et al., 2000; Coombes et al., 2021; Yan & Williams, 2022).

1.3 Research questions and hypotheses

Based on the literature reviewed above, further research is needed to assess the feasibility of CAWI among preschool- and elementary-school-aged children (i.e., generally children aged 5 to 11 years in Germany) and to identify appropriate approaches to its implementation. The present study addresses this issue through the following research questions and hypotheses: First, we aim to investigate the data quality of a CAWI survey with children aged 5 to 11 years (Research Question 1). We expect that data quality increases with more developed cognitive abilities, which generally improve with age (Borgers et al., 2000; de Leeuw, 2011; Haunberger, 2014; Scott, 1997) (Hypothesis 1). Moreover, previous research suggests that the

number of response options can affect response quality in surveys with children (Borgers et al., 2004). Accordingly, we anticipate that data quality decreases with a more complex survey design (i.e., six-point vs. four-point response format; Hypothesis 2). In addition, as a further aspect of feasibility, we examine how children subjectively experience the CAWI survey in terms of perceived difficulty, understanding, and enjoyment. (Research Question 2). Children's cognitive and linguistic development may facilitate their understanding of survey questions and reduce demands of the response process (Borgers et al., 2000; de Leeuw, 2011; Yan & Williams, 2022). Therefore, we expect that more developed cognitive abilities, which generally increase with age, are associated with lower perceived difficulty (Hypothesis 3) and higher perceived understanding (Hypothesis 4). We further hypothesize that a less complex survey design (i.e., four-point response scales) is associated with lower perceived difficulty (Hypothesis 5) and higher perceived understanding (Hypothesis 6), since more differentiated response formats may increase cognitive demands of the response task, particularly for younger children (Borgers et al., 2004; Coombes et al., 2021; Mellor & Moore, 2014). Given the limited state of research, we examine the associations of enjoyment with cognitive abilities and survey design complexity exploratorily, without formulating directional hypotheses.

2 Method

2.1 Design and procedure

To address the research questions and hypotheses, a pilot study was designed to investigate the feasibility of an age-appropriate CAWI questionnaire for children aged 5–11 years. Feasibility was examined using indicators of data quality (RQ1) and subjective survey experience (RQ2).

The study was conducted as a self-administered online survey using LimeSurvey. Parent-child dyads completed the survey in a home-based setting. Furthermore, the study included a between-subjects manipulation of response format. Parent-child dyads were randomly assigned within LimeSurvey to one of two questionnaire versions, using either a four-point or a six-point response scale format. In the four-point condition, all response options were verbally labeled. In the six-point condition, the endpoints were verbally labeled, whereas the intermediate response options were not verbally labeled. The response-format manipulation applied only to selected items. Some questionnaire sections were kept in a four-point format, whereas the remaining eligible items were presented either with four or six response options. The full survey included parent and child questionnaire sections and took approximately 30 minutes to complete in total. Completion on a computer, laptop, or tablet was recommended to ensure optimal display of the questionnaire content and visual response format.

2.2 Survey instrument and developmental adaptations

The child questionnaire contained 33 to 38 items, depending on whether children reported siblings and a best friend. The questionnaire was based on selected items from the AID:A 2023 survey, a nationwide panel study on the growing-up process of children, adolescents, and young adults in Germany conducted by the German Youth Institute (Kuger et al., 2024). The child questionnaire covered psychological well-being, self-esteem, school or kindergarten attitudes, selected family and peer relationships, and survey experience.

The questionnaire had previously been adapted for children aged 5–11 years for the AID:A 2023 survey, which was conducted as a computer-assisted personal interview (CAPI) (Linberg, Guglhör-Rudan et al., 2026; Linberg, Simm et al., 2026). This developmental adaptation included visual response formats and practice items to support children's understanding of the response scales and survey questions. For the present study, the adapted questionnaire was further developed for CAWI. The survey was designed to reduce cognitive, linguistic, and motivational demands on children, particularly for participants at the lower end of the target range. The design aimed to allow children to complete the questionnaire as independently as possible, without requiring reading or writing skills or parental assistance to provide answers. Additional adaptations for the web-based format included audio presentation of instructions and items, simple navigation, and an avatar-guided survey environment.

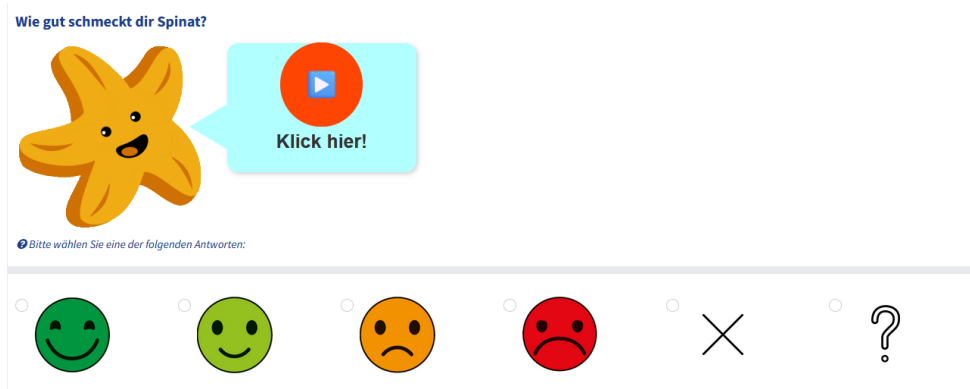
The avatar was designed as a child-friendly character that accompanied the participating children throughout the questionnaire and provided a consistent visual element across all survey pages. A speech bubble visually connected to the avatar contained a large play button, allowing children to listen to the corresponding instructions and questions. All audio recordings could be paused or repeated as needed.

