CHAPTER 3: DEVELOPMENT AND DEVELOPMENTAL EVALUATION SUB-STUDIES OF RESEARCH TOOLS

Chapter overview

3.1 TEST METHOD

Keller et al [184] and Jones [185] recommend that the first step in the systematic development of a nutritional assessment and screening tool should be to spell out the aim(s) of the instrument. Only when that, which is to be determined by the screening tool, is very clear, can the validity thereof be measured. Furthermore, since the context can also affect validity, the correct usage (for example the relevant outcome variable[s], target group and administration detail) should be specified.

3.1.1 Overall development rationale, aims and correct application information

Based on the possible purposes and uses that Wilkin et al [182] identified for health-related measures, the current screening tool should have primarily discriminative properties, since it should separate children with a high fat intake from those eating according to the prudent dietary guidelines as specified by the NCEP.

Within clinical practice, the dietary fat screener should enable health professionals to correctly classify individuals according to their fat intake. Since the valid measurement of usual, absolute dietary intake of individuals appears to be an unrealistic aim, the correct ranking of groups of children is the immediate aim, primarily for use in nutritional epidemiology. Thus, usual dietary intake of PFE, PSFE and cholesterol should be reflected by the dietary screener, even though quantitative precision in terms of these variables was not required.

Since the screener was supposed to be a tool for level 1 [177] nutrition care, the current tool had to be administrable by qualified professionals (that is 'individuals who are qualified by virtue of their education, experience, competence, applicable professional licensure, regulations, or certification, registration, or privileges') who have a basic nutrition knowledge and who have knowledge of and access to referral systems for children who are identified as having a nutritional risk. Alternatively, it should be suited for self-completion. No parental assistance should be required for providing the information. The setting was school-based requiring no specialised apparatus. The target group was urban, middle-class South African school children in grade six.

The following practical usage and development criteria were set:

- South-African food-based so as to provide a food-based starting point for behaviour change counselling in line with current thinking within nutrition education theory
- Requiring minimal literacy from participant

- Attractive: Pictorial, colourful, novel and practical
- Interpretable and action-oriented
- Administrable in group setting
- Suitable for school settings
- Quick to use

Following the developmental evaluation and comparative validation (thus prior to implementation), the scientific merits of the screener should be known in quantitative terms.

3.1.2 Format

The basic structure chosen for the test method was that of a FFQ, since a FFQ can accommodate most of the stated aims and pre-set criteria, for example estimation of usual intake, focusing on specific dietary components with relatively low respondent burden. Consequently the general guidelines for developing such a tool were followed [26, 27, 115] and in particular the principles regarding cultural appropriateness.[116, 261]

In the absence of representative food intake data for the target group, the US's NCEP MEDFICTS Dietary Assessment Questionnaire was chosen as starting point for the development. The MEDFICTS tool is widely published in standard nutrition text books (for example references [10, 26]) as well as on the world wide web (www.bgsm.edu/nutrition/medfic.htm; www.nih.gov/news/stepbystep/medficts.htm; http://www.nhlbi.nih.gov/guidelines/cholesterol/atp3full.pdf).

Since food questionnaires should be adjusted for the population group for which they are intended,[116] a number of changes were made to the original MEDFICTS tool during the initial development for this study. In the following sections the item list and the various elements of the quantification are each described in terms of background considerations and the application in the current tool. In addition, the developmental evaluation sub-studies that were conducted in terms of the above are briefly presented. The Research and Ethics Committee of the University of Pretoria approved each of the developmental evaluation sub-studies individually.

3.1.2.1 Item list

Creating an item list is the first step in the development of a FFQ type dietary assessment tool. This refers to the foods or rows in the questionnaire.

3.1.2.1.1 Background (original tool)

In the original MEDFICTS tool, the individual food item list is replaced by eight food categories (Meats, Eggs, Dairy, Fried foods, In baked goods, Convenience foods, Table fats, Snacks), which are the prime contributors of dietary fat, saturated fat and cholesterol in the North American diet.[201] Kris-Etherton et al [203] created three sub-categories (milk, cheese and dessert) for the dairy category. Furthermore, each category has a group 1 (high fat choices) and a group 2 (low fat choices).

3.1.2.1.2 Application (current tool)

A consensus workshop consisting of three private practicing dietitians (with extensive experience in dealing with the target group) and the researcher reviewed the original item list for face and content validity. It was decided to retain the main MEDFICTS categories and sub-categories. However, in the meats, eggs, fried foods, in baked goods, and convenience foods categories group 2 was excluded, since it was reasoned that in the South African context this would cause practical problems because of limited availability, access and target group awareness of these foods. Furthermore, including this group did not affect the score. The role of the group 2 foods in the original tool was only for clarification purposes.[203] Thus a total of ten food categories formed the basis of the scored part of the current tool with five clarifying food categories (see Table 3.2 for a text summary of the tool).

Subsequently a developmental evaluation sub-study was conducted to obtain a more objective indication of the content and face validity.

3.1.2.1.3 Sub-study 1: Content and face validity

Rationale: The face validity of a method describes, according to Johnson et al,[262] the extent to which the questions asked conform to current expert opinion relating to what the instrument is intending to measure. It is based on intuitive judgment of experts, and is considered a necessary step in measuring the validity of any new dietary tool.

Aim: To identify the food items recommended by local dietitians to healthy black and white children between the ages of nine to twelve years in order to meet selected food-based dietary guidelines.

