Public Health Nutrition: 13(2), 187–195 doi:10.1017/S1368980009990759

Relative validity of a self-completion 24 h recall questionnaire to assess beverage consumption among schoolchildren aged 7 to 9 years

Rebecca Muckelbauer*, Lars Libuda and Mathilde Kersting
Research Institute of Child Nutrition (FKE) Dortmund, Rheinische Friedrich-Wilhems-University of Bonn, Heinstueck 11, D-44225 Dortmund, Germany

Submitted 3 December 2008: Accepted 8 June 2009: First published online 4 August 2009

Abstract

Objective: Drinking habits in children are associated with diet quality, but validated assessment tools for large-scale studies in young children are lacking. Therefore, we validated a self-completion 24 h recall questionnaire (RQ) focusing on beverage consumption with a 24 h weighed record (WR).

Design: Thirty-five voluntary participants from the DONALD (Dortmund Nutritional and Anthropometric Longitudinally Designed) Study cohort aged 7–9 years completed the RQ. The illustrated RQ required ticking the number of glasses of seven beverage categories consumed in five time intervals in the previous 24 h. As a reference, parents completed weighed records of their child's diet. Agreement between the RQ and WR was tested by classification into consumers and non-consumers (kappa coefficients, κ), by the children's ability to estimate the exact beverage and total volume consumed (Wilcoxon signed-rank test, Spearman rank correlation), and by ranking children according to reported beverage volumes.

Results: The RQ and WR showed a good level of agreement for classifying participants into consumers and non-consumers of the single beverage categories (κ values between 0·78 and 0·94). Correlation coefficients for the volume of the single categories ranged between 0·81 and 0·91. The total beverage volume was overestimated in the RQ, on average, by 114 ml ($P = 0{\cdot}015$). Agreement in ranking into tertiles by beverage volume was moderate to good for juice/soft drinks ($\kappa = 0{\cdot}44$), milk ($\kappa = 0{\cdot}57$) and water ($\kappa = 0{\cdot}70$), but fair for the total beverage volume ($\kappa = 0{\cdot}23$).

Conclusions: Our self-completion 24 h RQ could estimate the consumption of several beverage categories among young children at the group level, but quantification of total beverage volume was flawed.

Keywords
Validation study
Self-report 24 h recall questionnaire
Dietary weighed record
Dietary assessment in children
Beverage consumption

Beverages are the main fluid source for meeting water requirements and may also contribute considerably to energy and nutrient intakes in children[1–4]. Therefore, beverage consumption influences total diet quality[4,5] and possibly childhood obesity, which is linked to the consumption of sugar-containing beverages[6,7]. Recently, intervention studies that focused on beverage consumption achieved beneficial effects on the body weight status of children and adolescents by promoting water consumption[8] or discouraging the consumption of soft drinks[9,10].

To investigate the long-term consequences of drinking habits in observational studies or to test for intervention effects, valid dietary measures are essential. However, the assessment of dietary intake in children is challenging owing to their limited cognitive abilities[11]. Especially for large-scale trials carried out in schools it is essential that the data collection methods are relatively quick, cost-effective, easy to implement and appropriate for the targeted age group. Methods without parental involvement may decrease participation barriers and the risk for selection bias in the sample. Assessment tools such as the observation of meals, weighed food records or food diaries are cost-intensive, have a high respondent burden and require parental involvement or trained interviewers[12], and are therefore more suited for small-scale trials. For large-scale trials self-report questionnaires seem to be more feasible. Among these, the FFQ is a popular method[12], although young children tend to lack the cognitive skills to recall and quantify their usual dietary intake over a long time period[11]. Moreover, FFQ seem to estimate children's diet less accurately than 24 h recalls[12], which can be applied by an interviewer, via computer or as a self-completion questionnaire.

*Corresponding author: Email muckelbauer@fke-do.de

Many studies have evaluated beverage consumption in children and adolescents but the number of validated assessment tools is limited[12]. To our knowledge, the validity of exclusively self-completion tools with a focus on beverage consumption has not been previously investigated in elementary-school children.

For a large-scale intervention trial focused on the prevention of overweight by promoting water consumption[8], we developed a self-completion semi-quantitative questionnaire based on the concept of a 24 h recall to assess changes in the beverage consumption of elementary-school children in the classroom setting. The objective of the present study was to test the validity of this 24 h recall questionnaire (RQ) in children aged 7 to 9 years using a parent-completed 24 h weighed diet record (WR) as the reference method.

Methods

Subjects

A subsample from the Dortmund Nutritional and Anthropometric Longitudinally Designed (DONALD) Study cohort was defined for participation. The DONALD Study is an ongoing longitudinal (open cohort) study established in 1985 at the Research Institute of Child Nutrition in Dortmund, Germany. It collects information on the nutrition, development, metabolism and health status of subjects between infancy and early adulthood. The regular assessments begin at 3 months of age, take place annually from the age of 2 years onwards, and include 3 d weighed dietary records, anthropometrics, urine sampling, interviews on lifestyle and medical assessments. Further details of the DONALD Study are provided elsewhere[13].