Response options were presented graphically rather than numerically. In line with the AID:A survey, three types of visual response scales were used covering various dimensions: a smiley scale for preference ratings (“very good” to “very bad”), a circle scale for frequency ratings (“very often” to “never”) and a thumbs scale for agreement ratings (“strongly agree” to “strongly disagree”). For each item, children could also select a non-substantive response option if they did not want to answer or could not answer the question. As in the AID:A 2023 survey, at the beginning of the child questionnaire, an introductory text and three practice items familiarized children with the survey procedure and response format. Parent-child dyads were free to decide whether a parent was present while the child completed its part of the questionnaire.

Figure 1 provides an example of the survey interface, including the avatar, audio play button, and graphical response options.

Figure 1

Example of the child survey interface featuring the avatar, audio play button, and graphical response options.



Note. The screenshot shows a four-point smiley response scale with two additional non-substantive response options. The original survey was administered in German.

2.3 Participants and recruitment

Participants were recruited using a convenience sampling approach. Recruitment was initiated primarily through staff members of the German Youth Institute (Deutsches Jugendinstitut, DJI), who were invited to forward the study invitation to potentially eligible families within their personal and professional networks. In addition, selected preschools, after-school care facilities, and leisure or holiday programs across Germany were contacted and invited to disseminate the study invitation to eligible families. Participating parent-child dyads were eligible to enter a voucher lottery after completing the survey.

In total, 127 parent-child dyads accessed the survey. Cases were excluded if child assent was not provided or if the survey was terminated before the child assent item was reached ($n = 58$). Further cases were excluded if they met at least one exclusion criterion ($n = 9$). Exclusion criteria were fewer than five answered items in the child questionnaire, fewer than five of seven answered sociodemographic items in the parent questionnaire, or an inconsistency between parental age report and the child's reported institutional context, namely children older than six years reporting preschool attendance. The final analytic sample consisted of $N = 60$ parent-child dyads. Participant characteristics and group allocation are summarized in Table 1.

Children in the analytic sample were between 5 and 11 years old ($M = 7.78$ years; $SD = 2.03$ years). For age-group comparisons, children were grouped by age, distinguishing 5- to 7-year-olds ($n = 25$) from 8- to 11-year-olds ($n = 35$). According to parental reports, 51.67% of the children were boys and 48.33% girls; no child was reported as having a diverse gender identity. Most children were born in Germany (98.33%). Approximately 75% of the parents held a university degree; this above-average level of education within the sample is taken into account when interpreting the results. 95.00% of the parents stated that they communicated exclusively in German with their children in daily life.

Table 1

Participant characteristics and distribution across study conditions

Characteristic	Value		
A. Participant characteristics			
Child characteristics			
Age in years, <i>M</i> (<i>SD</i>)	7.78 (2.03)		
Boys, <i>n</i> (%)	31 (51.67)		
Girls, <i>n</i> (%)	29 (48.33)		
Diverse gender identity, <i>n</i> (%)	0 (0.00)		
Born in Germany, <i>n</i> (%)	59 (98.33)		
Family characteristics			
Household language: German, <i>n</i> (%)	57 (95.00)		
Parent with university degree, <i>n</i> (%)	45 (75.00)		
B. Distribution by age group and response-format condition (absolute frequencies)			
Age group	Four-point format	Six-point format	Total
5–7 years	11	14	25
8–11 years	17	18	35
Total	28	32	60

Note. In Panel A, percentages are based on the final analytic sample of $N = 60$ parent–child dyads. Panel B reports absolute frequencies by age group and response-format condition; children were randomly assigned to the response-format condition, whereas age groups were defined based on children's age.

2.4 Measurements

2.4.1 Feasibility

Feasibility was evaluated using indicators of data quality (RQ1) and subjective survey experience (RQ2).

Data quality indicators were total completion time for the child questionnaire and overall item nonresponse. Overall item nonresponse included technical missing data as well as non-substantive responses such as “I do not want to answer” and “I cannot answer”, because these responses could not be assigned to the substantive target scale. Items that were not displayed due to filter-based question routing were excluded from the calculation.

Subjective survey experience was assessed at the end of the child questionnaire using items on understanding, perceived difficulty, and enjoyment. These aimed to capture the children's subjective experience and the burden associated with key aspects of the survey situation. Parents who had been present during the child questionnaire completed corresponding items on their child's understanding, perceived difficulty, and enjoyment. To allow comparability across the four-point and the six-point response format, all subjective survey experience items were rescaled to range 0 to 1. The difficulty item was recoded so that higher values indicated greater perceived difficulty.