Participants: Anonymous questionnaires were handed out to 120 dietitians attending Association of Dietetics in Southern Africa meetings in the period June to August 2001 in the branches Mpumalanga, Pretoria and Gauteng South (South Africa).

Methods: Dietitians were requested to list the ten most important foods for children in the specified age group to increase and to decrease in order to achieve selected food-based dietary guidelines. During data analysis final year B Dietetics students, who were unaware of this sub-study's aim, grouped the listed foods. The groups that were formed were: '*fatty meats'* consisting of regular minced meat, steak, bacon, biltong [that is dried beef], dried beef sausage, boerewors [that is beef sausage], lamb chops and sausages; '*processed meat*' included russions, polony, viennas, salami and corned beef; '*organ meat'* included tripe, liver and kidney; full cream cheeses, full cream yogurt, ice cream, cream, full cream inkomazi [that is soured milk] and milkshakes formed the '*full cream dairy product'* group; '*take-aways'* referred to hamburgers, toasted sandwiches and pizzas; '*animal fat*' included lard, Holsum [that is hydrogenated fat] and butter.

Results and discussion: A total of 50 questionnaires were timeously returned by prepaid postage. Ten of these had to be excluded because they did not meet the inclusion criterion of coming from dietitians who consult the target group at least once a week, resulting in a final sample size of 40. In Table 3.1 the findings regarding the ten most frequently mentioned foods / groups of foods to eat less of in order to lower fat intake are summarised. Seventeen and 23 dietitians reported that their respective majority clients were white and black children. They thus based their recommendations on that group. (Results regarding foods to increase in order to lower fat intake, and recommendations in respect of the other guidelines are not reported).

TABLE 3.1: FOODS TO EAT LESS OF FOR LOWERING FAT INTAKE: NUMBER (PERCENTAGE) OF DIETITIANS (n=40) WHO MENTIONED EACH FOOD / FOOD GROUP

White children (n=17)		Black children (n=23)		All children (n=40)	
Rank and food	Frequency (%)	Rank and food	Frequency (%)	Rank and food	Frequency (%)
1 Chocolate	14 (82)	1 Full cream dairy products	20 (87)	1 Full cream dairy products	33 (83)
2 Full cream dairy products	13 (76)	2 Fried chips	19 (83)	2 Fried chips	31 (78)
3 Fatty meats	12 (71)	3 Processed meats	14 (61)	3 Fatty meats	24 (60)
4 Fried chips	12 (71)	4 Fatty meats	12 (52)	3 Chocolate	24 (60)
5 Biscuits	9 (53)	5 Chocolate	10 (43)	5 Processed meats	19 (48)
5 Take-aways	9 (53)	6 Margarine	9 (39)	6 Margarine	17 (43)
5 Margarine	9 (53)	7 Crisps	9 (39)	7 Crisps	16 (40)
8 Crisps	7 (41)	8 Fish	8 (35)	8 Biscuits	14 (35)
8 Animal fat	7 (41)	9 Organ meats	8 (35)	8 Take-aways	14 (35)
10 Chicken	6 (35)	10 Cakes	7 (30)	8 Animal fats	14 (35)
		10 Coffee / tea creamers	7 (30)		
		10 Chicken	7 (30)		
		10 Animal fat	7 (30)		

From the survey the following circumstantial evidence emerged:

- In general, the foods listed showed some similarity with the MEDFICTS food categories, even though the groupings differed. Eggs, regardless of racial group, were a clear exception in the sense that they did not feature at all in the top ten listings.
- There tended to be more agreement between items in the test method (MEDFICTS) and the recommendations for white children, than between the test method and recommendations for black children. This confirmed that a valid item list is population-dependent.[111]
- In the top ten sources of fat in the diets of twelve-year-old Australian children Gracey [263] reported a list very similar to the MEDFICTS foods and the food listed in this sub-study, except for eggs which do not appear on that list.
- The list also showed some similarity with the findings of Johnson et al [262] except that meats, eggs and convenience foods did not feature on their top ten list for British school children, again emphasising the importance of contextualisation.

It was thus concluded that the test method had reasonable face validity for the current target group of white South African children. Further support for face and content validity comes from the fact that this list showed substantial similarity with numerous published fat screeners, for example Caan et al [49], who modified Block et al's original fat screener [206]; the 'Dietary Risk Assessment ' [213, 214]; the 20-item short questionnaire to qualitatively assess the intake of total fat, saturated, mono-, polyunsaturated fatty acids and cholesterol of Rohrmann & Klein [264]; the brief FFQ for fat, fiber, and fruit and vegetables intake of rural adolescents by Buzzard et al. [265] In addition, recent analyses of the dietary sources of nutrients among a nationally representative sample of US adults [22, 266] yielded on visual inspection similar results in terms of the food sources of fat, saturated fat and cholesterol.

3.1.2.1.4 Practical development

For the current tool, colour pictures of South African (branded and generic) foods within each food category were used instead of text for the following reasons:

- Jonsson et al [267] have argued that the reliability, validity and usability of a measure of dietary habits can be increased if the measure does not require sophisticated linguistic abilities, and concentrates more on visual and comprehension skills.
- In a society with many languages this has obvious additional advantages. An instrument of this nature is presumably more likely to be suitable for a trans- or multicultural target group.
- This approach would make the tool also suitable for children with limited literacy.
- Visual appeal can never be wrong, even though it may be possible that brands could also cause problems, for example amongst brand loyal consumers whose particular preference is not pictured, even though it is implied to be part of the category because of similar nutritive value. Equally, the depicted example food may be a particular brand not chosen by a respondent, or the availability of certain brands (for example yoghurt) may be regional. Finally, new foods tend to enter the market at a very fast rate.