In the present validation study, participants aged 7 to 9 years were enrolled from the DONALD Study cohort between September 2006 and September 2008. Children were invited for participation in the validation study at their visit for the annual assessment. Each child could participate once only.

The DONALD Study was approved by the Ethics Committee of the Rheinische Friedrich-Wilhelms-University, Bonn, Germany. Parents gave written informed consent for their child's participation.

24 h Recall questionnaire

The RQ asked for the number of glasses of seven beverage categories consumed at five time intervals over the previous 24 h. These intervals were named: (i) this morning for breakfast at home; (ii) this morning at school; (iii) yesterday at supper and afterwards; (iv) yesterday between lunch and supper; and (v) yesterday at lunchtime. Each time interval was dealt with on a single page of the RQ. The front page of the RQ described how to complete the questionnaire by ticking the glasses illustrated according to the number of glasses of each beverage category consumed. Children could choose between full glasses, half-full glasses and an empty glass. The beverage categories included: (i) tap water; (ii) tea; (iii) mineral water; (iv) milk (including milk drinks); (v) soft drinks (liquid or powdered, carbonated or non-carbonated, e.g. regular and diet soft drinks, iced tea, energy drinks, sport drinks); (vi) juices (fruit and vegetable juices, fruit drinks, juice mixed with sparkling water); and (vii) other beverages (e.g. coffee, drinks the child could not categorize). The identification of the appropriate category was facilitated by illustrations. Figure 1 exemplarily illustrates a page of the RQ.

24 h Weighed record – reference method

The dietary assessments of the DONALD Study are carried out using 3 d weighed dietary records on three consecutive days. Solely or predominantly the participants' parents weighed and recorded foods and fluids consumed, as well as leftovers, using electronic food scales, to the nearest 1 g. Semi-quantitative recording was allowed (e.g. number of glasses), but in 97 % of the WR analysed in the present study more than 90 % of the food items were weighed. For validation, only dietary data from the 3 d recording period that corresponded to the 24 h period assessed in the RQ were included.

Data collection and coding

Children and their parents who agreed to participate received the forms for both assessment tools, the RQ and WR. The forms were collected by study personnel of the DONALD Study at a visit the family's home. Children had to complete the RQ on the second or third day of the annual 3 d parent-completed weighed record just before lunchtime and preferably not on the weekend. Parents received a short letter with instructions on how their child should complete the RQ including definitions of the beverage categories. They were advised to explain the questionnaire to their child but not to interfere with their dietary recall.

The volume of consumed beverages reported in the RQ was converted from the glass unit into millilitres by using a predefined factor of 200 ml per glass. All beverages reported in the WR were coded by beverage category and time interval to match those on the RQ. The total 24 h beverage volume was calculated by summing up the seven beverage categories. Only fourteen (40 %) out of thirty-five children ticked the empty glass to indicate that a beverage category was not consumed. Therefore, non-consumption was assumed even if it was not explicitly marked.

Statistical analyses

The primary outcomes of the present validation study for measuring the agreement between the RQ and the WR as the reference method were: (i) the ability of the RQ to classify individuals into consumers and non-consumers by beverage category; (ii) the ability of the RQ to estimate

