Food Quality and Preference 48 (2016) 238–250

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Developing a novel tool to assess liking and wanting in infants at the time of complementary feeding – The Feeding Infants: Behaviour and Facial Expression Coding System (FIBFECS)

M.M. Hetherington [a,*], J. Madrelle [a], C. Nekitsing [a], C. Barends [b], C. de Graaf [b], S. Morgan [a], H. Parrott [a], H. Weenen [c]

[a] Institute of Psychological Sciences, University of Leeds, Leeds LS2 9JT, UK
[b] Division of Human Nutrition, Wageningen University, Wageningen, The Netherlands
[c] Danone Nutricia Research, 3584CT Utrecht, The Netherlands

ARTICLE INFO

Article history:
Received 11 November 2014
Received in revised form 24 September 2015
Accepted 25 September 2015
Available online 9 October 2015

Keywords:
Complementary feeding
Weaning
Like/dislike
Facial expressions
Wanting
Food refusal
Infant feeding

ABSTRACT

Introduction: Consumption of foods is determined in part by how much a food is liked. However, assessing liking in infants is difficult. Research with infants has often relied on indirect measures such as intake or subjective ratings from mothers. Therefore the aim of the present research was to devise a tool adapted from existing techniques which can directly and systematically measure liking in infants during the weaning period.
Method: A tool was developed by extracting items from previous studies. In all, 13 items were generated, which included 6 behaviours reflecting avoidance and approach: turning away, arching back, pushing spoon away, crying/fussy, leaning forward and rate of acceptance; also 7 facial expressions thought to reflect affective response; brow lowered, inner brow raised, squinting, nose wrinkling, upper lip raised, lip corners down and gaping. An e-training manual was developed with a certification test to train coders. The coding tool is based on coding the first 9 spoonfuls for each infant. 63 videos were coded by 4 raters, each video was coded by at least 2 different coders. For each spoonful the absence or presence of each item was recorded; for rate of acceptance, a four point scale was used.
Results: In the certification test most cues were high in agreement for all coders. Factor analysis indicated two dimensions, one which largely captured gross behaviours and the second featuring a cluster of facial expressions. Internal consistencies of the overall scale and the behaviour and facial expression subscales were acceptable as indicated by Cronbach's alpha >0.7. Intra-class correlation indicated moderate to high inter-rater reliability and test–retest reliability for most of the cues. Spearman correlations indicated significant associations of the total number of negative behaviours with rate of acceptance and overall facial expressions. Rejection behaviours corresponded with a low rate of food acceptance and a high rate of negative facial expressions. Two parameters occurred less frequently and did not appear to provide any further discriminatory ability, namely leaning forward and crying/fussiness, these can be removed from the scale for future use.
Conclusions: The Feeding Infants: Behaviour and Facial Expression Coding System (FIBFECS) is structurally valid and reliable for use by trained coders and those who are researching infant eating behaviour. The two factor structure of the tool suggests that the facial expression subscale reflects liking and the behaviour subscale wanting. The tool could also be adapted for mothers and professionals to detect liking and wanting through facial expression and behavioural cues respectively.

1. Introduction

Healthy eating habits begin with pregnancy (Mennella, Jagnow, & Beauchamp, 2001), breastfeeding (Maier, Chabanet, Schaal, Leathwood, & Issanchou, 2008; Mennella & Beauchamp, 1997; Mennella, Forestell, Morgan, & Beauchamp, 2009) and weaning

* Corresponding author at: School of Psychology, University of Leeds, Leeds LS2 9JT, England, UK.
E-mail address: M.Hetherington@leeds.ac.uk (M.M. Hetherington).

http://dx.doi.org/10.1016/j.foodqual.2015.09.010
0950-3293/

(Fildes & Cooke, 2012). Offering a variety of solid foods including vegetables during complementary feeding (weaning) increases their acceptance (Caton et al., 2014; Maier et al., 2008). However, beyond weaning, consumption of foods is linked to how much a food is liked (Bere & Klepp, 2005; Gibson, Wardle, & Watts, 1998; Olsen, Ritz, Kraaij, & Møller, 2012) which in turn is linked to energy density (Gibson et al., 1998). At weaning, infants appear to accept most solid foods, but later in childhood as neophobia sets in it is more difficult to encourage children to accept new foods (Caton et al., 2014).

It is assumed that infants like the foods they are offered at weaning since these are consumed, however, liking in infants is difficult to assess due to limited communication capacity. Scales which are used with older children such as ratings are not possible in infancy. Therefore, studies have used objective but indirect measures to assess infants' liking such as, intake (weight), ingestion ratio (based on volume consumed), duration of eating (Forestell & Mennella, 2012, 2007; Schwartz, Issanchou, & Nicklaus, 2009; Mennella & Beauchamp, 1997; Mennella et al., 2001; Mennella, Forestell, Morgan, & Beauchamp, 2009), and pace of eating (Forestell & Mennella, 2007; Mennella & Beauchamp, 1997; Mennella et al., 2009). Both the relevance and prominence of these parameters will depend on other factors such as state of hunger, eating traits and maternal feeding practices.

Liking has also been assessed by asking caregivers (usually mothers) to make a judgment of how much their infant likes a food (Forestell & Mennella, 2012, 2007; Mennella et al., 2001; Maier et al., 2008). Although this measure of liking can be direct, it is subjective, open to bias and the criteria used to assess liking are not always clear. The variety of cues which mothers draw upon to provide information on liking can make comparison across infants and across studies difficult. Alternatively an external observer (usually the researcher) may also rate perceived liking (Maier et al., 2008). It is important that the researcher is familiar and sensitive to the way infants express their liking but have no prior experience with the particular infants in the study to avoid bias. The ratings provided for different infants by the researcher are likely to be comparable within any one study, however, such ratings are also subjective and comparison across different studies is challenging.

In an attempt to obtain more accurate, direct and systematic measures of liking several studies have used video recordings of feeding sessions; see Table 1. The video recordings have allowed researchers to observe infants' reactions to foods in more detail, as the same assessor(s) can repeatedly observe the recordings which is not possible in real time observations or in natural settings. In order to code these video recorded sessions it is important to know what behaviours to search for and to document.