In order to minimise the limitations, yet keep the advantages of package recognition, a clarifying introduction (in standardised text) was always offered when a new category was presented to the respondent. This specified what the picture under discussion showed, for example, in the case of meats the introductory statement was as follows: "This is a picture of various sorts of meat. It ranges from beef like steak, boerewors, biltong and minced meat, to pork, mutton and chicken. Organ meat, like kidneys and liver, is also included. Processed meats, such as cold meat, bacon, polony and spreads also form part of this group." Table 3.2 provides an overview of the item list (food categories and pictured example foods) and in Addendum A are reduced copies of the pictures in the tool.

For the current study, the first question on the initial tool was always: "Do you eat foods such as those on the picture?" For those participants who responded positively to this filter question a follow-up question "Which one food do you eat most of the time?" was posed, because Koehler et al [92] suggested that children more accurately recall specific items rather than categories in their 'Yesterday's Food Choices' instrument.

TABLE 3.2: TEST METHOD: DESCRIPTION OF FOOD CATEGORIES, FOOD PICTURES INCLUDED, REFERENCE PORTION SIZE AND PORTION SIZE ESTIMATION AIDS (PSEA)

Category	Sub-category / group	Code	Food pictures	Reference portion size and PSEA's (P=photo), (D=diagram) (H=Household measure)
Meats	-	M	Beef: steak, boerewors, biltong, minced meat Pork: Chicken: Mutton: chops, roast Processed meat: cold meats, bacon, polony, spreads Organ meat: kidney, chicken liver	90mm diameter circle (D)
Eggs	-	E	Eggs	2 eggs (P)
Dairy	Milk, whole or fat reduced	DM1	Fresh or long life milk (full cream or 2%) Milk powder (full cream or blends) Coffee/tea creamers Condensed or evaporated milk Full cream yoghurt	250mL (H)
	Milk, skim/fat free	DM2	Fresh or long life fat free milk Powdered fat free milk Fat free / low fat yoghurt or drinking yoghurt	250mL (H))
	Cheese, full cream	DC1	Hard cheese (full cream) Processed; spread or wedges (full cream) Cottage/ cream cheese (full cream)	3 slices ea 30x90x2mm (D) 2 wedges (D) 125mL (H)
	Cheese, fat reduced	DC2	Hard, low fat cheese Processed low fat; spread or wedges Cottage cheese, low fat	3 slices ea 30x90x2mm (D) 2 wedges (D) 125mL (H)
	Dessert, full cream	DD1	Milkshakes Full cream ice cream Full cream custard Full cream dessert (blancmange type)	125mL (H)
	Dessert, fat reduced	DD2	Low fat flavoured milk drinks Low fat frozen dessert Frozen yoghurt	125mL (H)
Fried foods	-	F	Chips (French fries) Fried vegetables (eg onion rings) Fried chicken (eg whole or pieces) Fried fish (eg hake) or fried seafood (eg calamari) Fried meat (eg sausage) or fried eggs	90mm diameter circle (D) 125mL (H)
In baked goods	-	I	Cakes Cookies/ biscuits Sweet tarts / pastries (eg chelsea buns, doughnuts, eclairs) Savouries (eg samoosas, croissants, vetkoek) Rusks	120mm diameter circle (D)
Convenience foods	-	C	Tins / cans (eg spaghetti) Packaged (eg pasta sauces, noodles) Frozen meals (eg pizza)	100x100mm square (D) 250mL (H)
Table fats	High fat	T1	Butter Margarine brick wrapped in paper Mayonnaise / salad dressing Peanut butter	5mL (H)
	Fat reduced	T2	Low fat salad dressing / mayonnaise 'Lite' / medium fat tub margarine	5mL (H)
Snacks	High fat	S1	Chocolates Chips and cheese puffs (eg Niknaks) Regular savoury crackers Peanuts	Photo: 1 small packet chips 50g chocolate 22g peanuts 6 crackers, high fat
	Fat reduced	S2	Low fat crackers or pretzels Fruit bars Hard candy, liquorice, jelly sweets, marshmallows	3 ProVita (P) 1 fruit bar (P) 10 sweets, low fat (P) 30g pretzels, plain (P)

3.1.2.2 Quantification

3.1.2.2.1 Background (original tool)

According to Block et al [44] quantification includes the portion or serving size attributions of each line item and also the assumed nutrient content of each. In addition, the measurement of frequency of intake should also be considered since all of these eventually determine the relevance of the score obtained.

3.1.2.2.2 Application (current tool)

In the current tool the following three specific aspects of the quantification were investigated in more depth in the proposed target group: the reference portion size, the portion size estimation aids (PSEA) and the frequency of intake.

- ***Reference portion size***

Most of the reference portions in the original MEDFICTS tool are consistent with definitions given by the American Diabetes Association exchanges, American Heart Association, and the US Department of Agriculture handbooks.[203]

These original reference portion sizes were adjusted for most food categories since it was felt that the relevance for the target group was in some cases questionable. The same team of dietitians mentioned before, again held a consensus workshop. Recommendations were made based on their experience and in some by cases practically weighing and measuring example foods. In order to field-test these reference portion sizes a small survey was conducted.