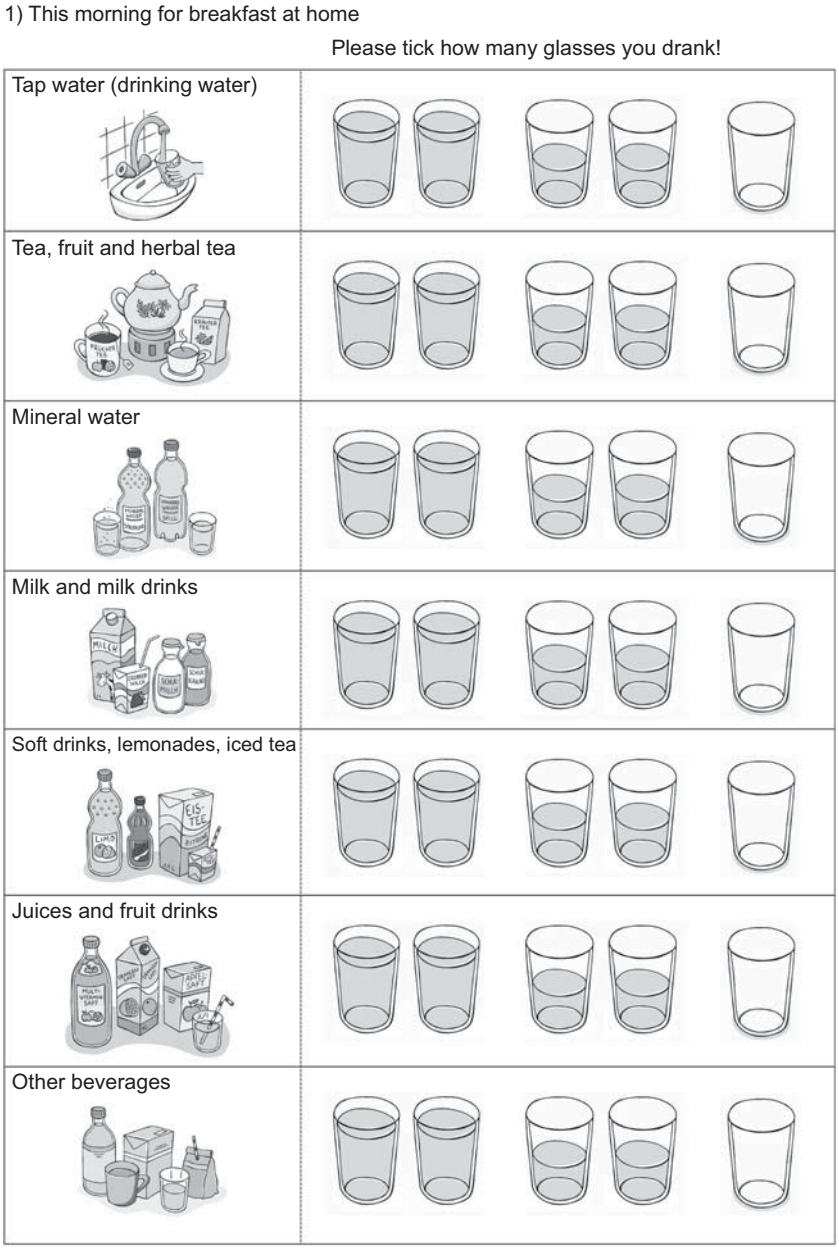

Fig. 1 An extract of the 24 h recall questionnaire translated from German showing the page that asked for beverage consumption in the first of the five time intervals, 'this morning for breakfast at home'

the exact beverage volume; and (iii) the ability of the RQ to rank individuals according to beverage volume.

For the first outcome we classified individuals into consumers and non-consumers over the total 24 h period for each of the seven beverage categories irrespective of the volumes reported in the RQ and WR. The agreement of reported consumption *v.* non-consumption of each beverage category between the two methods was assessed by designating individuals as matches (consumption or non-consumption reported on both the RQ and WR), omissions (consumption reported in the WR but not in the RQ) or intrusions (consumption reported in the RQ but not in the WR), and by calculating the kappa coefficient (κ).

For assessing the second outcome, the volume (in ml) of each beverage category and the total 24 h beverage volume were considered. To test for systematic differences between the volumes reported in the RQ and WR, the Wilcoxon signed-rank test was used. To assess the association between the volumes reported in the two methods, the Spearman rank correlation coefficient (r) was used. To reveal the quantitative agreement between the RQ and the WR across the range of total 24 h beverage volume, the Bland–Altman plot was used[14]. The difference between the RQ and WR was plotted against the average of the two methods. The mean difference indicated the bias of the RQ compared with the WR. The limits

https://doi.org/10.1017/S1368980009990759 Published online by Cambridge University Press

of agreement (LOA) were defined by the mean difference plus or minus two standard deviations. The association between the difference and the average of the two methods was tested by using Spearman rank correlation.

For assessing the third outcome, individuals, consumers and non-consumers, were categorised into tertiles according to their total 24 h beverage volume as reported in the WR and RQ and according to their beverage volume of those beverage categories with less than a third of non-consumers. The categories tap water and mineral water were merged into one summary category because rates of non-consumers of the single categories were higher than one-third. For the same reason the categories juices and soft drinks were also merged. We calculated the percentage of children classified into the same tertile and those into the opposite tertile. Kappa coefficients were also provided.

A secondary outcome of the validation study was to assess the ability of the RQ to differentiate between confusable beverage categories. Therefore, we evaluated whether the children correctly differentiated between tap water and mineral water, and between soft drinks and juice. For this analysis, we included all the matched cases of the consumption of water (mineral or tap water) and juices/soft drinks (juices or soft drinks), respectively, reported in both the RQ and WR. These matched cases of consumption were assessed for each of the five time intervals and summed up for the total 24 h period. Each case was categorised into 'match' if the beverage categories were reported correctly in the RQ compared with the WR as the reference, or into 'misclassification' if not. Percentages of matches and misclassifications were calculated for each category.

Kappa coefficients were interpreted using the guidelines provided by Altman[15]: $\kappa < 0{\cdot}20$, poor; $\kappa = 0{\cdot}21$–$0{\cdot}40$, fair; $\kappa = 0{\cdot}41$–$0{\cdot}60$, moderate; $\kappa = 0{\cdot}61$–$0{\cdot}80$, good; $\kappa = 0{\cdot}81$–$1{\cdot}00$, very good. All analyses were performed using the SAS statistical software package version 9·1·3 (SAS Institute, Cary, NC, USA). $P < 0{\cdot}05$ was considered statistically significant.