From a very early age infants are able to contract their facial muscles and express primary emotions, very similar to those of adults (Ekman & Oster, 1979). As a result many studies have used facial expressions to assess liking in infants (Forestell & Mennella, 2012, 2007; Mennella & Beauchamp, 1997; Mennella et al., 2001, 2009; Zeinstra, Koelen, Colindres, Kok, & de Graaf, 2009; Soussignan, Schaal, Marlier, & Jiang, 1997; Rosenstein & Oster, 1988). The Facial Action Coding System; FACS (Ekman & Friesen, 1978) is a widely used research tool for measuring facial expressions in humans by detecting facial movements. Within the FACS the actions of individual or grouped muscles contribute to "action units" which are identified and coded. The facial display in humans (and other non-human primates) is a means to communicate to others. For example core emotions conveyed by human facial expressions such as anger, disgust, happiness and sadness appear to be universally understood across cultures (Ekman & Friesen, 1978). It is assumed that facial expressions displayed by babies might also reflect underlying emotional states to encourage communication with their caregiver. Such behaviours will have strong evolutionary significance in promoting appropriate responses from carers. A similar tool known as BABY FACS was developed for infants by Oster (2004). It is necessary to adapt this tool for infants since morphological differences between adults and infants necessitate a specialised version. For example, there is more fat in the dermis of an infant's face, and more elastic skin than adults giving a different and a smoother facial response than for adults. Facial expression coding from BABY FACS has been implemented in several studies to assess food liking in both infants and older children (Forestell & Mennella, 2012, 2007; Zeinstra et al., 2009; Mennella et al., 2009; Soussignan et al., 1997; Rosenstein & Oster, 1988).

A number of studies which have measured infants' facial expressions as a measure of acceptance or liking, have correlated these with indirect measures of liking such as intake (weight), feeding duration, pace and mother's overall judgement of infant liking (Forestell & Mennella, 2012, 2007; Mennella & Beauchamp, 1997; Mennella et al., 2009). These studies found that infants who displayed fewer expressions of distaste were likely to eat more, have a faster eating pace and judgements of liking by mothers were higher. This shows good correspondence between facial expressions and more objective measures such as intake and pace of eating.

Video coding in addition to other measures of liking provide triangulation (maternal ratings, researcher judgements and independent coding from facial expressions) to improve levels of description and explanation. A study by Mennella et al. (2009) examined the effect of exposure to milk formulas with different sensory properties on liking and intake in infants aged 4–9 months. The infants who were familiar with drinking hydrolysed casein formulas (with noticeable bitter, sour, and savoury tastes compared with breast milk), displayed fewer negative facial expressions when eating bitter and savoury cereals compared to the infants who were unfamiliar with such milk. This suggests higher levels of liking or at least acceptance of these flavours as a function of experience. Thus differences in liking can be detected via facial expression analysis.

There are subtle but important differences in assessment of liking between indirect (e.g. how much is eaten) and direct measures (facial responses to the food). For example in one study it was shown that although intake was similar across conditions, facial expressions did differ by condition (Mennella et al., 2001). In this study the response of weaning age infants whose mothers consumed carrot juice during pregnancy or breastfeeding were filmed when given cereal with added carrot during weaning. Infants displayed fewer negative facial expressions to the carrot flavoured cereal compared to infants who had never been exposed to carrot before, but intake did not differ by exposure. Since mothers spoon-fed their babies, it is possible that intake did not vary according to liking as mothers were in control of amount eaten. Nevertheless it raises the possibility that liking can be expressed separately from amount eaten.

Some studies have incorporated other expressive behaviours related to feeding such as mouth movements in response to spoon approaching, vocalisation (Dearden et al., 2009; Mennella & Beauchamp, 1997), body movements (Mennella & Beauchamp, 1997) or other autonomic parameters such as skin temperature, regulation of respiration and eye movements (Soussignan et al., 1997). Clearly the use of these other behaviours indicates that facial expression alone may not capture all the information which mothers will observe during a meal. Therefore facial expressions in combination with other feeding related behaviours such as physical body and mouth movements indicating approach or avoidance could provide a detailed and reliable method of assessing responses to foods in infancy. It is clear that filming, then coding infant behaviours and facial expressions can be time consuming, since comprehensive training is required to recognise and code

Table 1
Behaviours (B1–B5) expressed during feeding sessions/interactions to indicate negative or positive affective responses.

	Hedonic value given	Negative			Positive	
	Reason for occurrence	Break needed/disengaging	Satiety	Overall negative (no explanation)	Hunger	Engagement
B1	Turns head away (from the stimulus) Looks away (up/down/in the corners, distracted)	North Dakota Department of Health NDDOH (2000)	Hodges, Hughes, Hopkinson, and Fisher (2008) Oster (2004) Women Infant Children (WIC (2007))	Mennella and Beauchamp (1997) Mennella et al. (2001)		
B2	Arches back/pulls body away	NDDOH (2000)	Hodges et al. (2008)			
B3	Crying/fussiness	NDDOH (2000)	Hodges et al. (2008) Oster (2004)	Mennella and Beauchamp (1997) Mennella et al. (2001)	NDDOH (2000)	
B4	Pushes bottle/spoon away (with hands) Playful/ Plays with or throws utensils		Hodges et al. (2008) WIC (2007)	Mennella and Beauchamp (1997) Mennella et al. (2001)		
B5	Reaches for food/leans forward Moves head toward the spoon/Tries to swipe food towards mouth (grab spoon)				WIC (2007)	Hodges et al. (2008)Dearden et al. (2009)

both behaviours and facial expressions. Nevertheless, these behaviours could be used to reveal important facets of human infant preferences and also appetite. If it is shown to be reliable and valid, such a system could be used within and across infants to explore food preferences, liking (hedonic response) and wanting (incentive salience) in early life. Since weaning is a particularly important stage in the development of eating habits and a period in which assessment of liking and wanting is difficult or subjective, the aim of the present research was to develop a systematic, reliable method to assess liking and wanting of foods by observing infants' facial expressions and behaviours during weaning. Therefore the present coding tool was developed to assess both feeding related behaviours and facial expressions observed during the feed.

2. Methods

The behaviours and facial expressions included in the current coding system were based on previous literature. These were extracted then applied and tested to assess structural validity and inter-observer reliability.