For the purpose of this study small portion was defined as half as much or less than the reference portion. A large portion was equivalent to one-and-a-half times or more the size of the reference portion. In the original tool 'less' or 'more' than the reference were seen as small and large respectively.

Sub-study 2: Reference portion size

Rationale: Ease of completion and brevity are desirable for screeners. Quantitative FFQ's (such as MEDFICTS) require that respondents specify their usual intake relative to a given medium portion. The question arose whether the proposed medium portion size was valid for the intended target group, and whether the screener could be simplified to a semi-quantitative format (in which medium intake was assumed).

Aims: To explore ten to twelve year old children's perception of a medium portion and the actual amounts dished up by them within the context of a given meal for three different example foods from the test method.

Participants: Fifty-two children (21 male and 31 female), who met the inclusion criteria of being familiar with and not allergic to the test foods were systematically chosen from the grade four, five and six class lists of an urban primary school.

Methods: Each child was requested to dish up the amount of test food (minced meat, chips and margarine) that was usually eaten. For minced meat the plate already contained cooked white rice and carrots. The chips had to be added to a fast food hamburger, and the margarine had to be spread onto a 40g slice of bread. The amount dished up was weighed on an electronic Soehnle kitchen scale. For assessing the perception of medium size portion, four different portion sizes of each test food were then presented in random order to each participant on separate plates: 50%, 100%, 150% and 200% of the proposed reference portions of the test foods, that is 90g minced meat, 45g chips and 5g margarine, keeping the accompanying foods identical. After an introductory definition ("A medium portion is the amount eaten by children to stay healthy"), the participant had to report which amount, in his/her opinion, reflected a medium portion. Test-retest reproducibility was checked by having each fifth (10%) subject repeat the assessment.

Results and discussion: The mean age of the participants was 10.8±0.8 years and mean BMI was 17.9±2.42kg/m^2. The 45g and 90g portions of minced meat were perceived by about equal numbers (23 and 25 respectively) of participants as "the amount a child eats to stay healthy", even though the mean amount actually dished up was about 45g (Table 3.3). The mean may to be age-related, particularly amongst the boys. The fact that almost 35% of children reported disliking minced meat (data not shown) may also be of relevance. Most children perceived the 40g portion of chips as the medium portion. This was in agreement with the mean amount dished up (44.6±18.92g). No age-related pattern could be detected, but girls always dished up less than the boys (Table 3.3). In the case of margarine, the 5g portion was perceived by the majority of children as 'medium', whilst the mean amount spread onto bread was 4.12±3.76g. Similar to minced meat, many reported disliking margarine (data not shown), and age may also here be a determinant, specifically in boys (Table 3.3).

Conclusion: Perceived medium portions were sometimes (for example for chips and margarine) perceived to be the same as the adjusted reference. Mean actual intake might also be similar (for example for margarine), but inter-individual variations (as evident from the large standards deviations) and possible age and gender-related differences, the small sample size and the limited number of example foods did not yet support the use of one assumed reference portion size per food category as in a semi-quantitative FFQ.

TABLE 3.3: PERCEIVED MEDIUM PORTION SIZE AND MEAN AMOUNT DISHED UP OF MINCED MEAT, POTATO CHIPS AND MARGARINE BY AGE AND GENDER (n=52)

Food	Perceived medium portion size (g)	Number of participants						Total
		10 years Total = 22		11 years Total = 19		12 years Total = 11		
		Male *Total=8*	Female *Total=14*	Male *Total=9*	Female *Total=10*	Male *Total=4*	Female *Total=7*	
Minced meat	45	3	9	1	8	0	2	23
	90	5	4	6	2	3	5	25
	135	0	0	1	0	1	0	2
	180	0	1	1	0	0	0	2
	Mean amount dished up (g)	36.3±12.7	36.5±9.7	59.9±26.2	37.2±9.9	75.0±18.1	44.7±11.3	44.7±18.9 g
		36.4±10.6		48.0±22.1		55.7±20.2		
Potato chips	20	0	4	2	1	0	1	8
	40	5	9	3	7	0	3	27
	60	2	0	3	1	3	3	12
	80	1	1	1	1	1	0	5
	Mean amount dished up (g)	46.9±16.4	32.1±8.8	68.3±23.6	39.3±10.7	59.3±12.8	35.7±8.1	44.6±18.9 g
		37.5±13.8		53.1±23.0		44.3±15.2		
Margarine	2.5	1	5	3	3	0	0	12
	5	4	7	4	6	2	6	29
	7.5	3	2	0	1	2	1	9
	10	0	0	2	0	0	0	2
	Mean spread on bread (g)	3.4±2.4	2.4±2.5	5.9±3.1	2.8±2.3	9.5±8.9	4.9±2.3	4.1±3.8
		2.8±2.4		4.3±3.1		6.6±5.7		

- ***Portion size estimation aids (PSEA)***

The original tool does not include PSEA's, but Kris-Etherton et al [203] do mention that the use thereof could enhance the value of the tool. Two-dimensional serving size measurement aids (life-size photos and geometric shapes) [26, 129] were developed for each food category to assist in estimating quantity of intake. Again, these were the result of a consensus workshop with the mentioned three private practicing dietitians, whose clientele included many children, and the researcher.