Results

Participants

Out of 114 children from the DONALD Study cohort with available WR who were invited to participate, forty-two (37 %) agreed and returned the RQ. Six participants were excluded because the RQ was completed on a day without a corresponding WR and one was excluded because it was completed by a parent. Consequently, thirty-five (83 %) out of forty-two pairs of WR and RQ were available for validation. Participants included in the analysis did not differ from invited but not included children with respect to gender (χ^2 test, $P = 0{\cdot}25$) and birth date (t test, $P = 0{\cdot}74$). The participants (fifteen boys, twenty girls) had a mean age of 8·0 (SD 0·8) years. Parents reported completion times for the RQ of between 5 and 15 min.

Recall by beverage category

Table 1 shows the agreement between the RQ and WR by matches, omissions and intrusions of beverage consumption *v.* non-consumption for each beverage category. The majority of children were correctly classified into consumers or non-consumers by the RQ as match rates ranged from 91 % to 97 %. Values of κ between 0·78 and 0·94 argue for a good to very good agreement between the RQ and WR. Omission rates ranging between 0 % and 3 % indicate low reporting of phantom foods. The highest intrusion rate was 6 % and observed for tap water.

Estimation of beverage volume

Volumes of beverage consumption reported in the RQ and WR are presented in Table 2. The median total 24 h beverage volume was higher in the RQ than in the WR ($P = 0{\cdot}015$). The reported volume of each of the single beverage categories did not differ between the RQ and WR. Spearman rank correlation coefficients between the RQ and WR ranged from $r = 0{\cdot}86$ to $r = 0{\cdot}91$ for the single

Table 1 Agreement between the 24 h recall questionnaire (RQ) and 24 h weighed record (WR, reference) in terms of classification into consumers and non-consumers by beverage category: German schoolchildren aged 7 to 9 years (*n* 35)

		Matches†		Omissions‡		Intrusions§			
	No. of consumers*	*n*	%	*n*	%	*n*	%	κ	95 % CI
Mineral water	21	33	94	1	3	1	3	0·88	0·72, 1·00
Tap water	9	32	91	1	3	2	6	0·78	0·55, 1·00
Milk	22	33	94	1	3	1	3	0·88	0·71, 1·00
Juices	23	34	97	0	0	1	3	0·94	0·81, 1·00
Soft drinks	14	33	94	1	3	1	3	0·88	0·72, 1·00
Tea	5	34	97	1	3	0	0	0·87	0·63, 1·00
Other	0	34	97	0	0	1	3		NA

NA, statistics not applicable due to missing consumption.
*Consumption reported in the WR.
†Consumption reported in both the RQ and WR, or non-consumption reported in both.
‡Consumption reported in the WR but not in the RQ.
§Consumption reported in the RQ but not in the WR.

https://doi.org/10.1017/S1368980009990759 Published online by Cambridge University Press

Table 2 Difference and correlation between beverage volumes reported in the 24 h recall questionnaire (RQ) and 24 h weighed record (WR, reference) for individual beverage categories and total 24 h beverage consumption: German schoolchildren aged 7 to 9 years (*n* 35)

	Beverage volume (ml)*						
	RQ§		WR				
	Median or mean	IQR	Median or mean	IQR	Mean difference (RQ−WR)	*P*†	*r*‡
Total 24 h volume	1200	1100, 1400	1125	875, 1277		0·015	0·72
	1289		1175		114		
Mineral water	300	0, 700	300	0, 658		0·906	0·90
	403		412		−9		
Tap water	0	0, 200	0	0, 124		0·383	0·81
	120		88		32		
Milk	200	0, 200	115	0, 250		0·341	0·89
	169		146		23		
Juices	300	0, 600	308	0, 550		0·525	0·91
	374		353		21		
Soft drinks	0	0, 400	0	0, 260		0·105	0·86
	180		132		48		
Tea	0	0, 0	0	0, 0		0·188	0·87
	26		43		−17		
Other	0	0, 0	0	0, 0		NA	NA
	17		0		17		

NA, statistics not applicable due to missing consumption; IQR, interquartile range.
*Values are medians and IQR (25th, 75th percentile) or means of all participants (*n* 35).
†For differences between RQ and WR obtained by the Wilcoxon signed-rank test.
‡Spearman rank correlation coefficient *r* with $P < 0{\cdot}0001$.
§Volume conversion: 200 ml = 1 glass.

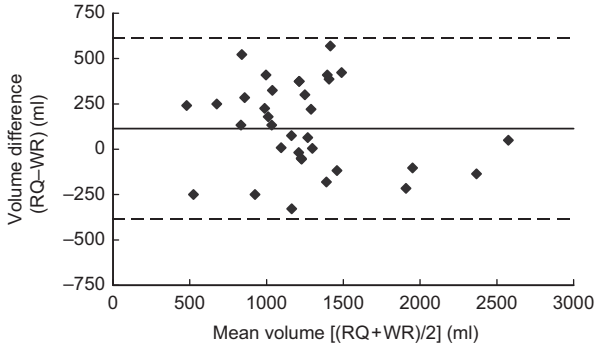

Fig. 2 Bland–Altman plot for analysis of agreement between total 24 h beverage volume reported in the 24 h recall questionnaire (RQ) and in the 24 h weighed record (WR) for German schoolchildren aged 7 to 9 years: – – –, upper and lower limits of agreement (mean ± 2 SD); ——, bias

beverage categories, whereas correlation was worse for the total 24 h beverage volume, $r = 0{\cdot}72$ (Table 2).

