



Enhancing cognitive accessibility in assessments for children with neurodisability: development and implementation of an adaptation tracking questionnaire

Magnus Ivarsson, Henrik Danielsson, Lena Almqvist & Christine Imms

To cite this article: Magnus Ivarsson, Henrik Danielsson, Lena Almqvist & Christine Imms (28 Jan 2025): Enhancing cognitive accessibility in assessments for children with neurodisability: development and implementation of an adaptation tracking questionnaire, *Disability and Rehabilitation*, DOI: [10.1080/09638288.2025.2455532](https://doi.org/10.1080/09638288.2025.2455532)

To link to this article: <https://doi.org/10.1080/09638288.2025.2455532>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 28 Jan 2025.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Enhancing cognitive accessibility in assessments for children with neurodisability: development and implementation of an adaptation tracking questionnaire

Magnus Ivarsson^a , Henrik Danielsson^a , Lena Almqvist^b and Christine Imms^c 

^aDepartment of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden; ^bDepartment of Psychology, School of Health, Care, and Social Welfare, Mälardalen University, Västerås, Sweden; ^cMurdoch Children's Research Institute, The University of Melbourne, Melbourne, Australia

ABSTRACT

Purpose: The range of impairments in children with neurodisability (ND) complicates data collection, yet individualising materials and procedures could enable more children to self-report. This study introduces the Cognitive Accessibility Tracking Questionnaire (CATQ), designed to monitor changes enhancing accessibility ("adaptations") in interview-administered patient-reported outcome measures (PROMs). The CATQ is used in a longitudinal study of mental health and participation in children with ND investigating adaptation use and its utility in assessing the risk of bias introduced by these adaptations.

Materials and methods: The 13-item CATQ was developed with experts in ND and augmentative and alternative communication. Predictors of PROM adaptations were analysed using linear regression; the overall change was tested with a t-test and item-specific agreement with Cohen's weighted kappa and proportion of agreement.

Results: Six interviewers conducted 69 interviews, interviewing 43 children once or twice. Common adaptations included explaining/replacing concepts (56.5% of interviews), exemplifying (60.9%), or repeating questions/instructions (50.7%). Child age, seizure history, verbal communication abilities, adaptive behaviour, and interviewer identity predicted adaptation use. Adaptation use did not differ between the two data collection points, 13 months apart.

Conclusion: The CATQ enhances methodological rigor by tracking adaptations and facilitating risk-of-bias-assessment by analysing adaptation changes and factors affecting their use.

ARTICLE HISTORY

Received 26 June 2024
Revised 17 December 2024
Accepted 15 January 2025

KEYWORDS

Self-report; adaptations; mental health; children; neurodisability; neurodevelopmental disorders; methods

> IMPLICATIONS FOR REHABILITATION

- More children with neurodisability may be able to report their subjective experiences by allowing some level of individualisation of interview-administered patient-reported outcomes.
- A novel interviewer-rated questionnaire called the Cognitive Accessibility Tracking Questionnaire (CATQ) can be used to keep track of such changes.
- CATQ can help researchers and clinicians monitor adaptations, assess the cognitive and communicative accessibility of PROMs, and evaluate assessment training.

Introduction

Children with neurodisability (ND) are heterogeneous in terms of type and level of impairments [1], making standardised data collection on subjective phenomena like well-being and mental health problems challenging [2]. Allowing for individual tailoring of materials and procedures to the needs of individual participants could enable more children with ND to provide self-ratings. However, a more flexible approach introduces the risk of measuring different phenomena between individuals or over time in the same individual. Therefore, it is essential to track deviations from standardised procedures. In this study, we introduce and test a novel instrument designed to measure the type and number of changes made to facilitate the understanding and response to

questions in patient-reported outcome measures (PROMs) when undertaken as an interview.

Levels of functioning and the need for support vary not only between different groups within the ND umbrella [1] but also within each diagnosis. For example, children with very severe intellectual disability (ID) largely rely on support persons' abilities to interpret their behaviours communicatively, whereas children with mild ID typically communicate verbally [3]. In age-mixed samples, developmental changes further contribute to this heterogeneity.

When collecting data on largely subjective experiences from groups characterised by such diversity, researchers and clinicians face the challenge of finding a data collection method suitable for the entire group. For children with the most severe cognitive

CONTACT Magnus Ivarsson  magnus.ivarsson@liu.se 

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

impairments, any data collection method reliant on the child's ability to self-assess internal states and self-report is unrealistic [2,4]. However, for many other children with ND, such as those with mild ID or movement-related impairments, there is no clear rationale for why the omission of self-reporting would be acceptable.

One frequent solution to this problem is to divide the broad ND group into diagnostic subgroups and study them separately with different methods (e.g. proxy and self-report). Unfortunately, this complicates direct comparisons between groups due to the relatively low inter-informant correlations [5,6]. Another approach is to develop PROMs accessible to children with ND, who may struggle to provide valid answers on tools designed for typically developing children [7,8]. However, accessible PROMs are scarce for many outcomes, and developing universally accessible PROMs for all impairment types is nearly impossible. In practice, many studies use methods that, while suboptimal, are feasible across different levels of functioning, typically relying on proxy ratings (e.g. [9,10]).

However, choosing data collection methods often involves tradeoffs between different aspects of validity. For example, parental ratings of children's emotional states offer the practical advantage of not relying directly on the child's cognitive and communicative abilities, thus allowing for a more diverse representation of functioning levels within the sample. This broader diversity may improve the representativeness of the sample and thereby enhance the generalisability of findings and the external validity of the study. However, there are both rights-based [11] and methodological [5,6,12,13] arguments that can be made for facilitating self-ratings. Importantly, proxy ratings may not fully capture the child's internal experience, potentially threatening internal validity. This threat becomes more pronounced as the subjectivity of the phenomenon increases. When studying intrinsically subjective phenomena such as life satisfaction and well-being, not involving subjective experiences is questionable [2,14]. Thus, advantages in internal validity relating to self-rating must be weighed against possible advantages in external validity in parent ratings when choosing between the strategies.

A supplementary approach that could make self-report a feasible option for a larger proportion of children with ND is to allow some degree of individual tailoring of the materials and/or procedures used in data collection to meet the needs of individual children (henceforth referred to as "adaptations"). For example, strategies used in augmentative and alternative communication (AAC), such as pictorial support, gestures, or sign language [15], or the literature on accessible PROMs (e.g. [7,8]), could be encouraged in cases where they increase the chances of the participant understanding and responding to a question as intended. Allowing for adaptations could tilt the balance between external and internal validity in favor of self-rating in some cases. Simultaneously, there is also the risk that adaptations change what is being measured. Thus, assessing the risk of bias introduced by allowing some degree of adaptation is essential.