In order to field-test the PSEA a developmental evaluation sub-study was conducted:

Sub-study 3: Portion size estimation aids (PSEA)

Rationale: Cost and space favour the use of two-dimensional (2D) PSEA in dietary screeners, but this should not be at the expense of validity and reliability. In adults 2D and three-dimensional (3D) PSEA's appear to yield similar results,[129] but this has not yet been demonstrated in children.

Aims: To investigate the ability of twelve-year old children to express real servings of example foods in terms of 2D and 3D PSEA's, and to compare 2D to 3D PSEA's.

Participants: Of 60 twelve-year old children randomly chosen from the class lists, 22 boys and 21 girls (total 43) and their parents provided timely written, informed consent for participation.

Methods: In an one-to-one, standardised encounter each child was shown a known amount of five true, ready to consume foods (milk, chicken, bread, butter and chips), and requested to indicate for each food which one of three different sized /graded 2D graphics (drawings of measuring cups and teaspoons, cutout circle shapes and rectangles) most closely reflected the quantity of the true food. Similarly, the child was asked to describe the true food as 'half as much', 'the same' or 'one and a half times as much' as a 3D food model (Nasko-Ford Atkinson, WI) of the respective food. Milk and chicken information was obtained twice to control for guessing and check for test-retest reliability. Error rate, defined as the total number of wrong answers expressed as a percentage of total respondents, was calculated for each food and PSEA. The McNemar test for symmetry was used to assess differences / bias.

*Results and discussion***:** In the case of milk, no significant difference between the two subsequent administrations for both the 2D (P=0.32) and the 3D (P=0.48) PSEA was found, thus suggesting test-retest reliability. For chicken, however, the two administrations significantly differed for both PSEA's (P=0.0016 for 3D and P=0.0000 for 2D, based on the fact that in both administrations all children answered wrongly). Thus the guess factor could not be ruled out. From Table 3.4 it is evident that error rates are lowest for milk (for 2D and 3D estimations of quantities) and highest for chicken. The mean error rates for all 2D and 3D estimations (milk 1, butter, bread, chips, chicken 1) were 36.26% and 54.42% respectively. For milk, butter and chips the difference between the 2D and 3D PSEA was not statistically significant (P>0.05). When comparing the error rate in the quantification of milk to foods of differing consistency (that is all other foods) a statistically significant difference (P<0.05) was found for all foods when using 3D PSEA, as well as for butter with the 2D PSEA (data not shown). When checking for gender differences, it appeared that only in the case of chips, using the 2D PSEA, the difference between boys and girls was statistically significant (P=0.01).

TABLE 3.4: ERROR RATES FOR 2D AND 3D PSEA[a], AND SIGNIFICANCE OF DIFFERENCE (P) BETWEEN THEM FOR VARIOUS FOODS

Food	Error rate (%)		P-Value[c]
	2D PSEA	3D PSEA	
Milk[b]	11.6	7.0	0.48
Chicken (1st administration)	100.0	72.1	NA
Chicken (2nd administration)	100.0	95.3	NA
Bread	9.3	51.2	0.03
Butter	34.9	95.3	0.65
Chips	25.6	34.9	0.91

[a] Two dimensional and three dimensional portion size estimation aids
[b] First administration only, because of consistency of response (see text)
[c] McNemar test for symmetry

Conclusions: Inconsistent responses and error rates confirmed that children had problems with estimating quantities of individual foods. The magnitude of the error varied across foods. Physical food consistency, but not gender, might have played a role. However, in line with findings of Cypel et al [129] for adults, 2D and 3D PSEA's did not appear to differ significantly. It is concluded that 2D PSEA's could be retained in the screener, but a reference method in a validation study should include weighing (in contrast to estimation) in order to describe the likely quantification error.

- ***Frequency of intake***

It has been claimed that the main determinant of variation in measuring dietary intakes is frequency of consumption of the food items in the list.[140] In the original MEDFICTS tool three response options are given for categorizing frequency of intake per week: Once or less, up to three times, four or more times. Kris-Etherton et al [203] assert that this frequency grouping would minimise the tendency of persons to underestimate intake, and that it resembles those used in existing questionnaires.

From the review of literature it was evident that reporting frequency of intake is a cognitively challenging task. It was not known whether children would be able to perform these tasks with the necessary accuracy. Another question that arose was whether a (graphic) depiction of frequency of intake, for example as bar charts or abacus type presentations could enhance responses. Whilst graphics might have eye-appeal, they might also be an abstraction, and might reduce rather than improve accuracy in the target group.

A developmental evaluation sub-study was conducted to assess the above-mentioned problems.

Sub-study 4: Frequency of intake

Rationale: The cognitive tasks involved in providing a correct frequency of intake response on a grouped weekly consumption FFQ, include at least the following steps:

1. Recognizing the individual foods / items on the list
2. Understanding the food categorisation (grouping) principle (which foods form part of a specific food category)
3. Recalling own intake
4. Counting number of times per week foods are eaten
5. Summation of counts of foods belonging to a food category
6. Conversion of summated intake to the frequency of intake format used in the data collection tool (recording the answer in the required questionnaire format)

Steps 4 and 6 were the focus of this developmental evaluation sub-study.