The Bland–Altman plot (Fig. 2) showed a mean difference between the two methods (RQ−WR) of 114 (SD 249) ml for the total 24 h beverage volume indicating that, on average, the RQ overestimated total 24 h beverage volume compared with the WR. The upper and lower LOA indicate that the RQ could estimate the total 24 h beverage volume within a range of 612 ml above to 385 ml below the volume measured in the WR. The individual differences between the two methods were not significantly associated with the average of the volumes measured by the two methods ($P = 0{\cdot}716$), which indicates that the variability and direction of the difference did not depend on the consumption level. However, the sample was quite small regarding the advice about using the Bland–Altman procedure.

Table 3 Agreement and kappa statistics for tertiles by beverage volume reported in the 24 h recall questionnaire (RQ) and 24 h weighed record (WR, reference) for selected beverage categories and total 24 h beverage consumption: German schoolchildren aged 7 to 9 years (*n* 35)

	Classification into					
	Correct tertile		Opposite tertile			
	n	%	*n*	%	κ	95 % CI
Total 24 h volume	17	49	1	3	0·23	−0·01, 0·47
Mineral/tap water	28	80	1	3	0·70	0·50, 0·90
Milk	25	71	0	0	0·57	0·36, 0·79
Juices/soft drinks	22	63	1	3	0·44	0·21, 0·68

Ranking by beverage volume

Table 3 shows the classification of the participants into tertiles according to beverage volume reported in the RQ and WR. Agreement for total 24 h beverage volume was fair as indicated by $\kappa = 0{\cdot}23$, and 49 % of the participants were classified into the correct tertile by the RQ. Based on κ values for the individual beverage categories agreement was moderate for milk and for the category juices/soft drinks, whereas it was good for the category mineral/tap water. A small percentage of participants (0–3 %) were grossly misclassified into the opposite tertile for the total 24 h consumption and for the different beverage categories.

https://doi.org/10.1017/S1368980009990759 Published online by Cambridge University Press

Table 4 Differentiation between tap water and mineral water, and between soft drinks and juice, in the 24 h recall questionnaire (RQ) compared with the 24 h weighed record (WR, reference) by the percentage of matched and misclassified beverages within the five time intervals: German schoolchildren aged 7 to 9 years (*n* 35)

		Matches†		Misclassifications‡	
	No. of occurrences*	*n*	%	*n*	%
Mineral/tap water					
Mineral water	51	50	98	1	2
Tap water	9	9	100	0	0
Juices/soft drinks					
Juices	41	39	95	2	5
Soft drinks	15	13	87	2	13

*Cases of matched or misclassified beverage category in the five time intervals.
†Beverage category correctly classified in the RQ as in the WR as reference.
‡Beverage category misclassified as the contrary category in the RQ as in the WR.

Differentiation of beverage categories

Based on the classification into matches and misclassifications in the RQ compared with the WR, tap water was not misclassified at all and mineral water was misclassified as tap water in one of fifty-one cases, indicating that the children were widely able to differentiate between tap and mineral water (Table 4). Regarding the differentiation between juices and soft drinks, children misclassified soft drinks as juices in 13 % of all cases and juices as soft drinks in 5 % of all cases.

Discussion

In the present study we validated a self-completion 24 h RQ to assess beverage consumption among elementary-school children at the group level with a 24 h WR as the reference. The results indicate that the children could recall well the beverage categories consumed. This semi-quantitative questionnaire also provided valid estimations of the volume consumed in the single beverage categories; however, estimation of the total 24 h beverage volume was unsatisfactory.

To our knowledge several self-completion dietary assessment tools for children aged 9 years or younger have been validated[16–26], but none of these solely targeted beverage consumption. Furthermore, for validation studies there is no established statistical standard for measuring validity[27], various methods have been used as a reference, and different main outcomes have been defined. As a result it is difficult to compare the validity of different dietary assessment tools.

Recall by beverage category

In our study the ability of the RQ to differentiate between consumers and non-consumers of the single beverage categories was good to very good as indicated by the match rates and kappa statistics. Omission rates were found to be very low. Other validation studies have shown that, in 24 h recall interviews, children are least likely to omit beverages compared with other foods consumed during school meals[28,29]. The validation of computerised 24 h recalls completed by children aged 11 to 14 years resulted in match, omission and intrusion rates similar to ours for several beverages[30]. However, the reference method in that study was a self-report food record that might have resulted in reporting errors in the same direction. Match rates slightly smaller than ours were found in a comparison between a 24 h questionnaire completed by children aged 9 to 11 years for several beverages and a 24 h recall interview[23].