2.4.2 Cognitive Abilities and Survey Complexity

Cognitive abilities and survey complexity were considered as independent variables in the present study. Cognitive abilities were approximated using children's age as a proxy. As described above, children were divided into two age groups for the analyses, distinguishing 5- to 7-year-olds ($n = 25$) from 8- to 11-year-olds ($n = 35$). This distinction follows Vogl (2015), who identifies approximately eight years of age as an important developmental transition, characterized by advances in cognitive abilities, including multidimensional thinking, mental reversibility, more differentiated logical reasoning, and the ability to generalize.

Survey complexity was operationalized through the complexity of the response format, comparing four-point and six-point response scales (see Section 2.1). Children were randomly assigned to one of the two response formats, resulting in approximately equal group sizes (four-point: $n = 28$; six-point: $n = 32$).

2.5 Analytic strategies

Analyses were conducted using IBM SPSS Statistics 31.0 (IBM Corp., 2025) and StataNow 19.5 (StataCorp., 2025). Descriptive statistics included frequencies, percentages, means, standard deviations, medians, and ranges. Differences between age group and response format conditions were examined using Mann-Whitney U

tests. For indicators with predefined directional expectations, namely completion time, overall item nonresponse, understanding, and perceived difficulty, one-tailed tests were used. Perceived survey enjoyment was analyzed using two-tailed tests, as no directional hypothesis was formulated for this indicator. Although the central comparisons were guided by directional hypotheses, the pilot character of the study and the limited sample size call for interpreting inferential results as preliminary evidence rather than confirmatory hypothesis tests. To account for the small sample size in the nonparametric group comparisons, exact rather than asymptotic p-values were used. Results with $p < .05$ were interpreted as statistically significant, while results with $p < .10$ were treated as exploratory trend-level evidence.

3 Results

3.1 Overall feasibility indicators

Table 2 summarizes the overall feasibility indicators regarding the child questionnaire. The child questionnaire took about 13 minutes to complete on average; overall item nonresponse was low with only a mean on 1.58 (SD: 3.97). In total, 37 of 60 children answered every item substantively. The mean item nonresponse was influenced by three outlying cases with 15 or 17 missing responses. Excluding these three outlying cases from the main analysis does not substantially change the results.

Children's subjective survey-experience ratings indicated high understanding and enjoyment, while perceived difficulty was low. Overall, parent ratings showed a similar pattern.

In sum, the descriptive findings indicate that the adapted CAWI instrument yielded indicators consistent with manageable participation and low levels of item nonresponse among children in the analytic sample.

Table 2

Overall feasibility indicators (data quality and survey-experience ratings)

Indicators	<i>M</i>	<i>SD</i>	Additional information
Total completion time in minutes	13.12	7.22	Median = 12.44; Range: 2.98 – 30.66
Overall item nonresponse	1.58	3.97	Median = 0; 37 out of 60 children answered every item substantively
Child-reported understanding	0.90	0.18	
Child-reported difficulty	0.15	0.19	
Child-reported enjoyment	0.85	0.22	
Parent-rated child understanding	0.84	0.20	
Parent-rated child difficulty	0.21	0.20	
Parent-rated child enjoyment	0.74	0.23	

Note. Values are means and standard deviations. Survey-experience indicators were rescaled to range from 0 to 1. Higher values indicate higher understanding, higher difficulty and higher enjoyment. Overall item nonresponse

includes technical missing data and non-substantive responses; items not displayed due to filter routing were excluded.

3.2 Results for data quality

To address Research Question 1, data quality was assessed using completion time and item nonresponse, with differences examined across age groups (H1) and response formats (H2).

Results for data quality regarding age-group differences are shown in Table 3. For total completion time, no statistically significant age-group difference was found. However, Levene's test indicated a significant difference in variance between age group, with younger children showing greater variability in completion time than older children, $F(1, 37) = 5.60, p = .023$. Under the exploratory $p < .10$ threshold, younger children showed higher item nonresponse than older children. Thus, although the findings partly aligned with the expected age-related pattern, Hypothesis 1 was not supported at the conventional significance level of $\alpha = .05$.

Table 3

Data quality by age group

Indicators	5–7 years		8–11 years		z	p(U)
	M	SD	M	SD		
Total completion time in minutes	15.23	9.07	11.81	5.58	1.10	.141
Overall item nonresponse	2.44	4.88	0.97	3.09	1.34	.093

Note. Values are means and standard deviations. z and p refer to Mann-Whitney U-tests. For directional hypotheses, one-tailed p-values are reported. Results with $p < .05$ were interpreted as statistically significant, and results with $p < .10$ as exploratory trend-level evidence.

Table 4 presents the results for data quality regarding response-format differences. No statistically significant response-format differences were found for total completion time or overall item nonresponse. Thus, the expected disadvantage of the six-point response format was not observed for these indicators, and Hypothesis 2 was not supported.