2.1. Development of the coding system (FIBFECS)

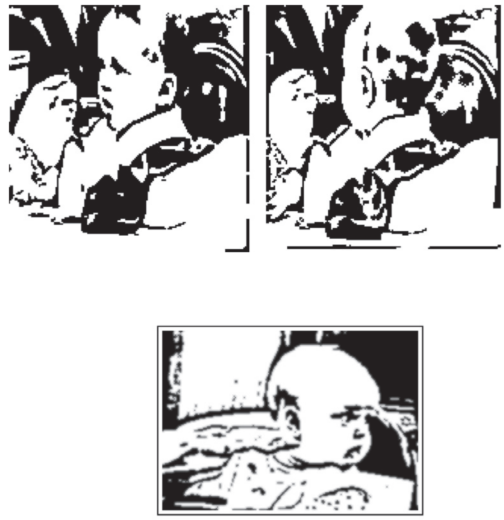
2.1.1. Selection of behaviours and facial expressions

A literature review was conducted to identify behaviours and facial expressions considered in previous studies as measures of the hedonic response to food. Cues for hunger and satiety were also included in this review as these are important to code when considering the context in which hedonic cues occur. For the avoidance/dislike category, four discrete behaviours were identified: B1 turns head away/looks away or looks down; B2 = pulls body away/arches back, B3 = becomes fussy/cries, B4 = pushes spoon away/becomes playful. For the incentive salience or wanting category the following behaviours were identified: B5 leans forward/reaches for food/puts spoon voluntarily in his/her mouth (see Table 1; also Fig. 1).

Infants prepare to eat something acceptable by opening their mouth early during the spoon offer and will do this more frequently if the food is wanted. Preparedness to accept the food is taken here as another measure of "wanting", and was measured by the "rate of acceptance". This was a judgement made on accepting the spoon offered in spatial rather than strictly temporal terms. Thus early acceptance occurred when the infant opened their mouth when the spoon was at a distance to their mouth (score = 3); late acceptance when the spoon was close to the mouth (score = 2); enforced acceptance was scored if the infant only opened their mouth when their lips were touched with the spoon by the mother (score = 1) and refusal was scored as 0. This approach was developed directly from a study by Mennella and Beauchamp (1997) see Table 1.

At the beginning of weaning infants tend to experiment with new textures and are learning how to chew and swallow, as a result, some foods may be ejected from the mouth. Therefore spitting out was not selected as an explicit avoidance behaviour as it was hard to decide whether or not a food was coming out intentionally due to dislike or uncontrollably due to a lack of experience.

For facial expressions a review was conducted to detect which action units (AUs; Rosenstein & Oster, 1988) were used to assess liking in babies. Several studies reported that facial expressions seemed more suited to measure 'dislike' but not 'like' (Mennella et al., 2001; Rosenstein & Oster, 1988; Zeinstra et al., 2009). Therefore only negative facial expressions were incorporated from the BABY FACS (Oster, 2004). The selection of cues was based on three criteria: (a) discriminative in terms of like/dislike, (b) minimum two publications have agreed on these cues as indicative of hedo-



A: Left panel – acceptance; Right panel - turns head away (Behaviour 1, B1); lower panel (B1)

B: Left panel – neutral expression; Right panel - cries/gets fussy (Behaviour 3, B3)

Fig. 1. Illustrations of infant behaviours; rate of acceptance and facial expressions. (A) Left panel – acceptance; Right panel – turns head away (Behaviour 1, B1); lower panel (B1). (B) Left panel – neutral expression; Right panel – cries/gets fussy (Behaviour 3, B3). (C) Left panel – neutral; Right panel – pushes spoon away (Behaviour 4, B4). (D) Left panel – open mouth; Right panel – learning forward, pulls spoon toward mouth (Behaviour 5; B5); Lower panel leaning forward to pull spoon toward mouth (B5). (E) – examples of early, late and enforced rates of acceptance. (D) – examples of facial expressions (nose wrinkling, brow lowering NW + BL; inner brow raising: IBR; squinting: Sq; gaping: Ga).

nic value and (c) expressions clearly detectable in video recordings. The following facial expressions were selected: inner brow raised (AU1), brow lowered (AU4), squinting (AU7 extreme), nose wrinkling (AU9), upper lip raised (AU10), lip corners down (AU15) and gaping (AU26/27) (see Table 2).

2.1.2. Behaviour sampling

The duration of coding time to assess liking has varied in previous studies (see Table 3) and research by Forestell and Mennella (2007) indicated that the first 2 min are sufficient to detect liking and disliking since satiety would be minimal during this interval. However based on published studies and the authors' observations a specific time duration may not account for individual differences (e.g. rate of eating, quantity eaten and pace of feeding), therefore the comparison between infants is challenging. In order to standardise food intake to some extent, the present coding system used only the first 9 spoonfuls (missed spoons are not counted, e.g. infant is looking away prior to the food offer and therefore misses the spoon) so that this would be uniform across sessions. These spoonfuls were paced to the infant's rate of eating and mothers were instructed to feed her infant at the customary pace, then to stop feeding after 3 refusals (as suggested by Mennella & Beauchamp, 1997). After the first mouthful, the infant responds to the taste. From the second spoonful the infant will display general behaviours expressing acceptance in reaction to the spoon approaching his/her mouth. Therefore it is assumed that the observed behaviours and facial expressions reflect the infant's experience of the food and are not attributable to other factors (such as gastrointestinal distress).

2.1.3. Coding scale and video coding

The current paper-based coding system is divided into two sections; part one codes behaviours prior to tasting food, upon the spoon approaching and includes four negative behaviours (turning head away, arches back, gets fussy, pushes spoon away), one positive behaviour (leans forward) and rate of acceptance based on mouth movements (early, late, enforced, refused). Part two included coding of seven facial expressions after the food had been tasted (nose wrinkling, upper lip raised, lip corners down, gaping, brow lowered, inner brow raised, squinting). Behaviours and facial expressions were assessed per spoonful and scored as yes/no (1/0) except the rate of acceptance which was assessed on a 4-point scale (early = 3, late = 2, enforced = 1, refused = 0). The coding system also included 3 non visible options (for mouth movement, upper face and lower face) i.e. when the visibility was obscured or obstructed. When a spoon offer was rejected, the infant's behaviours were coded for the spoon offer; however facial expressions were not coded because the infant had not tasted the food for that

C: Left panel – neutral; Right panel – pushes spoon away (Behaviour 4, B4).

D: Left panel – open mouth; Right panel – learning forward, pulls spoon toward mouth (Behaviour 5; B5); Lower panel leaning forward to pull spoon toward mouth (B5)

Fig. 1 (*continued*)

particular spoon. Codes were provided for each spoonful and a total count (frequency) at the end of the session. In addition we have a score per spoonful for each behavioural and expressive code. A complete table of results can be found in Nekitsing et al. (in press).

It is recognised that in the present context separating pre and post-tasting responses is somewhat unusual, since in the normal meal context these occur rapidly, in succession and at times in parallel. However, it was decided to code gross behaviours before food is actually tasted as a general marker of readiness to accept/reject the food and more "molecular" facial responses to the taste of food as a specific indicator of liking/disliking and acceptance/rejection.