In our study the intrusion rates of the different beverage categories (phantom food) were similar to the omission rates. This was also observed in validation studies with schoolchildren for a 24 h recall interview[28], but not for a recall questionnaire[23]. However, it has to be mentioned that omission and intrusion rates were differently defined and calculated in these studies.

Beverage volume

The ability of the RQ to rank individuals by beverage volume and to estimate the volume of single beverage categories was good to moderate, and no systematic over- or underestimation was observed for the single beverage categories. Similar to these results, a good estimation of beverage volume was observed in a 24 h recall interview among schoolchildren aged 8 to 10 years in which juices and milk were among the food groups of best estimated quantities[31]. Two self-report questionnaires targeting fruit and vegetable consumption[32,33] and a computerised 24 h recall[30] showed slightly lower validities of estimated volumes of various beverages in young adolescents compared with our beverage-targeted RQ.

The ability of our RQ to rank individuals according to their total 24 h beverage volume was only fair and the RQ overestimated the total volume systematically. However, the bias of 114 ml, as shown in the Bland–Altman plot, was small and independent of consumption level indicating good validity at the group level. In contrast, individual differences between the two methods were quite high as shown by the large LOA. In conclusion, the RQ was limited in its precision to quantify total 24 h beverage volume and may therefore also fail to detect small changes in total beverage intake in an intervention study.

A possible cause for the flaw in measuring the quantity may be the conversion factor of 200 ml for one glass that was applied as the assessment unit of the RQ. Haraldsdóttir *et al.* used the same conversion factor and found an overestimation of juice consumption by children aged 11 to 12 years[33]. For more accurate quantification, conversion of portion sizes could be sex-, age- and country-specific as was applied in the validation study of Cade *et al.*[22].

The quantity category 'empty glass' was included in the RQ to ensure that children consider each beverage category. This did not work because the majority of the children did not tick the category 'empty glass' correctly, although the rest of the questionnaire was completed properly.

The present study showed that children were able to differentiate between juices and soft drinks, but this may depend on country-specific drinking habits and food knowledge. In other countries the beverage categories could be defined differently, e.g. include both a regular and a diet soft drinks category if the difference is generally understood by the target groups. To address this research question, the questionnaire should be validated in other countries.

Advantages

An important strength of the current validation study is the parent-completed weighed 24 h food record as the reference method. The weighed record is one of the most accurate non-invasive methods of dietary assessment[34]. For assessing validity among children direct observation has been often used as a reference method[12], but this method of data collection is limited to short observation periods such as school breaks or meals. Estimated food records or recall techniques are also commonly used as a reference in validation studies[12]. Our reference method was carried out by the parents and not by the participants themselves, which might have reduced the risk of reporting errors of the same direction.

Our RQ is well suited for assessing beverage consumption among large samples of children owing to its self-completion design. If applied in the school setting it is independent from the involvement or assistance of parents, interviewers or computers that might increase participation barriers and costs. The illustrated questionnaire requires low writing and reading skills, and therefore it can be used in young schoolchildren and in immigrant children who have a different first language.

When completing a questionnaire, children may choose a particular response because of social desirability[35], especially if they are aware of the beverage of interest or the research aim. However, we could not discern any specific misreporting in the single beverage categories. Since the RQ included questions on many kinds of beverages, participants cannot deduce which beverage category is of research interest.

The illustrations of the listed beverage categories were meant to encourage the children to identify and to recognise consumed beverages because food listings in a questionnaire are supposed to improve the recall of a consumed food by prompting the memory[35]. This kind of support in a food recall may also lead to recognition errors, i.e. reporting of foods that were actually not consumed by the children, but in the present study intrusion rates of the 24 h RQ were found to be low.

Limitations

The study has several limitations. First, the RQ was designed for use at the group level in the school setting, but validity was assessed on an individual level at home under the supervision of parents instead of teachers. Furthermore, parental support in the completion of the questionnaire cannot be excluded completely although parents were repeatedly and personally instructed not to do so. Children needed 5 to 15 min to complete the questionnaire as estimated by their parents. Completion of the RQ by children aged 7 to 9 years in the classroom in the context of a large school-based intervention study[8] took between 20 and 40 min as reported by the supervising teachers. This discrepancy might indicate that completion of the questionnaire by the children could depend on the setting in which it is administered. Second, parental report of their child's dietary intake was the reference method. Parents may not be completely aware of children's snacking and out-of-home consumption and thus depend on their child's report. In addition, asking the parents to record and weigh the food may have enhanced the children's attention to the food eaten, possibly leading to improved recall accuracy or to alterations in the diet[36]. However, children and parents of the DONALD Study cohort are very used to the annual dietary recording. Third, the study was prone to selection bias as the participation rate was low, and sample size was small with thirty-five participants. As the present study was an addition to the quite extensive annual DONALD Study assessments and as there is an interest in keeping participants in this longitudinal study, no further effort was made to increase participation in the validation study. In addition, participants are derived from a cohort with a higher socio-economic status than the average in Germany[13], which might limit the transferability of our results. Finally, the outcome that can be assessed by the RQ is limited to the quantity of the main beverage categories but the exact nutrient or energy intake from beverages cannot be measured by this questionnaire.