Table 4

Data quality by response-format condition

Indicators	Four-point		Six-point		z	p(U)
	M	SD	M	SD		
Total completion for varied-format items in minutes	8.79	4.75	9.53	5.09	-0.25	.407
Overall item nonresponse for varied-format items	1.71	4.64	0.84	2.11	-0.33	.386

Note. Values are means and standard deviations. z and p refer to Mann-Whitney U-tests. For directional hypotheses, one-tailed p -values are reported. Results with $p < .05$ were interpreted as statistically significant, and results with $p < .10$ as exploratory trend-level evidence.

3.3 Results for survey experience

To address Research Question 2, survey experience was examined across age groups and response formats.

Regarding age-related differences, younger children reported lower understanding than older children (H4), with exploratory trend-level evidence for this difference ($p = .052$). Parent ratings showed the same pattern of rating the survey as less understandable for younger children ($p < .05$). Additionally, parents assessed the survey as more difficult for younger children ($p < .01$) (H3). No statistically significant age-group differences were found for children's perceived difficulty (H3) or enjoyment, or for parent-rated enjoyment. Thus, Hypothesis 3, predicting lower perceived difficulty with higher age, was supported by parent ratings but not children's own ratings. Hypothesis 4, predicting higher perceived understanding with higher age, was supported by parent ratings, while children's ratings showed a trend in the expected direction. Results for survey experience regarding age-group differences are shown in Table 5.

Table 5

Survey experience by age group

Indicators	5–7 years		8–11 years		z	p(U)
	M	SD	M	SD		
Child-reported understanding	0.85	0.21	0.93	0.15	-1.67	.052
Child-reported difficulty	0.14	0.18	0.15	0.21	0.06	.481
Child-reported enjoyment	0.83	0.26	0.86	0.19	0.14	.889
Parent-rated child understanding	0.79	0.16	0.87	0.22	-2.19	.015
Parent-rated child difficulty	0.29	0.20	0.15	0.18	2.54	.006
Parent-rated child enjoyment	0.74	0.23	0.74	0.23	-0.17	.872

Note. Values are means and standard deviations. z and p refer to Mann-Whitney U-tests. For directional hypotheses, one-tailed p -values are reported; for enjoyment, two-tailed p -values are reported. Results with $p < .05$ were interpreted as statistically significant, and results with $p < .10$ as exploratory trend-level evidence.

No statistically significant response format differences were found for child-reported survey experience, as shown in Table 6. In parent ratings, perceived difficulty was lower in the six-point condition than in the four-point condition, contrary to the expected direction (H5). Because the observed difference was opposite to the directional hypothesis, the hypothesis-directed one-sided test was not significant. However, two-sided testing indicated a statistically significant difference ($p < .001$). No significant response-format differences were found for parent-rated understanding (H6) or enjoyment. Thus, Hypothesis 5, predicting lower perceived difficulty with the four-point response format, was not supported, with parent ratings indicating a difference in the opposite direction. Hypothesis 6, predicting higher perceived understanding with the four-point format, was also not supported.

Table 6

Survey experience by response-format condition

Indicators	Four-point		Six-point		z	p(U)
	M	SD	M	SD		
Child-reported understanding	0.91	0.18	0.90	0.19	0.48	.330
Child-reported difficulty	0.17	0.24	0.12	0.15	0.64	.735
Child-reported enjoyment	0.88	0.20	0.83	0.24	0.96	.343
Parent-rated child understanding	0.79	0.25	0.87	0.14	-0.98	.839
Parent-rated child difficulty	0.32	0.20	0.12	0.16	3.48	< .001
Parent-rated child enjoyment	0.71	0.19	0.76	0.25	-0.84	.407

Note. Values are means and standard deviations. z and p refer to Mann-Whitney U-tests. For directional hypotheses, one-tailed p -values are reported; for enjoyment, two-tailed p -values are reported. Results with $p < .05$ were interpreted as statistically significant, and results with $p < .10$ as exploratory trend-level evidence.

4 Discussion

4.1 Feasibility of CAWI for children aged 5–11

The descriptive results provide initial evidence that the age-appropriate CAWI questionnaire was feasible among children in the present sample: the CAWI was generally perceived positively by children and parents, while the perceived level of difficulty was low and overall item nonresponse was also low. This pattern aligns with findings suggesting that computer-assisted or online-based methods can be successfully implemented with children, although their suitability depends on careful adaptation to the target age group and the specific assessment context (Borgers et al., 2000; Nelson et al., 2021; van Hattum & de Leeuw, 1999). The low rate of item nonresponse is consistent with earlier findings on computerized child and youth self-report measures, which suggest that computerized administration may reduce missing data by limiting the skipping of individual questions (Coombes et al., 2021; van Hattum & de Leeuw, 1999). However, the findings concerning survey experience should not be interpreted as indicating an objectively high or low level of burden, given the absence of normative benchmarks or established cut-off values. Nevertheless, they do suggest that the survey's technical and design implementation was generally manageable for participating children.