When allowing for adaptations, there are several hypothetical mechanisms through which a risk of bias might be introduced. For example, in longitudinal studies, the use of adaptations may change over time, distorting analyses of temporal trends. In studies with multiple data collectors, interviewer factors may lead to differences in adaptation strategies. Ideally, the choice of adaptations should be guided by child functioning and the nature of the information being communicated rather than by the preferences of the interviewer. Key to selecting adequate adaptations is to analyse the cognitive and communicative demands inherent

to a specific PROM. For children with intellectual and cognitive impairments, it is essential to analyse the cognitive accessibility of PROMs—i.e., the degree to which the design anticipates different levels of cognitive functioning and reduces cognitive demands [8].

A crucial first step in understanding the risk of bias associated with the use of adaptations in data collection is to systematically measure their use. Next, the rigor of studies allowing individualised adaptations as a strategy can be strengthened by analysing properties of these adaptations, such as the influence of irrelevant variables on adaptation use or shifts in adaptation use over time. Therefore, developing a measure to document the adaptations made to self-rated assessments is a priority.

Aims and research questions

The current study introduces the Cognitive Accessibility Tracking Questionnaire (CATQ), an instrument designed to measure adaptations made to procedures and materials in interview-administered PROMs. We used the CATQ in a longitudinal study of participation and mental health in children with ND, with multiple PROMs with varying content and layout complexity. The study aimed to (i) describe the CATQ's use in measuring the different forms of adaptations and their respective frequencies and (ii) explore how the CATQ could be helpful when assessing the potential for bias stemming from individualised adaptations. The second aim was addressed by analysing the factors relating to adaptation use and the consistency of such adaptations over time. The research questions were:

1. What types of adaptations do interviewers report using?
2. How did the frequency of adaptation use differ between PROMs?
3. Which factors are predictive of the quantity of various types of adaptations utilised in interviews?
4. How consistent is the use of different adaptations over time?

Materials and methods

The current study was conducted within the Child-Participation and Mental Health (Child-PMH) research program, funded by the Swedish Research Council (2018-05824_VR). The longitudinal component of Child-PMH, which includes the present study, aims to investigate the longitudinal trajectories and interrelationships between aspects of participation in everyday activities, mental health problems, and well-being. Ethical approval for the research was obtained after review by the Swedish National Ethics Review Authority in 2019 (2019-05028). Caregivers of the children provided written consent to participate after being informed about the project *via* mail. The children were informed about the project and consented verbally before commencing the interview.

Participants and procedure

Families participating in Child-PMH were recruited through the habilitation services in five regions in Sweden. In Sweden, child habilitation is eligible for children (age 0-16/18) with ND such as ID, autism, and cerebral palsy. Typically, children with more severe impairments are more likely to receive habilitation services than those with less severe impairments.

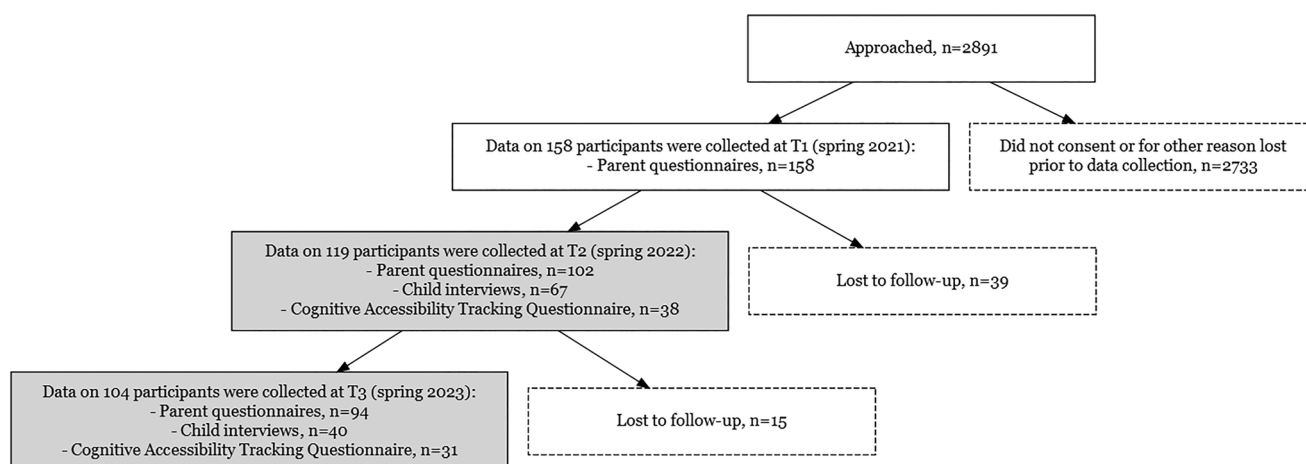


Figure 1. Flow chart displaying the data collected at the first (T1), second (T2), and third (T3) Wave of data collection (with grey representing data used in the current study).

Table 1. Characteristics of the participating children.

	ND group
N	43
Girls	24
Age (years)	
M	12.28
SD	2.71
Disability measured with TenQQ (%)	
Serious delay in sitting/standing/walking	18.84
Difficulties seeing	17.39
Difficulties hearing	15.94
Does not comprehend parent	15.94
Difficulties walking or moving arms	20.29
Has seizures/becomes rigid/lose conscience	5.80
Has not learned the same things as peers	39.13
Does not speak at all	8.96
Can not mention one object	7.25
Difficulties comprehending/is mentally slow	55.07

Note. Abbreviations used in the table: ND (Neurodisability), TenQQ (Ten Questions Questionnaire).

Written information about the project (a caregiver-version and a simplified child-version) was sent *via* mail to the caregivers of all children born in 2013-2015 and 2007-2009 who were enlisted at the habilitation centers in spring 2020. The information letters were written in Swedish but contained information about where to find Arabic, Somali, and English translations. Details about the data collection are displayed in Figure 1.

Child interviews were originally planned for all waves of data collection but could not be administered at T1 due to the Covid-19 pandemic (the parent-rated questionnaires were distributed *via* an online survey or mail). Further, only the child interviews for which the interviewer also assessed adaptations with the CATQ were included and analysed in the current study, resulting in data from six interviewers and 43 children being interviewed once or twice (on T2 and/or T3). Approximately 13 months ($m=397.22$ days) passed between T2 and T3. Interviewer and child characteristics are presented in Tables 1 and 2.

The child interviews were conducted at the preferred location of the families, including at schools, habilitation centers, and the children's homes. A majority (53.49%) of the interviews were conducted *via* a video communication tool (for a detailed description of this approach, see [16]). Interviewers were instructed to complete the CATQ directly after having finished an interview. The Ten Questions Questionnaire (TenQQ; [17,18]) was administered as a part of the caregiver questionnaire that was sent to caregivers *via* mail or e-mail in temporal proximity to the child interviews.

Table 2. Interviewer characteristics.