VLC media player was used to playback the videos because the tool was easily accessible, free to download and has various visual effects such as zoom, colour/contrast, frame by frame view and slow motion. These features were particularly useful when video quality was poor, for example due to reduced brightness in the room. Video effects such as slow motion and contrast were also necessary for recognising subtle facial expressions (e.g. inner brow raised) as occurrence of these were rapid and difficult to detect at times.

2.2. Training coders

A comprehensive self-teaching e-training manual was developed detailing the description of all behaviours and facial expressions including pictures/video extracts to illustrate each behaviour/facial expressions by one of the authors (JM). The manual also included information on possible misinterpretations/false alarms. The illustrations and video extracts were added after reaching consensus between two of the authors (JM and CB). One of the authors was also trained using the FACS which was extremely important for the validation process within the self-certification cue identification (CB).

The coders utilised the training material and practiced coding all behaviours and facial expressions from extraction of selected videos. To assess if coders were sufficiently consistent to begin the actual coding a certification test was developed with reference test scores compiled with agreement of two independent authors. The test included 33 spoonful extracts (equivalent to 4 videos of 9 spoonfuls) and covered all cues in the coding scheme. The coder's scoring was compared to the reference scores using Cohen's kappa to determine the accuracy. Kappa provided a quantitative measure of the agreement level between coders and a commonly cited Kappa agreement range was used: <0 indicating less than chance agreement, 0.01–0.20 slight agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement and >0.81 indicating almost perfect/perfect agreement (Landis & Koch, 1977; Viera & Garrett, 2005). Coder's scoring should match moderate Kappa ($\varkappa$) cutoff values, preferably >0.60 and Prevalence Adjusted Bias Adjusted Kappa; PABAK >0.5 (Byrt, Bishop, & Carlin, 1993; Sim & Wright, 2005). PABAK could be used due to distribution skew (mainly non occurrence of behaviours or facial expressions). Lower scores indicated that the coder should practice further and repeat the certification test until satisfactory scores had been achieved.

2.3. Pilot study

The reliability of the coding system was assessed in a randomized controlled trial; further details of the study are available elsewhere (Hetherington et al., 2015; Nekitsing et al., in press). In

E:– examples of early, late and enforced rates of acceptance

D – examples of facial expressions (nose wrinkling, brow lowering NW+ BL; inner brow raising: IBR; squinting: Sq; gaping: Ga)

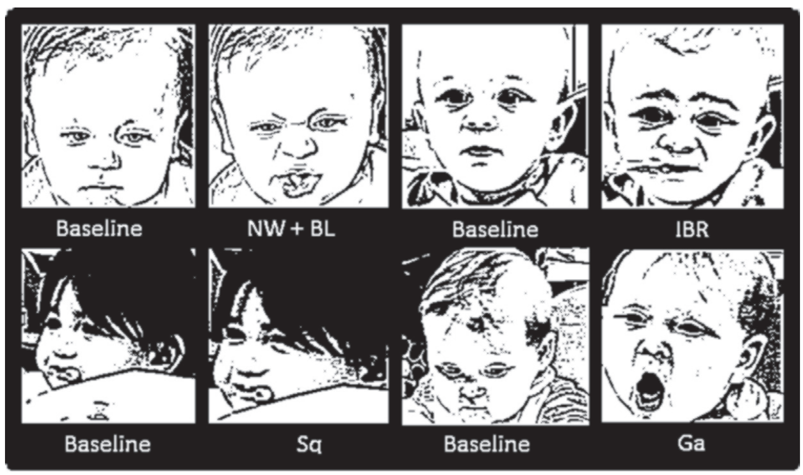


Fig. 1 (*continued*)

summary, 38 mother-infant dyads attended the laboratory on 5 occasions so that responses to vegetable puree served by mothers could be filmed in a comfortable, but standard infant laboratory. For the purposes of this study, complete data for 36 infants aged between 5 and 6 months were coded by four trained coders (2 coded behaviours, 1 coded facial expressions and 1 coded both parts), yielding 72 video episodes. If fewer than 6 spoonfuls were coded for an infant or if there was a discrepancy for non-applicable cues between 2 coders, these were excluded from the analysis. The final sample of videos coded ranged from $n = 63–68$ ($n = 68$ behaviours; $n = 64$ upper face and; $n = 63$ for lower face).

2.4. Structural validity: Psychometric analysis of factor structure and internal reliability

To assess coders' accuracy during the training phase (for the certification test) inter-rater reliability (IRR) for each cue was assessed by Cohen's kappa and PABAK.

For principle factor analysis, varimax rotation was applied, numbers of factors were set to automatic, initial communalities were squared multiple correlations (Rho) and stop conditions was by convergence = 0.0001/iteration = 50. 2 ID were removed due to missing data and the analysis was based on the average ratings.

For coding of facial expressions a minimum of 6 spoonfuls had to be coded for each infant. Data with <6 spoonfuls were removed from the analysis. Non visible data were treated as missing data during the analysis. For IRR also known as inter-observer agreement, checks were made to see if coders had coded the same spoons. Only spoons that were coded by a minimum of two coders were taken into account. If some behaviours or the facial expressions were not coded by a minimum of two coders (e.g. one coder coded behaviour or facial expressions as non visible) these variables were excluded from further analysis. The internal consistency of the scale was measured with Cronbach's Alpha-α (Cronbach, 1951). Leaning forward scores were reversed and rate

Table 2
Overview of facial expressions used in different studies as indicators of responses to food stimuli (odours/tastes).

Measurements	Forestell and Mennella (2012)	Mennella et al. (2001)	Forestell and Mennella (2007)	Zeinstra et al. (2009)	Mennella et al. (2009)	Soussignan et al. (1997)	Rosenstein and Oster (1988)
Inner brow raised (AU1)	x		x	x	x	x	x
Brow lowered (AU4) (+pulled together)* (+composite AU1 + AU4)**	x	x	x	x*	x	x**	
Squinting (AU7)+cheek raise (AU6)*	x		x		x	x*	
Nose Wrinkling (AU9)	x	x	x	x		x	x
Upper lip raised (AU10)	x	x	x	x	x	x	x
Lip corners down (AU15)				x		x	x
Gaping (AU26/27)	x	x	x	x		x	x

Table 3
Summary of video coding methods in previous research.