In the present study, features such as visual response formats, audio presentation, practice items, and simple navigation may have contributed to the generally positive pattern of feasibility indicators. It should be noted that feasibility should be evaluated through multiple indicators rather than completion time, missingness or subjective survey experience alone. Although the findings do not provide a basis for defining a clear lower age threshold within the examined age range, they suggest that younger children (e.g. in the preschool or elementary school age) can participate in age-appropriate web surveys with audiovisual support. Nevertheless, particular attention to the developmental needs of younger children remains important.

4.2 Data quality in child CAWI (RQ1)

Data quality was examined in terms of completion time and item nonresponse, with age (H1) and response format (H2) considered as potential methodological factors. The findings suggest that age may be relevant for selected indicators of data quality, although the expected age-related pattern was not consistently supported. For completion time, no statistically significant age-group difference was found. For overall item nonresponse, however, the difference pointed in the expected direction and reached the exploratory trend-level threshold, with younger children showing more missing or non-substantive responses than older children. Descriptively, younger children also showed longer average completion times. This pattern is consistent with the assumption that standardized questionnaires may place greater demands on younger children because language comprehension, verbal memory, and attention-related capacities are still developing during this age range (Borgers et al., 2004; Niehaus et al., 2017; Siegler et al., 2021). It is also notable that the variance

in completion times was statistically significant higher in the younger age group. This could indicate greater heterogeneity among younger children regarding their ability to complete the survey within a consistent timeframe.

The response-format experiment did not show the expected disadvantage of the six-point scale. For those items whose response format varied randomly between conditions, no statistically significant group differences emerged concerning completion time or item nonresponse. In terms of overall item nonresponse, children in the six-point condition showed descriptively slightly more favorable values, with an average of about half as many missing responses compared to children in the four-point group.

The absence of the expected disadvantage of the six-point format stands in partial contrast to previous research suggesting that the number and format of response options can affect children's ability to use response scales appropriate, with more differentiated response formats potentially posing particular challenges for younger children (Borgers et al., 2004; Coombes et al., 2021; Mellor & Moore, 2014). In the present study, the six-point format was embedded in an adapted survey environment that included age-appropriate graphic and auditory design components. These features may have supported children's understanding of the response task, as visual response formats and supporting auditory elements can facilitate understanding and reduce reading demands (Borgers et al., 2000; Coombes et al., 2021; Naab & Abraham, 2024), potentially mitigating some of the demands associated with a more differentiated response format. This interpretation remains tentative, because the study did not isolate the effects of individual design features.

Overall, the findings provide limited support for the expected association between higher cognitive abilities and better data quality (Hypothesis 1), while the hypothesized disadvantage of the six-point response format in data quality (Hypothesis 2) was not observed. However, the absence of statistically significant differences does not establish equivalence between the two response formats.

4.3 Survey experience in child CAWI (RQ2)

Survey experience was examined in terms of perceived understanding, difficulty, and enjoyment, considering both children's and parents' perspectives. Age-related differences were observed in selected indicators of subjective survey experience. Younger children reported lower understanding than older children, with exploratory trend-level evidence for this difference, while no statistically significant age-group differences were found for child-reported difficulty or enjoyment. Parent ratings showed a similar age-related pattern for understanding and indicated higher perceived difficulty among younger children. Therefore, the expected association between higher cognitive abilities and lower perceived difficulty (Hypothesis 3) was supported by parent ratings but not by children's own ratings. The expected association between higher cognitive abilities and greater understanding (Hypothesis 4) was supported by parent ratings, while children's own ratings showed the same pattern at trend level. These findings suggest that younger children may experience or display difficulties with selected aspects of the survey and

underline that child and parent perspectives capture related but not identical aspects of the survey situation. This aligns with the theory-based assumption that information provided by parents cannot simply be equated with children's own perceptions and evaluations (Erhart et al., 2009; Punch, 2002; Rees, 2017).

Regarding response-format differences, children in both conditions reported on average high levels of understanding, low perceived difficulty, and a high degree of enjoyment while answering the questions; no statistically significant response-format differences were found for these indicators. Thus, children's own ratings did not support the expected advantages of the four-point response format regarding perceived difficulty (Hypothesis 5) or understanding (Hypothesis 6).

The only clear response-format difference emerged from the parental assessment of perceived difficulty. Contrary to the hypothesized direction, parents rated the questionnaire as less difficult for children in the six-point condition than in the four-point condition; this difference was statistically significant in a two-tailed test. Therefore, parent ratings did not support Hypothesis 5 and instead indicated a difference in the opposite direction. However, this finding should be interpreted cautiously, as the response format used to assess survey experience varied between conditions and may have influenced parent's ratings. Moreover, this pattern was not observed in children's own difficulty ratings.

Overall, the present findings suggest that younger children may experience greater difficulties with selected aspects of web-based surveys, particularly regarding understanding. Regarding response formats, neither the four- nor the six-point format showed a consistent advantage in children's subjective survey experience. The exploratory findings for enjoyment likewise revealed no statistically significant differences between age groups or response formats.

4.4 Limitations and future research

As is typical in empirical research, this study is not without its limitations. Therefore, some limitations should be considered when interpreting the findings. Firstly, the study was based on a relatively small pilot sample of $N = 60$ parent-child dyads. Therefore, one might argue that statistical power was limited, and subgroup analyses were constrained.