	Interviewer					
	1	2	3	4	5	6
Interviewing children						
Self-rated level of expertise	Nov	Int	Int	Int	Exp	Int
Years of experience	1	20	5	20	17	3
Interviewing children with disability						
Self-rated level of expertise	Nov	Int	Int	Int	Exp	Int
Years of experience	0	15	15	20	17	3
Self-rated AAC expertise	Nov	Int	Nov	Int	Int	Int
Number of interviews	20	2	4	13	2	28

Note. Abbreviations used in the table: AAC (Augmentative and alternative communication), Exp (Expert), Int (Intermediate), Nov (Novice).

Material

Development of the cognitive accessibility tracking questionnaire

The CATQ was developed to measure the types and frequency of adaptations used by researchers and clinicians when supporting children to participate in interview-administered PROMs. The first draft contained 12 items reflecting different strategies to increase cognitive and/or communicative accessibility. The strategies were chosen based on literature from the fields of alternative and augmentative communication [15] and accessible PROM design [8]. The draft was independently reviewed by nine researchers and clinical experts within relevant fields, including occupational therapy, speech and language therapy, psychology, disability research, and child rehabilitation/habilitation. The reviewers were instructed to comment on whether the most relevant forms of adaptations were represented in the questionnaire and, if not, suggest new items and other changes to the questionnaire. The review led to items being merged, rephrased, added, and exemplified with concrete examples. The current Swedish version of the questionnaire (available at <https://osf.io/5rz9y/>) comprises 13 items, with a three-level Likert-style response scale: "Never (followed manual)" (scored as 0), "A few times (beyond manual)" (1), and "Several times (beyond manual)" (2). The CATQ is completed by the researcher directly after having administered one (or more) PROM(s) undertaken as an interview. Only adaptations that are not part of the standardised procedures and materials of a particular PROM are reported using CATQ. Apart from the 13 items, CATQ also contains an open-ended question about adaptations made during the interview not covered by the 13 items. A total frequency of adaptations score (scored 0-26) can be calculated by summing the score of all items in the instrument. Further, a total types of adaptations score can be calculated by

summing scores after transforming the items to binary variables by merging the two positive response alternatives (scored 0-13).

For the present study, six questions concerning the interviewer's background and experience in conducting interviews were added to the questionnaire: (1) Prior experience of interviews (structured conversations, testing, etc.) with children, (2) Prior experience of interviews (structured conversations, testing, etc.) with children with disabilities, (3) Knowledge about methods (i.e., AAC) to facilitate understanding and communication with children with disabilities, (4) Years of experience working with therapy, interviews, or another professional task with children in clinic and/or research, (5) Years of experience working with therapy, interviews, or another professional task with children with disabilities clinically and/or in research, and (6) Profession/higher education degree.

Ten questions questionnaire

To characterise the sample, and to explore possible predictors of adaptation use, the presence of different disability-related problems was measured with the TenQQ. TenQQ was developed to screen for ID, blindness, deafness, movement, and seizure disorders in low and middle-income countries. As such, it consists of ten questions for which the respondent (the primary caregiver) is to indicate the presence or absence of difficulties and/or delays ("Yes" = 1, "No" = 2). In the present study, some of the TenQQ variables were transformed so that a lower score ("1") indicated the presence of difficulty or delay in all cases.

PROMs administered in the child interviews

The participating children responded to items from five standardised PROMs during the interviews, which were administered in the same order in a single session, lasting approximately 45-60 min. The first PROM, Picture my Participation, measures participation in terms of the frequency of attendance and the intensity of involvement in 20 everyday activities (e.g. personal care, school). Developed for children and young people aged 5 to 21 years, Picture my Participation is described as a guided conversation rather than a self-rating scale [19]. It includes visual supports for activities and the three- and four-level Likert-style response scales, and it is administered using a Talking Mats approach [20]. Four dimensions of the frequency of attendance have been identified in an assessment of the structural validity of the PROM, with acceptable internal consistency for the total frequency of attendance scale in children with ID [21].

The second measure was the self-report version of the Strength and Difficulties Questionnaire [22], a 25-item PROM designed to assess conduct problems, emotional problems, hyperactivity, peer problems, and prosocial behaviour in children and adolescents aged 11-17 years over the last six months. Each item is scored on a three-point Likert-type: "not true", "somewhat true", and "certainly true". The total problems scale and the subscales (except peer problems) have been found to have acceptable internal consistency in children with ID [23]. However, evidence also suggests that a three-factor solution may be more appropriate than the original five-factor model [24].

The third part of the interview consisted of two items (pain/discomfort and being sad/unhappy) from EQ-5D-Y [25], a measure of health-related quality of life developed for use in children from eight years of age. Both items are rated on a three-point Likert-type scale: "no", "some", and "a lot". EQ-5D-Y has been found to be feasible for use in children and adolescents with functional

motor, orthopedic and medical disabilities, with acceptable discriminative validity [26].

The fourth PROM consisted of eight items from the Health Behavior in School-aged Children (HBSC) survey [27] which measures factors related to friendships and attitudes toward various aspects of the school environment in children aged 11, 13, and 15 years. The HBSC is a cross-national study conducted in collaboration with the World Health Organization, focusing on health, well-being, and related behaviours in school-aged children. The selected items for the present study varied in both the number of response options (four or five steps) and the overall design of the scale. The HBSC survey has been administered in diverse populations, including children with physical disabilities such as cerebral palsy and spina bifida [28].

The fifth PROM used was a modified version of the Cantril Ladder [29], which measures global life satisfaction using a single-item approach. Participants were presented with an image of a ten-step ladder and instructed to select the step that best represented their current life satisfaction, where the top step represented the best possible life they could imagine and the bottom step represented the worst possible life. The Cantril Ladder has demonstrated good test-retest reliability and convergent validity in a sample of adolescents aged 11 to 15 years [30].

Data analysis

Handling of missing data

Missing data was imputed before running the regression analysis through a combination of strategies. For missing TenQQ values (11.55%), data was imputed with the equivalent value from the subsequent wave of data collection (T2 data to T1; T3 data to T2). This procedure was justified by the relative stability of TenQQ items across waves (the average T1-T2 correlation was 0.72). The average rate of missing data across the CATQ items was low (1.11%). Missing CATQ data was replaced by multivariate imputation by chained equations with the *mice* R-package [31]. Missing data for the CATQ items were imputed in one step based on all CATQ items, using one iteration.

Identifying predictors

The relationship between child age, sex, disability status (as measured by the TenQQ), and the items in the CATQ were analysed in two steps. First, correlations were examined with the *mixedCor()* function from the *psych* R-package [32]. This function determines the appropriate correlation coefficient for variables based on the characteristics of the data. Secondly, a multiple linear regression approach was employed to examine the factors influencing the number of adaptations used in the interviews. The dependent variable was regressed on the following predictors: (1) child age, (2) sex, (3) disability (as indicated by the item in the TenQQ), and (4) interviewer identity (i.e., who it was that conducted the interview). To include the interviewer identity in the analysis, a dummy coding scheme was applied. Five binary dummy variables were created to represent the interviewer identity, with the sixth interviewer (the one with the least experience in interviewing children with ND) serving as the reference category. The analyses utilised the *lm()* function in R [33].