Source	Infants' age	Video extract length	Coding method/tool	Rater status	Inter-observer reliability	Scoring
Forestell and Mennella (2012)	6.3 ± 0.1 months $n = 92$	First 2 min	Noldus Observer; frame-by-frame; FACS	1st trained rater, certified in FACS and blind to infant's identity. 2nd rater scored facial expressions for 30 sessions	Pearson's $r = 0.85$	Frequency of facial distaste including facial expressions per spoon offer
Mennella and Beauchamp (1997)	−5.7 ± 0.2 months $n = 16$	Entire feeding episode	Noldus Observer; real time	Trained raters (unspecified), blind to experimental conditions	None reported	Frequency of spoon acceptance (4 categories; at a distance, when near, when in contact, or mouth closed) and duration of facial expressions (negative, neutral, positive)
Mennella et al. (2001)	5.7 ± 0.2 months $n = 46$	First 2 min	Real time	Trained raters (unspecified), blind to the experimental conditions	Pearson's $r > 0.80$	Frequency of negative behaviour and facial responses (e.g. head turning, gaping, brow lowering)
Forestell and Mennella (2007)	5.6 ± 0.2 months $n = 45$	First 2 min	Noldus Observer; frame by frame	1 trained rater, FACS certified, blind to the conditions	None reported	Frequency of facial expressions (distaste) per spoonful offered
Zeinstra et al. (2009) (pilot study)	5–13 years $n = 6$	First 6 s	Noldus Observer, FACS	1 FACS trained coder coded each video on 2 separate days. Coder aware of the conditions but blind to the order of the stimuli	Each video was coded twice by the same coder, however analysis for test–retest reliability are not reported	Frequency count of selected facial expressions
Dearden et al. (2009)	Time 1: 12 months, $n = 40$ Time 2:17 months $n = 51$	Entire feeding episode	DVD; frame by frame	6 coders (1 health professional and 5 students), trained by authors (achieved at least 80% agreement for each items)	Training required 80% agreement for each item to be coded however, no analyses were reported for actual video coding	Child interest in food was coded including verbalization, physical action and position of the child (e.g. leaning forward, reaching for spoon, rejection). Care givers' behaviours were also observed
Mennella et al. (2009)	6.18 ± 0.2 months $n = 97$	First 2 min	Noldus Observer; frame by frame	A FACS certified rater, blind to experimental condition	None reported	Number of facial expressions of distaste made per spoonful and rate of specific facial expressions (e.g. gape, smile)
Soussignan et al. (1997)	3.35 ± 1.5 days $n = 46$	45 s (30 s preceding, 10 s exposure and 5 s post stimulus)	Videotape; slow motion and frame by frame. Polygraphic recording BABY FACS	1 FACS certified coder (Test score >70%) blind to the stimuli. 2nd coder blind to the aim of the study and the quality of the odorant stimuli viewed 20 randomly selected videotaped segments	Spearman rho = 0.97 for mouthing behaviours but no reliability scores given for facial expressions	Regularity of respiration, eye/head/limbs movements/vocalisation and food related mouth movements (sucking, licking chewing, munching)
Rosenstein and Oster (1988)	2 h $n = 12$	2 consecutive 30 s presentations of each of the 4 solutions (separated by two 90 s water rinses)	Videotaped; slow motion, frame by frame. BABY FACS	2 certified FACS coders, blind to the stimulus conditions (2nd coder also blind to the hypothesis)	65–85% agreement for each action unit	Facial expressions and other gross behaviours such as spitting sucking head turning

of acceptance scores were reversed and recoded to fit 9 point scale. Therefore the higher the overall rate of acceptance and leaning forward, the lower the scores.

Data were corrected for non-applicable and non visible upper/lower or both part of the facial expressions by calculating the average of facial expressions occurred for the total spoon coded and then multiplied by 9 (to get overall measure for 9 spoonfuls). IRR output from both corrected and non corrected data are given below. IRR was assessed with an intra-class correlation coefficient (intra-class correlation coefficient) as this is widely used method for observational data (Hallgren, 2012; McGraw & Wong, 1996; Shrout & Fleiss, 1979). For the intra-class correlation coefficient analysis model selected was two-way mixed method with absolute agreement, confidence interval (CI) was set to 95%.

Spearman rho correlations were performed to explore relationships between all variables within the current coding scheme (average from coders' ratings were applied).

Statistics were performed using IBM SPSS (v20, Chicago, USA) except for the factor analysis for which XLSTAT version 2014.3.02 was used.

3. Results

3.1. Certification test

During the certification test, coders coded behaviours and facial expressions in comparison to a referent of each. Table 4 indicates Cohen's Kappa $\varkappa$ and PABAK score for each author's certification test (highest score if more than one certification attempt). Most scores were high in agreement. However, scores for inner brow raised, upper lip raised (for one author) and lip corners down were moderate to low. Therefore, two coders (C and D) took further practice and at least one further certification test, before coding was resumed.

3.2. Factor analysis and internal consistency

Initially factorability of 13 items including all behaviours and facial expressions was assessed. Based on eigenvalues >1.0 (Kaiser, 1960) two dimensions emerged with 37% of total variance explained. Factor values for all items were highest in either of these two dimensions, except for crying which was highest in dimension 3. After varimax rotation, factor values for all behavioural cues were highest in dimension 1 (19% of variance explained) and factor values for all FE cues were highest in dimension 2 (18% of variance explained). In other words, all behaviours were loaded onto dimension one, these included food avoidance behaviours turns head away, arches back, pushes spoon away and crying/fussy (see Fig. 2), leaning forward and rate of acceptance which are food approach behaviours also loaded on to dimension one, but inversely. All facial expressions loaded onto dimension two (see Fig. 2).

Only one factor with eigenvalue >1.0 emerged when behaviours alone were considered with 39% variance explained. All behaviours loaded on factor one, however for crying/fussiness (+ve) and leaning forward (−ve) factor loadings were low <0.4. When only facial expressions were considered, only one factor emerged with eigenvalue >1.0 explaining 35% variance (see Fig. 1).

Cronbach's alpha indicated that the scale with all 13 items, then behaviour subscale alone (6 items) and FE subscale alone (7 items) had acceptable internal consistency as indicated by α value of 0.77, 0.70 and 0.76 respectively (George & Mallery, 2003). See Table 5 for α value if item is deleted.

3.3. Inter-rater reliability; intra-class correlation coefficient

IRR was used to assess consistency between coders' ratings for behaviour and facial expressions. The resulting intra-class correlation coefficients were within an acceptable range (intra-class correlation coefficient range = 0.63–0.96, 95%CI −0.05–.0.98) for most of the variables, except for squinting (intra-class correlation coefficient = 0.57, 95%CI 0.31–0.74) and lip corners down (intra-class correlation coefficient = 0.31–0.08–0.57). This indicates that for most cues raters had similar ratings and good agreement for behaviours and for facial expressions. The cues with high intra-class correlation coefficient indicated minimal measurement error by coders and therefore were deemed to be reliable. Table 6 shows intra-class correlation coefficient with 95%CI for non-corrected and corrected data.