Aim: To assess the ability to correctly count and record frequency of food intake in three example cases (reflecting low, medium and high weekly consumption), as well as preferred depiction of this frequency of intake (MEDFICTS response categories versus graphics and absolute numbers), in ten to twelve year old school children.

Participants: A systematic sample of 39 ten to twelve year old children (19 male, 20 female) was selected.

Methods: On an individual, random rotation principle, each child was exposed to three example cases reflecting low, medium and high weekly consumption (corresponding to MEDFICTS grouping). They then had to orally indicate "How many times in a week did [the example case] eat [example food]". This was noted by the field worker. The child then had to record his / her answer on each of four different response formats (MEDFICTS grouping; table format with short sentences and numbers; horizontal abacus-type format; vertical bar-graph format with numbers underneath). Finally the four response formats had to be ranked in order of preference.

Results and discussion: Table 3.5 summarises the findings. As expected, the low frequency of intake case caused no problems either in terms of providing a correct answer or in terms of recording the answer into any of the four given response formats.
For both, the medium and high frequency of intake scenarios, however, five subjects (12.5%) provided wrong answers to the question, two of which overlapped between the two intake

scenarios. This was in line with the findings of Hammond et al [112] who found that food items eaten frequently were, on average, least accurately classified compared with those eaten with intermediate frequency. An inability to give a correct answer and an inability to correctly code the own answer will result in an invalid response. Error rate (defined as the percentage responses that were not correctly answered and correctly coded for each scenario and response format) ranged from 0% to 17.9%, with the majority of error within each of the error rates contributed by wrong answers, in contrast to recording errors. In terms of preference, the words with numbers format was chosen as first or second choice by a total of 32 (82.1%) of subjects, compared to 29 (74.4%) who chose the bar diagram with numbers in first or second place. What these two formats have in common was the provision of numbers. The abacus format and the MEDFICTS grouping were chosen by eleven and six subjects respectively as first or second choice in terms of preferred response format. Apart from a horizontal versus a vertical alignment respectively, the abacus and bar diagram were very similar, except that the bar diagram also provided the corresponding numbers below each bar. It appears that graphics represent an abstract representation or distraction that is disliked rather than assisting correct response. Furthermore, response categories (pre-set groupings) also require an additional step of placing an answer in a suitable group (category), again rather contributing to dislike than making the recording process easier.

Conclusions and recommendations: Children do make reporting mistakes when required to indicate weekly consumption of individual foods, particularly in medium to high intake situations. More of the error was attributable to giving wrong answers in the first place, than to incorrectly recording an answer in a particular response format. Thus, the focus should be on helping the children to correctly count their weekly consumption. However, a more preferred recording format would include the use of concrete numbers (not groupings) and avoiding abstractions (for example graphics). Avoiding groupings would have the additional advantage of eliminating the overlap in the MEDFICTS consumption categories ‘rarely or never’ with ‘3 times or less’. Since summation problems may have made it difficult for the children to calculate weekly intake in the high frequency scenarios, a ‘per day’ response option could be considered for foods with an expected high frequency of intake.

TABLE 3.5: CORRECTNESS OF ANSWERS AND RECORDING FOR THREE FREQUENCY OF INTAKE SCENARIOS, AS WELL AS PREFERRED WAY OF RECORDING OF 10-12 YEAR OLD CHILDREN (n=39)

Scenarios	Answers	Recording							
		MEDFICTS Categories		Words with numbers		Abacus		Bars with numbers	
		Correct *n(%)*	Wrong *n(%)*	Correct *n(%)*	Wrong *n(%)*	Correct *n(%)*	Wrong *n(%)*	Correct *n(%)*	Wrong *n(%)*
High frequency intake	Correct (n=34)	32(82.1)	2(5.1)	34(87.2)	0(0)	34(87.2)	0(0)	33(84.6)	1(2.5)
	Wrong (n=5)	4(10.3)	1(2.5)	5(12.8)	0(0)	5(12.8)	0(0)	5(12.8)	0(0)
	Error rate[a]	17.9%		12.5%		12.5%		15.4%	
Medium frequency intake	Correct (n=34)	34(87.2)	0(0)	34(87.2)	0(0)	32(82.1)	2(5.1)	34(87.2)	0(0.0
	Wrong (n=5)	5(12.8)	0(0)	5(12.8)	0(0)	5(12.8)	0(0.0)	5(12.8)	0(0.0)
	Error rate[a]	12.5%		12.5%		17.9%		12.5%	
Low frequency intake	Correct (n=39)	39(100.0)	0(100.0)	39(100.0)	0(100.0)	39(100.0)	0(0.0)	39(100.0)	0(0.0)
	Wrong (n=0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)
	Error rate[a]	0.0%		0.0%		0.0%		0.0%	
First choice		2(5.1)		14(35.9)		5(12.8)		18(46.2)	
Second choice		4(10.3)		18(46.2)		6(15.4)		11(28.2)	
Third choice		6(15.4)		5(12.8)		20(51.3)		8(20.5)	
Least preferred		27(69.2)		2(5.1)		8(20.5)		2(5.1)	

[a]Error rate = ((Total - (Number correctly answered and recorded) / Total) expressed as percentage; that is (39 minus (number correctly answered AND correctly recorded) divided by 39 and multiplied by 100)

3.1.2.2.3 Scoring

Portion size: In the original tool small, average and large portion sizes (relative to the reference) are respectively scored 1, 2 and 3 for the high fat group. The same scoring was used for the current tool. In the original tool no points are given for low fat choices, except if in the meat group 2 portion size was large. In this case 6 points are given. Since no meat group 2 was included in the current tool, these points were of no relevance.