Change over time

To assess changes in the overall use of different forms of adaptations from T2 to T3, a dependent t-test was conducted.

Additionally, agreement between T2 and T3 at the CATQ item level was analysed by calculating the percentage of agreement (number of T2-T3 agreements divided by the total number of assessments) and linearly weighted Cohen's kappa (κ_w) with the *vcd* R-package [34] for participants with longitudinal data. The interpretation of κ_w results followed [35] benchmarks for the strength of agreement: <0.00 = "Poor", $0.00-0.20$ = "Slight", $0.21-0.40$ = "Fair", $0.41-0.60$ = "Moderate", $0.61-0.80$ = "Substantial", $0.81-1.00$ = "Almost perfect".

Results

Characteristics of the adaptations used

Most interviews (85.51%) involved one or more adaptations to the standardised procedure and/or material. Both the median types of adaptations used and the median total frequency of adaptations to the interviews were 4 (with an IQR = 5 in both cases). As illustrated by Figure 2, some forms of adaptations were deployed rarely or not at all in the interviews (i.e., technical aids, pictures/objects, manual signs, and or adaptations to the physical

environment). The two most frequent forms of adaptations were to explain/replace a concept in a question and to clarify the meaning of a question with examples. Depending on the interviewer, the minimum number of different types of adaptations employed in an interview ranged from 0 to 4 and the maximum from 3 to 10 (see Figure 3).

Adaptations occurred with all administered PROMs but to different degrees. The PROM for which adaptations were made most frequently was Picture my Participation (73.53%), followed by the items from the Health Behavior in School-aged Children survey (55.88%), Strength and Difficulties Questionnaire (30.88%), the items from EQ-5D-Y (16.18%), and Cantril Ladder (8.82%). This order was preserved when comparing adaptations to PROMs from the two waves of data collection.

Some types of deviations from the procedure and/or material described in the study protocol that were not covered by the Likert-style items in CATQ appeared in the answers to the open-ended question. Most common among these ($n=10$) was some level of involvement of a third person (parent, translator, assistant, or teacher) in the interview. For example, the use of a sign language interpreter, a parent translating a concept to the

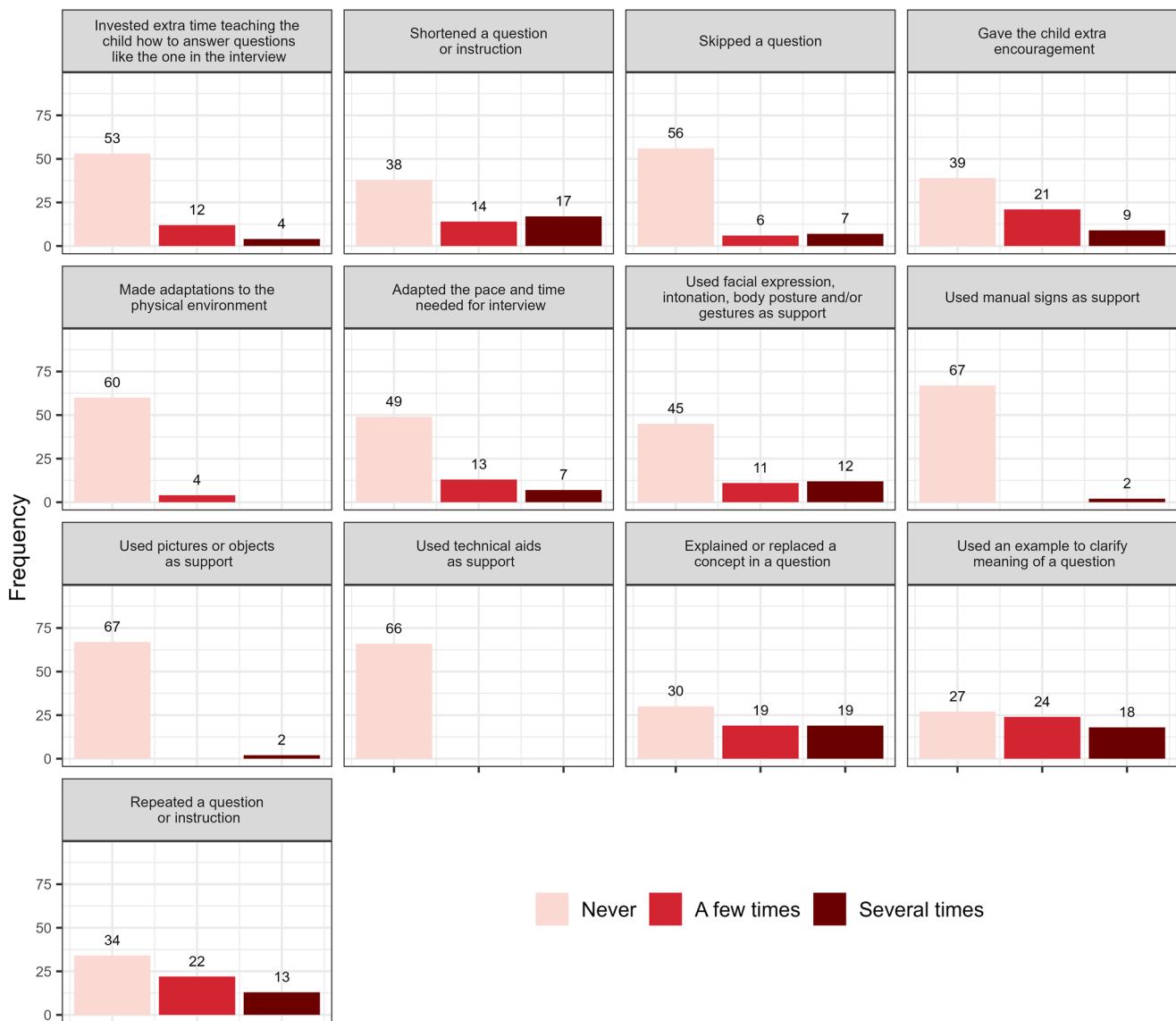


Figure 2. Cognitive accessibility tracking questionnaire frequencies.