3.4. Test–retest reliability; intra-class correlation coefficient

Test–retest reliability (was also performed to assess consistency in coding over time. From the collection, 10% videos (7 videos) were randomly selected by a number generator and re-coded by the same author (coder D). The time duration between the two coding session was 12 weeks. The intra-class correlation coefficient indicated very high consistency for most variables (>0.81, 95%CI,

Table 4
Certification test Kappa and PABAK scores for each author.

		Cohen's Kappa ($\varkappa$)				PABAK			
Coder		A	B	C	D	A	B	C	D
Behaviour	Turns head away	0.86	0.65	0.69	–	0.88	0.76	0.76	
	Arches back	0.61	0.72	0.87	–	0.82	0.82	0.76	
	Cries/fussy	0.65	0.78	1	–	0.94	0.94	1	
	Pushes spoon away	0.87	0.80	0.87	–	0.94	0.88	0.93	
	Leans forward	0.90	0.80	0.80	–	0.94	0.88	0.94	
	Rate of acceptance	0.80	0.74	0.72	–	NA	NA	NA	
	Overall Behaviour	0.82	0.65	0.72	–	0.90	0.85	0.86	
Facial expression	Brow lowered	–	–	0.69	0.78			0.76	0.82
	Inner brow raised	–	–	0.15	0.46			0.46	0.70
	Squinting	–	–	0.63	0.78			0.88	0.94
	Nose wrinkled	–	–	0.60	0.74			0.70	0.82
	Upper lip raised	–	–	0.80	0.50			0.88	0.47
	Lip corners down	–	–	0.15	0.31			0.64	0.64
	Gaping	–	–	0.65	0.78			0.94	0.94
	Overall facial expressions [a]	–	–	0.54	0.56			0.75	0.77

– Certification test not taken for this part.
[a] For PABAK analysis non-visible cues were excluded.

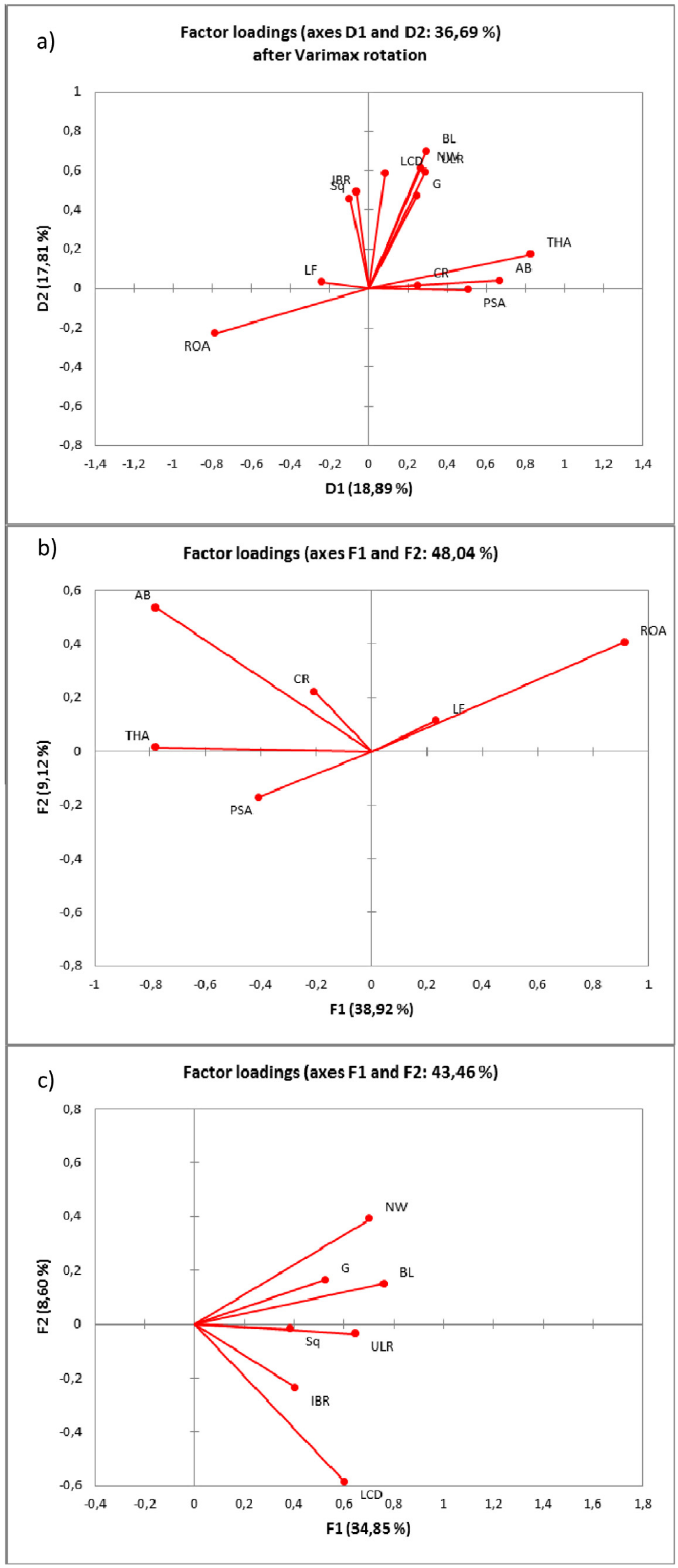


Fig. 2. Factor analysis results showing factor loadings for: (a) All factors, (b) behaviours only and (c) facial expressions only. AB: arching back, PSA: pushing spoon away, LF: leaning forward; CR: crying; ROA: rate of acceptance; IBR: inner brow raised, BL: brow lowered, Sq: squinting, NW: nose wrinkled; ULR: upper lip raised; LCD: lip corners down; G: gaping.

Table 5
Internal consistency and Cronbach's alpha if one item is deleted.