Secondly, the study used a convenience sampling approach, and the analytic sample was selective with a strong overrepresentation of participating parents reporting a university degree (75%). Although this is not a significant factor in terms of the experimental design (Meltzer, Naab & Daschmann, 2012), future studies should focus more closely than this study on environmental factors as potential mediators or moderators of questionnaire comprehension, such as parents' educational background or parental support during survey participation. Although the child questionnaire was designed to be completed independently by the participating children, parental presence was not prohibited. Therefore, it cannot be ruled out that parents provided explicit or implicit support. Particularly in the case of online data collection involving younger children, assistance from adults, such as parents

or other caregivers, is a methodologically relevant factor that should be taken into account when interpreting and planning such studies (Kooijmans et al., 2022; Nelson et al., 2021). Future studies should include additional information on the completion setting, such as interruptions, environmental distractions and parental assistance.

Thirdly, another limitation arises from the necessarily limited operationalization of feasibility and data quality. Completion time and item nonresponse are useful indicators, but they do not provide a comprehensive assessment of response quality, as other aspects, such as careless responding, may also affect the quality of online survey data (Brühlmann et al., 2020). Longer completion times may reflect increased difficulties or uncertainties, but they do not necessarily indicate poorer response quality, as they may also indicate more extensive processing during the response process (Bassili & Scott, 1996; Yan & Tourangeau, 2008). Similarly, item nonresponse captures only one aspect of potential data-quality problems. Future work should therefore include additional indicators, such as response consistency, response patterns and reliability.

Furthermore, the two response-format conditions differed not only in the number of response options but also in their verbal labeling (fully labeled response options vs. end-point labeling). Consequently, potential effects of the number of response options cannot be disentangled from those of verbal labeling.

Finally, the age grouping carried out in the present pilot study provided only a broad distinction between younger and older children. Although this comparison was useful for initial evaluation, it doesn't fully capture continuous developmental differences in cognitive, linguistic, and reading-related abilities. Future projects with larger and more diverse samples should examine age more fine-grainedly and consider functional criteria such as reading ability or language comprehension. Cognitive interviews could provide additional insight into how children of different ages understand the posed questions and response formats, and they may be particularly useful for identifying age-specific comprehension difficulties. Recent research supports the relevance and feasibility of this method, even with young children (Babarović et al., 2025; Gale et al., 2025). Further psychometric validation of the adapted scales is also needed before broader use of large-scale child CAWI studies.

Taken together, these limitations indicate that the present findings should be interpreted as initial pilot evidence and require replication in larger and more diverse samples.

4.5 Concluding remarks

Overall, this pilot study highlights the methodological potential of age-appropriate CAWI for collecting self-reports from children aged 5–11 years. Children generally completed the survey with low item nonresponse, high understanding and enjoyment, and low perceived difficulty. The findings suggest that children's cognitive abilities, approximated by age, may play a role in data quality and survey

experience, whereas increasing survey complexity, operationalized through the number of response options, did not consistently compromise either dimension. However, neither response format demonstrated a clear advantage across the indicators examined. Given the exploratory nature of the study, these findings do not establish equivalence between response formats or a definitive lower age threshold for child CAWI. Rather, they underscore the importance of aligning survey design with children's developmental capabilities and systematically evaluating both data quality and survey experience before broader implementation.