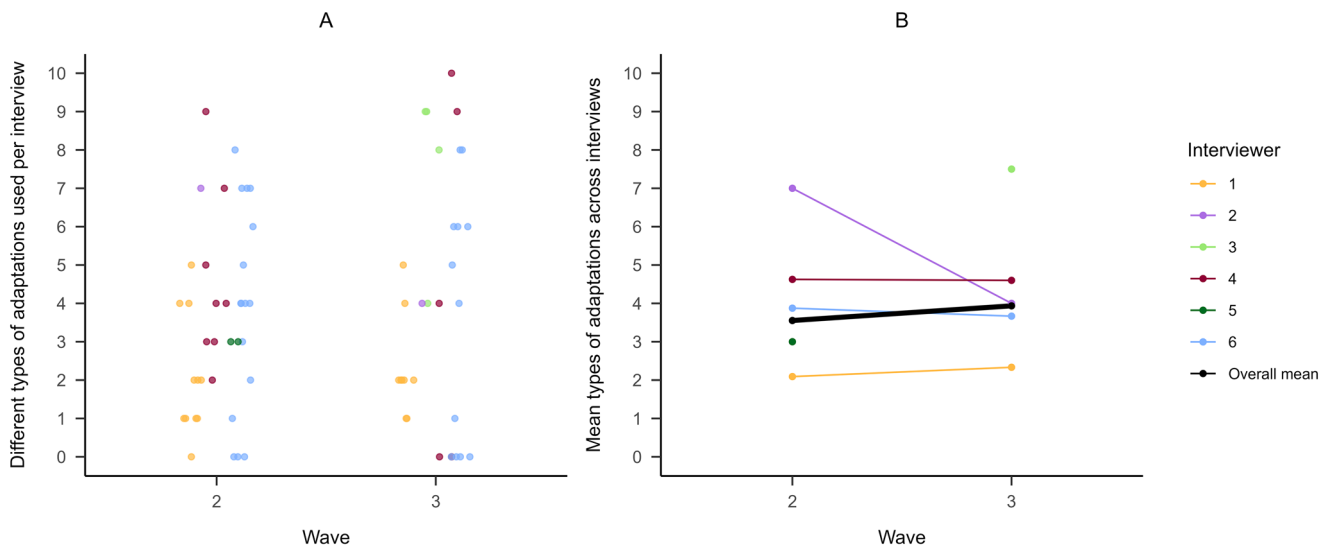


Figure 3. Sum of the different types of adaptations made (A) and the mean number of adaptations made (B) to the procedure and material during interviews at the second and third wave of data collection.

Table 3. Correlations between child age, sex, disability status, and the sum of the different forms of adaptations used during the interviews.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Age	1.00												
2 Sex	-0.20	1.00											
3 Serious delay sitting, standing, or walking	0.06	0.26	1.00										
4 Difficulties seeing	-0.13	0.08	0.43	1.00									
5 Difficulties hearing	0.26	0.02	0.52	0.23	1.00								
6 Comprehends when asked to do something	0.04	0.02	0.02	0.22	0.01	1.00							
7 Difficulty walking, moving arms	0.19	-0.17	0.14	0.01	0.12	-0.32	1.00						
8 Seizures, rigidity, or loses consciousness	0.05	0.27	0.36	0.53	0.34	0.20	0.58	1.00					
9 Has learned to do the same things as peers	-0.15	0.23	0.45	0.31	0.24	-0.05	-0.39	0.10	1.00				
10 Speak at all	0.05	-0.07	0.07	-0.56	0.17	0.21	-0.62	-0.66	0.16	1.00			
11 Can mention at least one object	0.29	-0.16	0.03	-0.63	0.22	-0.61	-0.08	-0.60	0.19	0.57	1.00		
12 Difficulties comprehending/is mentally slow	0.06	0.06	0.52	0.34	0.33	0.49	-0.33	0.17	0.77	0.30	-0.04	1.00	
13 Total types of adaptations used	-0.35	-0.09	-0.30	-0.21	-0.23	-0.36	0.29	-0.49	-0.58	-0.80	-0.43	-0.57	1.00

child's mother tongue, or exemplifying an abstract concept with examples from the child's everyday life. The other adaptations mentioned—occurring in 1–3 interviews each—included the interviewer translating questions or words to English, allowing the child to answer exclusively with the annotate function in the video meeting application (enabling marking answers directly in a shared presentation slide), activating the child to maintain focus, and play to socially connect with the child.

Predicting the total types of adaptations made

There was a strong (>0.7) correlation between the total types of adaptations made during the interviews and the absence of verbal expressive communication in the child (see Table 3). Two other factors related to disability/chronic health—adaptive behaviour deficits (i.e., not having learned the same things as peers) and difficulties comprehending/being mentally slow—correlated with the types of adaptations in the moderate range (>0.5 – 0.7). However, when combined in the multiple linear regression model, partly different variables emerged as significant predictors (see Table 4). Specifically, a lower child age at the time of the interview, a history of seizures, adaptive behaviour deficits, a lack of verbal expressive communication, and the interview being conducted by interviewer number three all contributed to predicting a higher number of different adaptations being deployed in an interview. In all, the model significantly predicted the total types of adaptations made to interviews, $R^2 = 0.74$, $F(17,49) = 8.05$, $p < 0.001$.

Change over time

For the children with longitudinal data ($n = 26$), the average number of different types of adaptations deployed during interviews did not change significantly between the second and third wave of data collection, $M_D = 0.50$, 95% CI $[-0.13, 1.13]$, $t(25) = 1.64$, $p = 0.114$. Panel B of Figure 3 displays the change in the number of adaptations deployed from T2 to T3 for each interviewer. For a single child, changes between T2 and T3 in the sum of types of adaptations deployed spanned from -4 to 2 .

The percentage of agreements between the use of each type of adaptation at T2 and T3 ranged from 53.85% to 100.00%. The tests for κ_w were significant, indicating a nonrandom level of agreement, for six of 13 items (see Table 5). The κ_w coefficients were in the almost perfect range for one item ("Sign language"), in the moderate for four ("Explained/exchanged term", "Used examples", "Repeated", and "Shortened"), and in the fair for one ("Body language").

Discussion

This study introduced a novel scale to measure adaptations in materials and procedures made spontaneously in an interview situation to facilitate the self-rating of subjective phenomena in children with ND. Three potential applications of the CATQ were demonstrated: (i) tracking types and frequency of adaptations, (ii) analysing factors associated with adaptation use, and (iii) assessing

Table 4. Statistics for the predictor variables in the linear regression analysis.