Scale	Overall scale 13 items α = 0.77	Behaviour subscale 6 items α = 0.70	Facial expressions subscale 7 items α = 0.76
Items	(*Cronbach's alpha if Item deleted*)		
Turns head away	0.74	0.58	
Arches back	0.76	0.65	
Cries/fussy	0.77	0.72	
Pushes spoon away	0.77	0.68	
Leans forward	0.79	0.73	
Rate of acceptance	0.73	0.54	
Brow lowered	0.72		0.67
Inner brow raised	0.76		0.75
Squinting	0.77		0.77
Nose wrinkled	0.75		0.72
Upper lip raised	0.73		0.68
Lip corners down	0.75		0.74
Gaping	0.75		0.74

−0.04–1.00). The test–retest reliability for arching back was moderate (intra-class correlation coefficient = 0.63, 95%CI −2.07–0.94).

As only 7 videos were coded the occurrences of some cues were infrequent. Occurrence of gaping and squinting did not match during the two time-points and the intra-class correlation coefficient score was zero (gaping intra-class correlation coefficient = 0.00, 95%CI −2.25–0.80; squinting intra-class correlation coefficient = 0.00, 95%CI −4.82–0.83). However the non occurrence of these facial expressions did match on 5 out of 7 videos (71% agreement). The discrepancy for occurrence was therefore due to the low frequency of these facial expressions. Some cues scored lower for the TRR however, the overall scores for total negative behaviours and total facial expressions were high over time, intra-class correlation coefficient = 0.96 (95% CI 0.77–0.99) and intra-class correlation coefficient = 0.97 (95% CI 0.80–0.99) respectively.

3.5. Association between behaviours and facial expressions

Associations between each cue are presented in Table 7. All significant results were in the direction expected. Thus the sum of avoidance behaviours was positively associated with the sum of all facial expressions and inversely associated with rate of acceptance. Also, some of the specific facial expressions such as brow lowered and upper lip raised as well as total facial expressions were inversely associated with rate of acceptance. This indicates that infants who expressed avoidance behaviours as the spoon was approaching were unlikely to accept the spoon later (i.e. when near to the mouth or when spoon had touched their lips) and were also likely to express more negative facial expressions after the food had been tasted. Turns head away was associated with all facial expressions and strongly associated with total negative facial expressions. Pushes spoon away was weakly associated with brow lowered but arches the back and crying/fussy were not associated with any facial expressions. Leaning forward was only associated with rate of acceptance, indicating that infants who express positive behaviours of leaning forward will accept the spoon early during the feed (i.e. opening mouth when the spoon is further away from mouth). However although leaning forward was correlated with rate of acceptance these measures may be conflated. Since infants may lean forward at the same time as opening their mouth to accept the food, it is hard to determine the added value of coding leaning forward separately from rate of acceptance. Therefore, it is proposed that only rate of acceptance is retained as an indicator of "wanting" and that leaning forward is removed from the coding scheme (see Nekitsing et al., submitted).

4. Discussion

Overall, the findings of this study indicate that the 13 item Feeding Infants: Behaviour and Facial Expression Coding System (FIBFECS) developed from existing published studies has a reasonable internal consistency, structural validity and reliability. With sufficient training the coding system can be used to explore food preferences, liking and wanting in early life. Factor analysis and measures of internal consistency confirmed structural validity. Since subjective measures can be unreliable due to bias and since indirect measures such as intake and pace of eating may be dependent on other factors such as hunger, this tool can be used to assess

Table 6
Intraclass correlation coefficients (intra-class correlation coefficient) for raw and corrected[a] data for each variable.

Intra-class correlation coefficient		Raw data	(Confidence interval)	Corrected data	(Confidence interval)
Behaviour	Turns head away	0.90	0.85–0.94	0.90	0.85–0.94
	Arches back	0.68	0.53–.80	0.68	0.53–0.80
	Cries/fussy	0.84	0.77–0.90	0.84	0.77–0.90
	Pushes spoon away	0.84	0.75–0.89	0.84	0.75–0.89
	Leans forward	0.84	0.73–0.90	0.84	0.73–0.90
	Rate of acceptance	0.96	0.94–0.98	0.96	0.94–0.98
	Total −ve behaviours	0.91	0.86–0.94	0.91	0.86–0.94
Facial expression	Brow lowered	0.71	0.01–0.87	0.64	0.05–0.84
	Inner brow raised	0.59	0.32–0.75	0.63	0.39–0.78
	Squinting	0.68	0.44–0.82	0.57	0.31–0.74
	Nose wrinkled	0.78	0.53–0.88	0.71	0.42–0.84
	Upper lip raised	0.75	0.58–0.85	0.74	0.57–0.84
	Lip corners down	0.28	−0.13–0.55	0.31	−0.08–0.57
	Gaping (G)	0.84	0.83–0.90	0.69	0.49–0.82
	Overall FE	0.91	0.84–94	0.86	0.28–0.95

[a] Data were corrected for non-applicable and non-visible upper/lower or both part of the facial expressions, see Section 2.3.

Table 7
Spearman's rho correlations between variables (behaviours and facial expressions).

	Turns head away	AB	Crying	PSA	Leaning forward	Rate of acceptance	Brow lowered	Inner brow raised	Squinting	Nose wrinkled	Lip corners down	Upper lip raised	G	Total −ve Bs
AB	.23													
Crying	.13	**.29***												
PSA	**.27***	.21	.07											
Leaning forward	.00	−.17	.11	.04										
Rate of acceptance	**−.50****	**−.49****	−.14	**−.48****	**.33****									
Brow lowered	**.79****	.17	.10	**.26***	−.05	**−.35****								
Inner brow raised	**.49****	.03	.23	−.12	.10	−.18	.31*							
Squinting	**.36****	−.08	.15	−.11	.15	−.03	.21	.22						
Nose wrinkled	**.69****	.13	.14	.21	.02	−.25	**.62****	.11	**.32***					
Lip corners down	**.60****	.24	−.06	−.03	−.01	−.22	**.38****	**.34****	.15	.22				
Upper lip raised	**.81****	.16	.06	.19	−.07	**−.40****	**.51****	**.27***	**.25***	**.49****	**.51****			
G	**.53****	.16	.06	.19	−.07	−.21	**.46****	.19	.18	**.44****	.16	**.48****		
Total −ve Bs	**.38****	**.76****	**.30***	**.51****	−.16	**−.75****	**.30***	−.03	−.08	**.26***	.11	**.30***	**.31***	
Total −ve FE	**.97****	.19	.15	.22	.01	**−.44****	**.83****	**.49****	**.40****	**.74****	**.55****	**.80****	**.60****	**.33****

* $p < 0.05$.
** $p < 0.01$.

liking in infants. However, two parameters occurred infrequently and did not appear to provide any further discriminatory ability, namely leaning forward and crying/fussiness, these can be removed from the scale for future use.