Based on the results of our validation study we suggest several modifications for improving practicability and validity of the 24 h RQ.

General modifications

1. The quantity category 'empty glass' could be omitted as most children did not tick it in the case of non-consumption of a beverage.

https://doi.org/10.1017/S1368980009990759 Published online by Cambridge University Press

2. The category 'other beverages' may invite children to misclassify beverages into this category and could be replaced by beverages of specific research interest such as probiotic drinks.

Population-adapted modifications

1. The volume of one glass should be adapted to the sex- and age-specific portion size of the target population.
2. The definition of beverage categories may be adapted to country-specific consumption habits.
3. In children older than 9 years of age, beverage consumption is expected to increase and thus the questionnaire should be adapted by adding one more full glass to mark, resulting in a maximum volume per time interval and per beverage category of four glasses.

Conclusion

Our 24 h RQ was able to estimate the consumption of different beverage categories among schoolchildren at the group level. Estimation of total 24 h beverage volume by the questionnaire was less accurate. The self-completion questionnaire is applicable even in young children of elementary-school age as its illustration-based design assumes low reading and writing skills. Whether this 24 h RQ can serve as a practical tool for the evaluation of children's drinking habits in large-scale studies also in the school setting should be affirmed by further validation studies.

Acknowledgements

Sources of funding: The study was supported by grant 05HS026 of the German Federal Ministry of Food, Agriculture, and Consumer Protection with R.M. and L.L. receiving research funding from this grant. *Conflict of interest declaration:* None declared. *Authorship responsibilities:* All authors designed the study. R.M. and L.L. were responsible for data collection. R.M. conducted the analysis and drafted the manuscript. All authors made substantial contributions and revised the manuscript. *Acknowledgments:* We thank C. Chada, R. Schäfer, B. Holtermann, S. Twenhöfen and U. Kahrweg for their support in data collection.

References

1. Nielsen SJ & Popkin BM (2004) Changes in beverage intake between 1977 and 2001. *Am J Prev Med* **27**, 205–210.
2. Storey ML, Forshee RA & Anderson PA (2006) Beverage consumption in the US population. *J Am Diet Assoc* **106**, 1992–2000.
3. Subar AF, Krebs-Smith SM, Cook A & Kahle LL (1998) Dietary sources of nutrients among US children, 1989–1991. *Pediatrics* **102**, 913–923.
4. Marshall TA, Eichenberger Gilmore JM, Broffitt B, Stumbo PJ & Levy SM (2005) Diet quality in young children is influenced by beverage consumption. *J Am Coll Nutr* **24**, 65–75.
5. Libuda L, Alexy U, Buyken AE, Sichert-Hellert W, Stehle P & Kersting M (2009) Consumption of sugar-sweetened beverages and its association with nutrient intakes and diet quality in German children and adolescents. *Br J Nutr* **101**, 1549–1557.
6. Vartanian LR, Schwartz MB & Brownell KD (2007) Effects of soft drink consumption on nutrition and health: a systematic review and meta-analysis. *Am J Public Health* **97**, 667–675.
7. Malik VS, Schulze MB & Hu FB (2006) Intake of sugar-sweetened beverages and weight gain: a systematic review. *Am J Clin Nutr* **84**, 274–288.
8. Muckelbauer R, Libuda L, Clausen K, Reinehr T, Toschke AM & Kersting M (2008) Promotion and provision of drinking water in schools for overweight prevention: a randomized controlled cluster trial. *Pediatrics* **123**, e661–e667.
9. Ebbeling CB, Feldman HA, Osganian SK, Chomitz VR, Ellenbogen SJ & Ludwig DS (2006) Effects of decreasing sugar-sweetened beverage consumption on body weight in adolescents: a randomized, controlled pilot study. *Pediatrics* **117**, 673–680.
10. James J, Thomas P, Cavan D & Kerr D (2004) Preventing childhood obesity by reducing consumption of carbonated drinks: cluster randomised controlled trial. *BMJ* **328**, 1237.
11. Livingstone MB & Robson PJ (2000) Measurement of dietary intake in children. *Proc Nutr Soc* **59**, 279–293.
12. McPherson RS, Hoelscher DM, Alexander M, Scanlon KS & Serdula MK (2000) Dietary assessment methods among school-aged children: validity and reliability. *Prev Med* **31**, issue 2, S11–S33.
13. Kroke A, Manz F, Kersting M, Remer T, Sichert-Hellert W, Alexy U & Lentze MJ (2004) The DONALD Study. History, current status and future perspectives. *Eur J Nutr* **43**, 45–54.
14. Bland JM & Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310.
15. Altman DG (1999) *Practical Statistics for Medical Research*, 1st ed. Boca Raton, FL: Chapman & Hall.
16. Edmunds LD & Ziebland S (2002) Development and validation of the Day in the Life Questionnaire (DILQ) as a measure of fruit and vegetable questionnaire for 7–9 year olds. *Health Educ Res* **17**, 211–220.
17. Baranowski T, Islam N, Baranowski J, Cullen KW, Myres D, Marsh T & de MC (2002) The food intake recording software system is valid among fourth-grade children. *J Am Diet Assoc* **102**, 380–385.
18. Baranowski T, Smith M, Baranowski J, Wang DT, Doyle C, Lin LS, Hearn MD & Resnicow K (1997) Low validity of a seven-item fruit and vegetable food frequency questionnaire among third-grade students. *J Am Diet Assoc* **97**, 66–68.
19. Economos CD, Sacheck JM, Kwan Ho Chui K, Irizzary L, Guillemont J, Collins JJ & Hyatt RR (2008) School-based behavioral assessment tools are reliable and valid for measurement of fruit and vegetable intake, physical activity, and television viewing in young children. *J Am Diet Assoc* **108**, 695–701.
20. Domel SB, Baranowski T, Leonard SB, Davis H, Riley P & Baranowski J (1994) Accuracy of fourth- and fifth-grade students' food records compared with school-lunch observations. *Am J Clin Nutr* **59**, 1 Suppl., 218S–220S.
21. Domel SB, Baranowski T, Davis H, Leonard SB, Riley P & Baranowski J (1994) Fruit and vegetable food frequencies by fourth and fifth grade students: validity and reliability. *J Am Coll Nutr* **13**, 33–39.