Predictor	<i>b</i>	95% CI	<i>t</i>	<i>df</i>	<i>p</i>
Intercept	24.44	[17.09, 31.80]	6.68	49	<0.001
Child age	-0.32	[-0.51, -0.13]	-3.41	49	0.001
Child sex	-0.37	[-1.24, 0.51]	-0.84	49	0.405
Serious delay in sitting/standing/walking	0.46	[-0.81, 1.73]	0.73	49	0.468
Difficulties seeing	-1.16	[-2.50, 0.17]	-1.75	49	0.086
Difficulties hearing	0.59	[-0.71, 1.89]	0.91	49	0.368
Does not comprehend parent	-0.79	[-2.10, 0.53]	-1.20	49	0.235
Difficulties walking or moving arms	1.16	[-0.14, 2.46]	1.79	49	0.080
Has seizures/becomes rigid/lose conscience	-2.73	[-4.77, -0.68]	-2.68	49	0.010
Has not learned the same things as peers	-2.56	[-3.93, -1.19]	-3.76	49	<0.001
Does not speak at all	-4.53	[-6.50, -2.55]	-4.60	49	<0.001
Can not mention one object	0.27	[-1.71, 2.25]	0.28	49	0.783
Difficulties comprehending/is mentally slow	0.38	[-0.96, 1.72]	0.57	49	0.572
Interviewer 2	2.39	[-0.43, 5.21]	1.70	49	0.095
Interviewer 3	3.52	[1.38, 5.65]	3.31	49	0.002
Interviewer 4	0.53	[-0.89, 1.94]	0.75	49	0.459
Interviewer 5	-0.89	[-3.69, 1.91]	-0.64	49	0.526
Interviewer 6	0.51	[-0.64, 1.65]	0.89	49	0.378

Table 5. Percent of agreement and weighted kappa (κ_w) statistics for the cognitive accessibility tracking questionnaire items at the second and third wave of data collection.

Item	Second wave			Third wave			%agree.	κ_w			
	Never	A few times	Several times	Never	A few times	Several times		Estimate	ASE	<i>z</i>	<i>p</i>
1. Extra time	23	2	1	19	6	1	73.08	0.22	0.16	1.37	0.17
2. Adapted pace	18	6	2	19	5	2	69.23	0.24	0.17	1.40	0.16
3. Body language	17	3	6	19	5	2	61.54	0.34	0.17	2.03	0.04
4. Sign language	25	0	1	25	0	1	100.00	1.00	0.00	∞	0.00
5. Pictures/objects	26	0	0	26	0	0	100.00				
6. Technical devices	26	0	0	26	0	0	100.00				
7. Explained/ exchanged term	11	6	9	13	6	7	69.23	0.59	0.13	4.42	0.00
8. Used examples	8	12	6	14	7	5	53.85	0.41	0.14	2.92	0.00
9. Repeated	12	10	4	16	5	5	69.23	0.57	0.14	4.13	0.00
10. Shortened	15	3	8	15	5	6	65.38	0.57	0.13	4.40	0.00
11. Skipped	21	4	1	22	1	3	76.92	0.38	0.22	1.72	0.09
12. Encouragement	16	7	3	15	7	4	53.85	0.25	0.16	1.58	0.11
13. Physical environment	24	2	0	26	0	0	92.31	0.00	0.00	0.00	1.00

Note. It was not possible to calculate κ_w for items 5 and 6 since there was full agreement between waves of data collection. Abbreviations used in table: %agree. (percent agreement), ASE, Approximate Standard Error. Cognitive Accessibility Tracking Questionnaire items are described in Figure 2 note.

consistency over time. The results indicated that a variety of adaptations were applied during interviews in a longitudinal study of mental health and participation in children with ND. The most common adaptations involved aspects of verbal communication (explaining or replacing concepts and clarifying through examples), with the total number of adaptations varying across PROMs. Adaptation use was predicted by the child's age, a history of seizures, adaptive behaviour deficits, lack of verbal communication, and the person conducting the interview. Additionally, the overall use of adaptations did not change between the two data collection points, although agreement over time varied for specific types of adaptations.

Factors related to the use of adaptations

Theoretically, the use of adaptations can be influenced by several factors, including those related to the characteristics of the adaptation (e.g. availability and cost), the layout and content of the PROM(s), the child (e.g. cognitive and communicative functioning), and the interviewer (e.g. knowledge and experience of AAC and other strategies to increase accessibility). Ideally, the adaptation that most efficiently bridges the gap between the child's capabilities and the intrinsic demands of

the PROM is prioritised. The interviewer's knowledge of AAC methods and/or attitudes toward a specific method should only influence the process when multiple adaptations are equally effective in increasing accessibility without changing what is being measured.

In line with these theoretical considerations, the present study found that less resource-demanding adaptations (i.e., verbal behaviours) were preferred over more complex adaptations, such as the use of technical aids or pictures/objects. This suggests that adaptation simplicity may be part of the explanation for the inconsistent use of different adaptations. Furthermore, differences in the degree to which PROMs were adapted also suggest that PROM complexity may have influenced the decisions.

Picture My Participation was adapted most frequently, which may seem counterintuitive given that it was specifically designed for children with disabilities [19,21]. Unlike many other PROMs, Picture My Participation incorporates visual representations of activities and response options. Although it has an accessible design, Picture My Participation does tap into abstract and complex constructs, such as "involvement in activities". Therefore, one possible explanation for the large number of adaptations being used is that more complex outcomes may require greater adjustments.

The PROM with the fewest adaptations, the Cantril Ladder, measures “life satisfaction”, which may be perceived as less complex than the constructs in Picture My Participation. It is important to emphasise that Picture My Participation was designed as a guided interview, not a self-rating scale, meaning that individualisation and flexibility in administration are intentional features. Taken together, these points suggest that while PROMs vary in their cognitive accessibility, differences in the level of abstraction in their core concepts may still influence the need for adaptations.

The results from the regression analysis show that child-related factors—such as younger age, differences in learning compared to peers, and non-verbal communication—were predictive of the extent of adaptations used. These findings are not unexpected: they indicate that appropriate factors were at least partly considered when selecting adaptations. More concerning, however, is that who conducted the interview was a significant predictor. This is one indication that risk of bias may have been introduced by the adaptation strategy, as it could indicate that interviewer preference, rather than a combination of child and PROM factors influenced the choice of adaptations.

Interviewers’ self-rated expertise and knowledge about child interviewing and AAC varied considerably. It is possible that the interviewer training in Child-PMH, which was restricted to general information rather than training in specific techniques, was not sufficient to compensate for these differences. However, the interviewers’ reported AAC competence, prior experience, and expertise in interviewing children seemed sufficient to help them recognise when a child did not understand a question, as all of them reported using some adaptations.

Changes over time

Tracking deviations from the data collection protocol is important in any study, but even more so when individualised adaptations are part of an explicit strategy. In longitudinal studies, changes in adaptation use over time can introduce bias. The present study found that the mean use of adaptations remained stable between the two waves of data collection, suggesting that any effect of adaptations at an aggregated level might be consistent over time. However, the levels of agreement between specific types of adaptations over time varied, falling within the fair to almost perfect range in slightly more than half of the cases where agreement could be analysed. Such variability in adaptations could pose challenges if different forms of adaptations introduce varying levels of bias. However, to the authors’ knowledge, no studies have quantified the risk of bias introduced by different specific adaptations in the administration of assessments.

While stability in the mean use of adaptations over time may suggest consistency, it could also theoretically mask changes in how specific PROMs were adapted across the two data collection waves. However, the results indicate that the most frequently adapted PROMs remained consistent across both data collection points. Taken together, the overall stable pattern of adaptations observed suggests that the risk of bias related to changes in adaptation use may have been limited.