Weekly consumption: In the original tool intakes of three or less servings per week receive three points, whereas a weekly consumption of four or more servings gets rated seven points. This applies to the high fat groups (group 1) of each food (sub)category. No points are given for intakes in group 2 (the low fat group) of the food categories. This same scoring system was used in the present study, based on the rationale given by Kris-Etherton: [203] Foods eaten rarely or never were assumed to contribute an insignificant amount of fat to the diet. Points for weekly consumption were derived from the assumption that a person consuming four or more servings per week (maximum seven) would average 5.5 servings per week. Similarly, persons consuming three or fewer servings per week would average two servings per week.

Four or more servings per week contributed approximately 73% (5.5/[5.5+2]) of the intake for this category of foods and three or fewer servings per week contributed approximately 27%. The simplified multiplication factors under weekly consumption became seven and three respectively.

Category score and final score: For each (sub) category the weekly consumption point was multiplied by the serving size point to obtain a category score. The category scores were added to yield a final score that could thus range from 0 to 216. In the original instrument a final score of 40 to 70 reflects a Step 1 diet, whereas a final score of less than 40 suggests a Step 2 diet. In the Revision 2000 of the American Heart Association the 'Step 1' designation has been replaced by 'major guidelines for the general population' and the 'Step 2' by 'medical nutrition therapy'. As the tool was intended for community-based use, specifically for children, it was decided to only dichotomise the final score into 'high fat' and 'prudent' intake (that is the major guideline), thus omitting the medical nutrition therapy.

3.2 REFERENCE METHODS

As indicated previously two reference methods were used in this study: The food record and parental completion of the screener.

3.2.1 Reference method 1: Food record

3.2.1.1 Background

The three-day food record was the primary reference method in this study. It was therefore considered essential to conduct field-testing in order to optimise the data collection.

3.2.1.2 Sub-study 5: Food record

The completion of the three-day food record was intended to be a mathematics assignment in the new South African outcome-based education approach (Curriculum 2005).[108] The school identified for testing agreed to participate, but requested that all the children in the grade be included. No incentives were given, however, the data collection and recording tasks were considered a practical assignment, which would contribute to each pupil's practical mark. Pupils were also informed that the recorded information would be analysed by them later; again in the form of a mathematics work sheet. (This was only done after all data were collected in order to prevent that the learning effect and increased food awareness affected the subsequent data collection). Worksheets on which learners performed curriculum-related data analyses (summation, rank ordering, calculation of means and proportions, and compilation of graphs) were set up for meaningful integration of nutrition and mathematics and for feedback. The intended learning outcomes and mathematics assignments based on the food records were jointly compiled by the researcher and the involved teacher.

3.2.1.2.1 Form

The format for the three-day food record resembled a simplified version of the one often used by the USDA in the Beltsville Human Nutrition Research Centre.[268] It was an open-ended form with each of the three days for recording printed on a different colour paper (Addendum B). For each recording day three A4-sized papers were printed with suitable headings. On the first page several examples were filled in. Each child also received written instructions.

3.2.1.2.2 Time frame and programme

In a two-week time span (10/9/2000 to 21/9/2000 that is early spring in South Africa), three recording periods of each three consecutive days (that is Sunday-Monday-Tuesday, Thursday-Friday-Saturday, and Tuesday-Wednesday-Thursday) were identified for documenting dietary intake. The programme schedule was set up in such a way that, for the sample as a whole, all days of the week were represented, and both, weekdays and weekend-days, formed the starting point for record keeping. This approach prevented recording fatigue and inaccurate reporting as observed by Gersovitz [50] to fall on a particular day. These researchers found a "significant association of actual and recorded values on the first two days of recording, but decreasing accuracy of recording afterwards".

Thus, in total, intakes of seven weekdays and two weekend-days were measured. Tuesday and Thursday were each represented twice (Table 3.6).

TABLE 3.6: NUMBER OF TIMES THAT EACH DAY OF THE WEEK FORMED FIRST, SECOND OR THIRD RECORDING DAY

Day of week	**Day of recording**			
	First	Second	Third	Total
Sunday	1	0	0	1
Monday	0	1	0	1
Tuesday	1	0	1	2
Wednesday	0	1	0	1
Thursday	1	0	1	2
Friday	0	1	0	1
Saturday	0	0	1	1
Total	3	3	3	9

3.2.1.2.3 Sample composition and grouping

All 176 grade six pupils from an Afrikaans speaking, primarily middle-class, public primary school in urban Pretoria, South Africa, were included in the testing study. All pupils in the grade were white. Three classes were taught by the same mathematics teacher. From each of these, eleven pupils were randomly chosen from the alphabetic class list and allocated to the electronic scale group (only eleven electronic scales were available). These children were given a Soehnle digital scale to take home. From the remaining children those that had kitchen scales at home and who had parental permission to use these, were requested to record intakes using these scales. Children indicated the type of scale (spring, balance, digital) they had used on the record form. The rest of the children were provided a set of household measuring cups (250mL) and spoons (15mL tablespoons and 5mL teaspoons).