22. Cade JE, Frear L & Greenwood DC (2006) Assessment of diet in young children with an emphasis on fruit and vegetable intake: using CADET–Child and Diet Evaluation Tool. *Public Health Nutr* **9**, 501–508.
23. Moore GF, Tapper K, Murphy S, Clark R, Lynch R & Moore L (2007) Validation of a self-completion measure of breakfast foods, snacks and fruits and vegetables consumed by 9- to 11-year-old schoolchildren. *Eur J Clin Nutr* **61**, 420–430.
24. Moore L, Tapper K, Dennehy A & Cooper A (2005) Development and testing of a computerised 24-h recall questionnaire measuring fruit and snack consumption among 9–11 year olds. *Eur J Clin Nutr* **59**, 809–816.
25. Lillegaard IT, Loken EB & Andersen LF (2007) Relative validation of a pre-coded food diary among children, under-reporting varies with reporting day and time of the day. *Eur J Clin Nutr* **61**, 61–68.
26. Lillegaard IT & Andersen LF (2005) Validation of a pre-coded food diary with energy expenditure, comparison of under-reporters *v.* acceptable reporters. *Br J Nutr* **94**, 998–1003.
27. Masson LF, McNeill G, Tomany JO, Simpson JA, Peace HS, Wei L, Grubb DA & Bolton-Smith C (2003) Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic. *Public Health Nutr* **6**, 313–321.
28. Baxter SD, Thompson WO, Litaker MS, Frye FH & Guinn CH (2002) Low accuracy and low consistency of fourth-graders' school breakfast and school lunch recalls. *J Am Diet Assoc* **102**, 386–395.
29. Baxter SD, Thompson WO, Davis HC & Johnson MH (1997) Impact of gender, ethnicity, meal component, and time interval between eating and reporting on accuracy of fourth-graders' self-reports of school lunch. *J Am Diet Assoc* **97**, 1293–1298.
30. Vereecken CA, Covents M, Matthys C & Maes L (2005) Young adolescents' nutrition assessment on computer (YANA-C). *Eur J Clin Nutr* **59**, 658–667.
31. Weber JL, Lytle L, Gittelsohn J, Cunningham-Sabo L, Heller K, Anliker JA, Stevens J, Hurley J & Ring K (2004) Validity of self-reported dietary intake at school meals by American Indian children: the Pathways Study. *J Am Diet Assoc* **104**, 746–752.
32. Field AE, Colditz GA, Fox MK, Byers T, Serdula M, Bosch RJ & Peterson KE (1998) Comparison of 4 questionnaires for assessment of fruit and vegetable intake. *Am J Public Health* **88**, 1216–1218.
33. Haraldsdottir J, Thorsdottir I, de Almeida MD, Maes L, Perez Rodrigo C, Elmadfa I & Frost Andersen L (2005) Validity and reproducibility of a precoded questionnaire to assess fruit and vegetable intake in European 11- to 12-year-old schoolchildren. *Ann Nutr Metab* **49**, 221–227.
34. Bingham SA, Cassidy A, Cole TJ *et al.* (1995) Validation of weighed records and other methods of dietary assessment using the 24 h urine nitrogen technique and other biological markers. *Br J Nutr* **73**, 531–550.
35. Baranowski T & Domel SB (1994) A cognitive model of children's reporting of food intake. *Am J Clin Nutr* **59**, 1 Suppl., 212S–217S.
36. Thompson FE & Byers T (1994) Dietary assessment resource manual. *J Nutr* **124**, 11 Suppl., 2245S–2317S.