In summary, the clearest indication that some degree of bias may have been introduced by allowing spontaneous adaptations is that interviewer identity was a significant predictor of adaptation use. While the exact degree of bias introduced by specific adaptations remains unclear, it is reasonable to assume that some level of bias exists and may vary by adaptation type. Whether this risk is acceptable or not depends on how self-report is valued. Two key aspects should be considered in this valuation: the

subjectivity of the phenomena being studied and the increase in accessibility provided by adaptations. Thus, the inherently subjective phenomena in focus for the present study suggested that the strategy to allow for spontaneous adaptations could be justifiable despite the signs of some risk of bias being introduced.

Implications for practice and future research

While the sample in the current study was limited to children with ND, there is no apparent reason why the CATQ should not apply to groups of adolescents and adults with similar diversity in functioning. In studies where the sample is more functionally homogeneous, designing a standardised procedure accessible to the entire study group is appropriate, and the need for individual tailoring of materials and procedures should be reduced. In diverse ND populations, the CATQ has several potential applications in both research and clinical settings:

- Helping interviewers track and consistently apply adaptations in studies with repeated measures designs.
- Assessing the risk for bias as a consequence of interviewer factors (e.g. experience and expertise) influencing the use of adaptations rather than child functioning.
- Contributing to the assessment and validation of interviewer training by indicating successful training if interviewer factors do not predict the use of adaptations.
- Evaluating the cognitive accessibility of PROMs, with more frequent adaptations suggesting less accessible designs.
- When developing new PROMs, the CATQ could be used to suggest strategies for a more accessible material and/or standardised procedure.
- Serving as a variable for control in statistical models, given large enough samples.

Future studies should investigate the psychometric properties of the CATQ, including aspects of structural validity, inter-rater reliability (potentially using filmed interviews), and internal consistency. Some items may be redundant while others may be missing, but establishing this requires assessment across multiple interviewers with varying levels of expertise. Since the involvement of a third person (e.g. parent, interpreter) was noted in the open-ended question of the CATQ in the current study, adding an item on this topic could strengthen the scale

Limitations

The results of the present study need to be interpreted with caution due to the limited number of interviewers participating in the study. This limitation prevented a quantitative analysis of the influence of prior interviewer experience and expertise on the use of adaptations. Unfortunately, this limitation is not unique to the present study. To explore the influence of such factors, data from several comparable studies may need to be aggregated. Another limitation of the CATQ is its reliance on the subjective experience of the interviewer. It is possible that interviewers either under- or overestimate the use of adaptations in interviews. Therefore, exploring the inter-rater reliability of the scale is a priority together with other aspects of reliability and validity. Furthermore, there is no clear evidence that all adaptations in the CATQ increase cognitive and/or communicative accessibility (i.e., lower cognitive demands without changing what is measured). The adaptations were chosen based on experience and

convention [8,15] rather than evidence. However, all items reflect strategies used in clinical and research settings and are, therefore, relevant.

Conclusions

The interviewer-rated instrument CATQ introduces the opportunity to measure and quantify the use of adaptations whenever interview-administered PROMs are used with groups of children with ND and diverse levels of functioning. As demonstrated by the present study, CATQ can be used to track changes and predictors of adaptations. Both are important in the assessment of the risk of bias introduced by a procedure allowing for individualised adaptations. The stability of adaptations, as observed in this study, suggests that possible biases introduced when allowing for adaptations may be of less importance when interpreting longitudinal changes. However, the retention of interviewer-related factors in the prediction models indicates the introduction of some level of bias. Thus, the results highlight the need for a common level of knowledge, skills, and attitudes toward adaptations among interviewers in projects allowing individual tailoring of material and procedures to reduce the risk of bias.

Author note

The authors made the following contributions. Magnus Ivarsson: Conceptualisation, Formal analysis, Investigation, Methodology, Project administration, Visualisation, Writing - original draft, Writing - review & editing; Henrik Danielsson: Data curation, Funding acquisition, Methodology, Supervision, Validation, Writing - review & editing; Lena Almqvist: Funding acquisition, Methodology, Supervision, Writing - review & editing; Christine Imms: Funding acquisition, Methodology, Supervision, Writing - review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Swedish Research Council (grant number 2018-05824_VR).

ORCID

Magnus Ivarsson  <http://orcid.org/0000-0002-5456-1597>
Henrik Danielsson  <http://orcid.org/0000-0002-0446-0827>
Christine Imms  <http://orcid.org/0000-0001-9055-3554>

Data availability statement

A reproducible version of the manuscript, including the code, and a synthetic dataset with similar properties as the original dataset can be found at <https://osf.io/5rz9y/>.