Literature

- Babaročić, T., Krpanec, E., Blažev, M., Dević, I., Downey, S., Huttunen, I., Panico, L., Perron, Z., Santos, A., Taylor, L. K., Upadaya, K., Symonds, J., & Pollock, G. (2025). Children's understanding of well-being related questions: Results of cognitive interviews in four European countries. *International Journal of Social Research Methodology*, 28(2), 163–178. <https://doi.org/10.1080/13645579.2024.2312621>
- Bandilla, W. (2016). *GESIS survey guidelines: Web surveys* (Version 2.0). GESIS – Leibniz Institute for the Social Sciences. https://doi.org/10.15465/gesis-sg_003
- Bassili, J. N., & Scott, B. S. (1996). Response latency as a signal to question problems in survey research. *Public Opinion Quarterly*, 60(3), 390–399. <https://doi.org/10.1086/297760>
- Ben-Arieh, A. (2005). Where are the children? Children's role in measuring and monitoring their well-being. *Social Indicators Research*, 74(3), 573–596. <https://doi.org/10.1007/s11205-004-4645-6>
- Borgers, N., de Leeuw, E., & Hox, J. (2000). Children as respondents in survey research: Cognitive development and response quality. *Bulletin de Méthodologie Sociologique*, 66(1), 60–75. <https://doi.org/10.1177/075910630006600106>
- Borgers, N., Hox, J. J., & Sikkels, 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. <https://doi.org/10.1093/ijpor/15.1.83>
- Borgers, N., Sikkels, D., & Hox, J. (2004). Response effects in surveys on children and adolescents: The effect of number of response options, negative wording, and neutral mid-point. *Quality & Quantity: International Journal of Methodology*, 38(1), 17–33. <https://doi.org/10.1023/B:QUQU.0000013236.29205.a6>
- Brühlmann, F., Petralito, S., Aeschbach, L. F., & Opwis, K. (2020). The quality of data collected online: An investigation of careless responding in a crowdsourced sample. *Methods in Psychology*, 2, Article 100022. <https://doi.org/10.1016/j.metip.2020.100022>
- Callegaro, M., Manfreda, K. L., & Vehovar, V. (2015). *Web survey methodology*. SAGE.
- Christensen, P., & James, A. (Eds.). (2017). *Research with children. Perspectives and practices* (3rd ed.). Routledge.
- Coombes, L., Bristowe, K., Ellis-Smith, C., Aworinde, J., Fraser, L. K., Downing, J., Bluebond-Langner, M., Chambers, L., Murtagh, F. E. M., & Harding, R. (2021). Enhancing validity, reliability and participation in self-reported health outcome measurement for children and young people: A systematic review of recall period, response scale format, and administration modality. *Quality of Life Research*, 30, 1803–1832. <https://doi.org/10.1007/s11136-021-02814-4>
- Daikeler, J., Roßmann, J., Bretsch, D., Gummer, T., & Silber, H. (2025). Instructed response item attention checks in mixed-mode questionnaires: Evidence from a probability-based mail and online panel. *Field Methods*, 37(3), 228–243. <https://doi.org/10.1177/1525822X241274103>
- de Leeuw, E. D. (2011). *Improving data quality when surveying children and adolescents: Cognitive and social development and its role in questionnaire construction and pretesting*. Report for the European Social Survey.
- Erhart, M., Ellert, U., Kurth, B. M., & Ravens-Sieberer, U. (2009). Measuring adolescents' HRQoL via self reports and parent proxy reports: An evaluation of the psychometric properties of both versions of the KINDL-R instrument. *Health and Quality of Life Outcomes*, 7, Article 77. <https://doi.org/10.1186/1477-7525-7-77>
- Facca, D., Gladstone, B., & Teachman, G. (2020). Working the limits of "giving voice" to children: A critical conceptual review. *International Journal of Qualitative Methods*, 19. <https://doi.org/10.1177/1609406920933391>
- Falkenberg, T., Ukasoanya, G., & Krepski, H. (2021). Students' understanding of student well-being: A case study. *McGill Journal of Education*, 56(2–3), 201–224. <https://doi.org/10.7202/1096452ar>
- Gale, V., Powell, P. A., & Carlton, J. (2025). Young children (6–7 years) can meaningfully participate in cognitive interviews assessing comprehensibility in health-related quality of life domains: A qualitative study. *Quality of Life Research*, 34, 1633–1646. <https://doi.org/10.1007/s11136-025-03940-z>
- Haunberger, S. (2014). Item nonresponse in face-to-face-interviews with children. *Journal of Official Statistics*, 30(3), 459–477. <https://doi.org/10.2478/jos-2014-0029>
- IBM Corp. (2025). *IBM SPSS Statistics for Windows, Version 31.0* [Software]. <https://www.ibm.com/de-de/products/spss-statistics>
- Kooijmans, R., Langdon, P. E., & Moonen, X. (2022). Assisting children and youth with completing self-report instruments introduces bias: A mixed-method study that includes children and young people's views. *Methods in Psychology*, 7, Article 100102. <https://doi.org/10.1016/j.metip.2022.100102>
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5(3), 213–236. <https://doi.org/10.1002/acp.2350050305>
- Kuger, S., Bernhardt, J., Herz, A., Linberg, A., Recksiedler, C., Simm, I., Ulrich, S., Braun, D., Fischer, A., Schacht, D., & Quellenberg, H. (2024). AID:A 2023 – Design des Surveys. In *Aufwachsen in Deutschland: Alltagswelten 2023 (AID:A 2023). Haupterhebung 2023* (1. Aufl., S. 7–15). Deutsches Jugendinstitut. <https://doi.org/10.3278/9783763978007>
- Liedl, B., & Steiber, N. (2024). Führen Online-Befragungen zu anderen Ergebnissen als persönliche Interviews? Eine Schätzung von Moduseffekten am Beispiel eines Mixed-Mode Surveys. *Österreichische Zeitschrift für Soziologie*, 49, 1–22. <https://doi.org/10.1007/s11614-023-00532-4>