3.2.1.2.3 Training and briefing

The mathematics teacher who taught the three classes was trained to brief all the children. All data collection and briefing materials were provided to him, including:

(i) A <u>reference file</u> containing

- Copies of the data collection materials, that is
    - general and specific instructions
    - food record forms
    - food description flow diagrams
- Enlarged transparencies of the above, partially completed with relevant examples
- Class lists for recording groupings (according to type of quantification method used by each child [provided electronic scale, home scale or household measure] and for indicating which scale was issued to which child)
- Transparency pens

(ii) A <u>demonstration kit</u> for practically showing the correct use of the electronic scale containing:

- Electronic scale
- Plate and knife
- Slice of bread
- Margarine
- Cheese
- Coffee mug
- Apple

(All foods were in separate containers)

The training session focused on the following:

- Proper administrative and logistical matters
- Correct completion of the food record form, specifically in respect of describing food using the provided flow diagrams
- Appropriate use of the TARA function of the scale (practical demonstration: slice of bread first spread with margarine, then cheese; zeroing of scale after utensils, that is plates and mugs have been weighed.)

During the briefing six food description flow diagrams (meats, fats, bread, fluids, fruit / vegetables, snacks) similar to those proposed by Wold et al [156], but adjusted for South African circumstances were explained to the class using overhead transparencies. The diagrams were also included in each child's instruction folder for continuous referral in compiling the detailed description of foods eaten (see Addendum C). The children were requested to hand in wrappers of purchased foods in a plastic envelope, which was also part of the handout.

The children were instructed to bring their completed forms to school after the first day of recording to check for compliance.

3.2.1.2.5 Findings and conclusions

Each returned food record was rated by the researcher using comprehensiveness of completion of the "description of foods" column as only criterion. This eliminated discrimination based on type of quantification (that is type of scale or household measure, or provided versus own scales), potential measurement accuracy (again based on type of quantification method available to a child) or number of foods recorded (that is presumed completeness of record).

The marks given by the researcher are stated in the bottom row of Table 3.7.

Assessment of the analysis assignments was purely a school mathematics activity.

TABLE 3.7: NUMBER OF SUBJECTS IN THE THREE INTAKE QUANTIFICATION GROUPS, AS WELL AS MARKS ACHIEVED

Class and recording period		Quantification method			Total
		Electronic scale	Own non-electronic scale	Household measure	
Grade 6S (Sunday to Tuesday)		11	19	5	35 out of a class of 36
Grade 6T (Thursday to Saturday)		12	20	1	33 out of a class of 35
Grade 6P (Tuesday to Thursday)		10	18	4	32 of a class of 35
Grade 6E (Thursday to Saturday)		1	21	14	36 out of a class of 36
Grade 6H (Thursday to Saturday)		2	26	6	34 out of a class 35
Total		36	104	30	170 out of 176 learners
Marks for assignment (out of 10)	10	16 (44.4%)	11 (10.5%)	2 (6.7%)	29 (17.1%)
	9	9 (25.0%)	21 (20.0%)	2 (6.7%)	32 (18.8%)
	8	3 (8.0%)	22 (21.1%)	5 (1.7%)	30 (17.6%)
	=<7	8 (22.2%)	50 (48.0%)	21 (70.0%)	79 (46.5%)

Four children had electronic scales at home. In Table 3.7 these are indicated together with the handed out electronic scales. One of the randomly chosen children in this group fell ill, resulting in a total of 36 respondents who measured food intake electronically. Thirty children assessed their intake with household measures provided to them, and the rest (n=104) used home scales (all of which were reported to be spring type scales).

In total 61 learners returned detailed records (nine or more marks given for assignment). The food record of 30 pupils was rated eight out of ten in terms of comprehensiveness of food description, whilst 79 received seven marks or less. Six records were not returned (in two cases because of illness). This represents a response rate of 96.5%.

As evident from Table 3.7, the response rate was not related to a particular quantification method or class. However, the data quality (in terms of comprehensiveness of food descriptions) was clearly superior in the group that weighed their intake on electronic scales. Forty four percent of this group achieved full marks, whilst 70% of those who had been given household measures scored seven or less marks out of ten.

It was concluded that the electronic scales may have acted as a form of motivator and since the food record was supposed to be the reference method in the main study, it was decided to obtain more electronic scales (even though it was not financially possible to do this for all children). The finding from sub-study 3 strengthened this decision. Furthermore, for logistical reasons and for consistency in terms of administration it was decided to only use classes taught by one teacher in the main (validation) study.

3.2.2 Reference method 2: Screener by parents

For the second reference method, a text version of the dietary fat screener was compiled for completion by parents on behalf of their respondent child (Addendum D). This was very similar to the original NCEP tool (apart from the South African food examples within the food categories, the omission of group 2 sub-categories to match the test method, the translation to Afrikaans and layout to fit one page), and was therefore not subjected to developmental evaluation. Colleagues of the researcher checked understandability and technical aspects.

3.3 REFLECTION

As part of the confirmation that dietary assessment is never perfect, the developmental evaluation sub-studies on the test method greatly contributed to the understanding of some of the errors to be expected from children completing a FFQ type assessment. Where feasible, adjustments could be made and consequently the main study could be approached with confidence. This was strengthened by the knowledge that reference method 1, the three-day food record, was not only accepted, functional and standardised in the mathematics context, but that it could make a meaningful contribution to the curriculum.