References

- [1] Morris C, Janssens A, Tomlinson R, et al. Towards a definition of neurodisability: a delphi survey. *Dev Med Child Neurol.* 2013;55(12):1103–1108. doi: [10.1111/dmnc.12218](https://doi.org/10.1111/dmnc.12218).
- [2] Fujiura GT, the RRTC Expert Panel on Health Measurement. Self-reported health of people with intellectual disability. *Intellect Dev Disabil.* 2012;50(4):352–369. doi: [10.1352/1934-9556-50.4.352](https://doi.org/10.1352/1934-9556-50.4.352).
- [3] American Psychiatric Association. *Diagnostic and statistical manual of mental disorders.* 5th ed. Arlington (VA): American Psychiatric Association; 2013. doi: [10.1176/appi.books.9780890425596](https://doi.org/10.1176/appi.books.9780890425596).
- [4] Finlay WML, Lyons E. Methodological issues in interviewing and using self-report questionnaires with people with mental retardation. *Psychol Assess.* 2001;13(3):319–335. doi: [10.1037/1040-3590.13.3.319](https://doi.org/10.1037/1040-3590.13.3.319).
- [5] Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychol Bull.* 1987;101(2):213–232. doi: [10.1037/0033-2909.101.2.213](https://doi.org/10.1037/0033-2909.101.2.213).
- [6] De Los Reyes A, Augenstein TM, Wang M, et al. The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychol Bull.* 2015;141(4):858–900. doi: [10.1037/a0038498](https://doi.org/10.1037/a0038498).
- [7] Kooijmans R, Mercera G, Langdon PE, et al. The adaptation of self-report measures to the needs of people with intellectual disabilities: a systematic review. *Clin Psychol: Sci Practice.* 2022;29(3):250–271. doi: [10.1037/cps0000058](https://doi.org/10.1037/cps0000058).
- [8] Kramer JM, Schwartz A. Reducing barriers to patient-reported outcome measures for people with cognitive impairments. *Arch Phys Med Rehabil.* 2017;98(8):1705–1715. doi: [10.1016/j.apmr.2017.03.011](https://doi.org/10.1016/j.apmr.2017.03.011).
- [9] Downs J, Blackmore AM, Epstein A, et al. The prevalence of mental health disorders and symptoms in children and adolescents with cerebral palsy: a systematic review and meta-analysis. *Dev Med Child Neurol.* 2018;60(1):30–38. doi: [10.1111/dmnc.13555](https://doi.org/10.1111/dmnc.13555).
- [10] van Steensel FJA, Bögels SM, Perrin S. Anxiety disorders in children and adolescents with autistic spectrum disorders: a meta-analysis. *Clin Child Fam Psychol Rev.* 2011;14(3):302–317. doi: [10.1007/s10567-011-0097-0](https://doi.org/10.1007/s10567-011-0097-0).
- [11] United Nations Convention on the Rights of the Child. 1989. <https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-child>.
- [12] Emerson E, Felce D, Stancliffe RJ. Issues concerning self-report data and population-based data sets involving people with intellectual disabilities. *Intellect Dev Disabil.* 2013;51(5):333–348. doi: [10.1352/1934-9556-51.5.333](https://doi.org/10.1352/1934-9556-51.5.333).
- [13] Huus K, Granlund M, Bornman J, et al. Human rights of children with intellectual disabilities: comparing self-ratings and proxy ratings. *Child Care Health Dev.* 2015;41(6):1010–1017. doi: [10.1111/cch.12244](https://doi.org/10.1111/cch.12244).
- [14] Schalock RL, Brown I, Brown R, et al. Conceptualization, measurement, and application of quality of life for persons with intellectual disabilities: report of an international panel of experts. *Mental Retardation.* 2002;40(6):457–470. doi: [10.1352/0047-6765\(2002\)040<0457:CMAAQ>2.0.CO;2](https://doi.org/10.1352/0047-6765(2002)040<0457:CMAAQ>2.0.CO;2).
- [15] Beukelman DR, Light JC. *Augmentative & alternative communication: supporting children and adults with complex communication needs.* 5th ed. Baltimore (Md): Paul H. Brookes Publishing Company, Incorporated; 2020.
- [16] Ivarsson M, Andersson AK, Almqvist L. Self-rating via video communication in children with disability – a feasibility study. *Front Psychol.* 2023;14:1130675. doi: [10.3389/fpsyg.2023.1130675](https://doi.org/10.3389/fpsyg.2023.1130675).
- [17] Durkin M, Wang W, Shrout PE, et al. Evaluating a ten questions screen for childhood disability: reliability and internal structure in different cultures. *J Clin Epidemiol.* 1995;48(5):657–666. doi: [10.1016/0895-4356\(94\)00163-K](https://doi.org/10.1016/0895-4356(94)00163-K).

- [18] Durkin M, Zaman S, Thorburn M, et al. Population-based studies of childhood disability in developing countries: rationale and study design. *Int J Ment Health*. 1991;20(2): 47–60. doi: [10.1080/00207411.1991.11449195](https://doi.org/10.1080/00207411.1991.11449195).
- [19] Arvidsson P, Dada S, Granlund M, et al. Content validity and usefulness of Picture My Participation for measuring participation in children with and without intellectual disability in South Africa and Sweden. *Scand J Occup Ther*. 2020;27(5):336–348. doi: [10.1080/11038128.2019.1645878](https://doi.org/10.1080/11038128.2019.1645878).
- [20] Cameron L, Murphy J. Enabling young people with a learning disability to make choices at a time of transition. *Brit J Learn Disabil*. 2002;30(3):105–112. doi: [10.1046/j.1468-3156.2002.00165.x](https://doi.org/10.1046/j.1468-3156.2002.00165.x).
- [21] Arvidsson P, Dada S, Granlund M, et al. Structural validity and internal consistency of Picture My Participation: a measure for children with disability. *African Journal of Disability*. 2021;10. doi: [10.4102/ajod.v10i0.763](https://doi.org/10.4102/ajod.v10i0.763).
- [22] Goodman R. The strengths and difficulties questionnaire: a research note. *J Child Psychol Psychiatry*. 1997;38(5):581–586. doi: [10.1111/j.1469-7610.1997.tb01545.x](https://doi.org/10.1111/j.1469-7610.1997.tb01545.x).
- [23] Emerson E. Use of the strengths and difficulties questionnaire to assess the mental health needs of children and adolescents with intellectual disabilities. *J Intellect Dev Disabil*. 2005;30(1):14–23. doi: [10.1080/13668250500033169](https://doi.org/10.1080/13668250500033169).
- [24] Haynes A, Gilmore L, Shochet I, et al. Factor analysis of the self-report version of the strengths and difficulties questionnaire in a sample of children with intellectual disability. *Res Dev Disabil*. 2013;34(2):847–854. doi: [10.1016/j.ridd.2012.11.008](https://doi.org/10.1016/j.ridd.2012.11.008).
- [25] Wille N, Badia X, Bonsel G, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual Life Res*. 2010;19(6):875–886. doi: [10.1007/s11136-010-9648-y](https://doi.org/10.1007/s11136-010-9648-y).
- [26] Burström K, Bartonek Å, Broström E, et al. EQ-5D-Y as a health-related quality of life measure in children and adolescents with functional disability in Sweden: testing feasibility and validity. *Acta Paediatr*. 2014;103(4):426–435. doi: [10.1111/apa.12557](https://doi.org/10.1111/apa.12557).
- [27] Roberts C, Currie C, Samdal O, et al. Measuring the health and health behaviours of adolescents through cross-national survey research: recent developments in the Health Behaviour in School-aged Children (HBSC) study. *J Public Health*. 2007;15(3):179–186. doi: [10.1007/s10389-007-0100-x](https://doi.org/10.1007/s10389-007-0100-x).
- [28] Steele CA, Kalnins IV, Rossen BE, et al. Age-related health risk behaviors of adolescents with physical disabilities. *Soz Praventivmed*. 2004;49(2):132–141. doi: [10.1007/s00038-004-3056-4](https://doi.org/10.1007/s00038-004-3056-4).
- [29] Cantril H. *The pattern of human concern*. New Brunswick, NJ: Rutgers University Press; 1965.
- [30] Levin KA, Currie C. Reliability and validity of an adapted version of the cantril ladder for use with adolescent samples. *Soc Indic Res*. 2014;119(2):1047–1063. doi: [10.1007/s11205-013-0507-4](https://doi.org/10.1007/s11205-013-0507-4).
- [31] van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Soft*. 2011;45(3):1–67. doi: [10.18637/jss.v045.i03](https://doi.org/10.18637/jss.v045.i03).
- [32] Revelle W. *Psych: procedures for psychological, psychometric, and personality research [Manual]*. Evanston, IL: Northwestern University; 2022.
- [33] R Core Team. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing; 2022.
- [34] Meyer D, Zeileis A, Hornik K, et al. *vcd: visualizing categorical data; 2023. [Manual]*. <https://CRAN.R-project.org/package=vcd>.
- [35] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174. doi: [10.2307/2529310](https://doi.org/10.2307/2529310).