- Linberg, A., Guglhör-Rudan, A., Chabursky, S., Langmeyer, A., Maly-Motta, H., Naab, T., Simm, I., & Kuger, S. (2026). *The AID:A 2023 child interview: Expanding direct self-reports to the 5–8 age cohort (AID:A Survey Report 4)*. Deutsches Jugendinstitut. <https://doi.org/10.36189/DJI202620>
- Linberg, A., Simm, I., Guglhör-Rudan, A., Naab, T., Langmeyer, A., & Kuger, S. (2026). *Durch die Augen von Kindern: Repräsentative Kinderbefragungen im Vor- und Grundschulalter am Beispiel von AID:A* [Manuscript submitted for publication]. Deutsches Jugendinstitut.
- McCafferty, P., & Mercado García, E. (2025). Children's participation in social work research. A secondary data analysis of an umbrella review of children's participation in child welfare. *The British Journal of Social Work*, 55(4), 1574–1597. <https://doi.org/10.1093/bjsw/bcaf017>
- Mellor, D., & Moore, K. A. (2014). The use of Likert scales with children. *Journal of Pediatric Psychology*, 39(3), 369–379. <https://doi.org/10.1093/jpepsy/jst079>
- Meltzer, C. E., Naab, T., & Daschmann, G. (2012). All Student Samples Differ: On Participant Selection in Communication Science. *Communication Methods and Measures*, 6(4), 251–262. <https://doi.org/10.1080/19312458.2012.732625>
- Mühlböck, M., Steiber, N., & Kittel, B. (2017). Less supervision, more satisficing? Comparing completely self-administered web-surveys and interviews under controlled conditions. *Statistics, Politics and Policy*, 8(1), 13–28. <https://doi.org/10.1515/spp-2017-0005>
- Naab, T., & Abraham, M. (2024). Wie ein Spielzeug zum Helfer in Kinderbefragungen werden kann: Zum Einsatz von Audio-Stiften zur Unterstützung von Befragungen mit Kindern im Vor- und Grundschulalter. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 60, 73–91. <https://doi.org/10.21240/mpaed/60/2024.04.24.X>
- Nelson, P. M., Scheiber, F., Laughlin, H. M., & Demir-Lira, Ö. E. (2021). Comparing face-to-face and online data collection methods in preterm and full-term children: An exploratory study. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.733192>
- Niehaus, S., Volbert, R., & Fegert, J. M. (2017). *Entwicklungsgerechte Befragung von Kindern im Strafverfahren*. Springer. <https://doi.org/10.1007/978-3-662-53863-0>
- Punch, S. (2002). Research with children. The same or different from research with adults? *Childhood*, 9(3), 321–341. <https://doi.org/10.1177/0907568202009003005>
- Rees, G. (2017). *Children's views on their lives and well-being. Findings from the children's worlds project*. Springer.
- Riley, A. W. (2004). Evidence that school-age children can self-report on their health. *Ambulatory Pediatrics*, 4(4), 371–376. <https://doi.org/10.1367/A03-178R.1>
- Scott, J. (1997). Children as respondents: Methods for improving data quality. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 331–350). Wiley.
- Siegler, R., Saffran, J. R., Gershoff, E. T., & Eisenberg, N. (2021). *Entwicklungspsychologie im Kindes- und Jugendalter* (5. Aufl.). Springer.
- StataCorp. (2025). *Stata statistical software: Release 19* [Software]. StataCorp LLC.
- van Hattum, M. J. C., & de Leeuw, E. D. (1999). A disk by mail survey of pupils in primary schools: Data quality and logistics. *Journal of Official Statistics*, 15(3), 413–429.
- Vogl, S. (2015). *Interviews mit Kindern führen: Eine praxisorientierte Einführung*. Beltz Juventa.
- Yan, T., & Tourangeau, R. (2008). Fast times and easy questions: The effects of age, experience and question complexity on web survey response times. *Applied Cognitive Psychology*, 22(1), 51–68. <https://doi.org/10.1002/acp.1331>
- Yan, T., & Williams, D. (2022). Response burden: Review and conceptual framework. *Journal of Official Statistics*, 38(4), 939–961. <https://doi.org/10.2478/jos-2022-0041>
- Zaal, E., Ongena, Y., van der Velden, N., Loughnan, D., & Hoeks, J. (2026). Unraveling honest responding: A systematic review on the effectiveness of social desirability bias reduction methods in survey research. *Quality & Quantity*, 60, 10359–10391. <https://doi.org/10.1007/s11135-026-02664-7>

Appendix

Ethics statement

The study received a positive ethics vote from the Ethics Committee of the German Youth Institute (DJI) in June 2025 (application number 2023/001). Parents provided informed consent, and children provided assent before participation. Participation was voluntary, and data was collected anonymously.

CRedit Author Contributions

Allegra Lenz: Conceptualization, Methodology, Software, Investigation, Data curation, Formal analysis, Project administration, Writing – original draft, Writing – review & editing. **Mirjam Weis:** Conceptualization, Methodology, Supervision, Project administration, Writing – review & editing. **Anja Linberg:** Conceptualization, Methodology, Supervision, Project administration, Writing – review & editing. **Sabine Düval:** Data curation, Formal analysis. **Thorsten Naab:** Methodology, Writing – review & editing. **Alexandra Langmeyer:** Methodology, Writing – review & editing. **Susanne Kuger:** Supervision.

Preprints published so far

1/2025

BlindtextBlindtext Blindtext

1/2025

BlindtextBlindtext Blindtext

1/2025

BlindtextBlindtext Blindtext

1/2025

BlindtextBlindtext Blindtext

1/2025

BlindtextBlindtext Blindtext

Deutsches Jugendinstitut e. V.

Nockherstraße 2

D-81541 München

Postfach 90 03 52

D-81503 München

Telefon +49 89 62306-0

Fax +49 89 62306-162