It was expected that facial expressions in response to food intake would provide reliable and valid information on liking. Facial expressions at the beginning of the meal when food is first tasted are considered indicative of like/dislike in pre-verbal infants (Mennella et al., 2001; Rosenstein & Oster, 1988; Soussignan et al., 1997; Steiner, Glaser, Hawilo, & Berridge, 2001). The work of pioneers such as Steiner (1977) and Rosenstein and Oster (1988) demonstrated that for sweet and umami tastes infants' express relaxed facial responses, begin to suckle, elicit tongue protrusions and they may even smile. Conversely, for bitter and sour tastes which are innately rejected, a negative facial reaction is displayed (e.g. aversion expressed with a nose wrinkle). Other studies have since confirmed that basic tastes and odours can be discriminated by infants (Mennella et al., 2001; Soussignan et al., 1997). Each infant may have differing responses to foods, nevertheless facial expressions appear to be robust when offered intense stimuli, but more subtle yet reliable with other tastants. Further exploration could be used to develop a taxonomy of liking and disliking in infant facial response in the same way that this has been done for adult emotions.

Currently the FACS is not aimed at detecting feeding expressions and does not seek to explain the potential relationship between avoidance behaviours and like/dislike. Therefore, the tool developed here with the combination of both behaviours and facial expressions can be useful to professionals and carers in recognising liking and wanting in infancy. Recognising these cues in conjunction may help mothers to facilitate healthy eating and responsive feeding. For example if readiness to eat indicated by rate of acceptance is recognised independently from cues of distaste, mothers can continue to offer "disliked" foods such as green, leafy vegetables to their infants. It has also been suggested previously by Forestell and Mennella (2007) that mothers often hesitate to continue feeding foods to their infant which they consider are disliked. This study encouraged mothers not to rely solely on indicators of distaste but also focus on infants' willingness to accept the food. The present tool can provide caregivers with information on both infant's willingness to eat (behaviours and rate of acceptance) and measures of like/dislike (facial expressions).

The weaning period around 6 months is crucial for establishing healthy eating habits (Harris, Thomas, & Booth, 1990). This period has been observed as a window of opportunity for taste exposure and learning, as infants may learn to accept fruit and vegetables which are disliked (Maier, Chabanet, Schaal, Issanchou, & Leathwood, 2007). Studies have shown that methods such as repeated exposure can provide flavour experience and enhance liking (Cooke et al., 2004; Forestell & Mennella, 2007; Hetherington et al., 2015; Maier et al., 2007; Sullivan & Birch, 1994) and mothers are more likely to try feeding novel foods if infants are willing to accept these (Mennella & Beauchamp, 1997). Researchers may also use this tool to measure liking over time during repeated exposures to observe the ways in which indications of dislike decline with experience and signals of the willingness to accept increase. The tool can also be helpful to professionals for product sensory analysis in infants.

4.1. Limitations and future directions

On occasion facial expressions were not easy to code due to video quality and specificities of the subtle facial expressions, such as lip corners down. When visibility is poor or obstructed during video coding, the option to code the spoon as non visible whilst useful also produced discrepancies between raters. This then affected the adjusted scores for the individual facial expressions. Thus the inter-rater reliability was lowered for certain cues such as squinting and inner brow raised. Another limitation of the present study is that test–retest was only performed on 10% of the videos and over a 12 week period. As occurrence of some cues such as gaping are infrequent coding further videos may have provided higher reliability scores. The training for coding and certification was time consuming. However, the training was devised to enable the scorers to get sufficient technical skills in order detect cues accurately without the need for FACS certification. To our knowledge previous studies have not developed such a comprehensive, referent-based training and certification test for coding behaviours and facial expressions during feeding sessions.

It is recognised that on their own, the action units do not convey hedonic valence, rather these can only be interpreted in the context of the meal, thus a gaping mouth before a food is tasted is interpreted as readiness to eat (an approach behaviour) but in response to a food already tasted, gaping indicates disgust (an avoidance behaviour). It is suggested that the coding system as described is limited to the context of the spoon-fed meal but further testing in different contexts is warranted to assess its generalisability.

The FIBFECS as developed can be used by researchers and others with an interest in early infant feeding. It permits simultaneous coding by multiple coders and uses freely available media tools, so other than the cost of the webcam or video recording equipment, it is relatively inexpensive. Studies using video coding will benefit by ensuring the room settings such as position of mother-infant and camcorder angle/focus are adequate for detecting facial expressions. In the present study mothers occasionally obscured their infant's lower face during spoon feeding. However, the need to standardise the feeding episode to allow unobscured viewing must be balanced against the need for mothers to feed without too many constraints.

Research in adults suggests that liking and wanting can be dissociated and the FIBFECS needs further investigation for assessing elements of each in infants to determine dissociability of these constructs. Our future aim is to validate the tool by investigating feeding behaviours and facial expressions under different conditions of foods offered and experience. The tool can also be developed further to observe maternal behaviour to assess mealtime interactions (Bergmeier, Skouteris, & Hetherington, 2014). Future studies may test and use only part of the coding system or test the tool with a different age range.

4.2. Conclusions

The Feeding Infants: Behaviour and Facial Expression Coding System (FIBFECS) can be useful for researchers and professionals to understand behaviours and facial expressions which infants may communicate during mealtime. Research on food refusal in infants is increasing due to the growing concern about both food fussiness and identifying problem eating early in life. Knowledge of refusals, aversion behaviours and facial expressions can be the early formation of healthy eating habits during the weaning period. As mothers are less likely to offer foods which are disliked by their infants, including green vegetables, if mothers can recognise the different elements of the infants' responses, they will be more willing to offer disliked but healthy foods to their infants. Clearly, it is the mother's willingness to offer these foods early in life which will predict later food preferences.

Acknowledgements

Funding received through the European Community's Seventh Framework Programme (FP7/2007-2013) under the IAPP 230637 "VIVA: V is for Vegetable – Applying Learning theory to increase liking and intake of vegetables"; coordinated by Professor Marion M Hetherington. http://www.vivacongress2013.co.uk.

The certification and training manuals are available as a hyperlink from the authors. Please contact Professor Marion M Hetherington (m.hetherington@leeds.ac.uk) for access and permissions.

References

Bere, E., & Klepp, K. (2005). Changes in accessibility and preferences predict children's future fruit and vegetable intake. *International Journal of Behavioral Nutrition and Physical Activity, 2*, 15.

Bergmeier, H., Skouteris, H., & Hetherington, M. (2014). A systematic research review of observational approaches used to evaluate mother–child mealtime interactions during preschool years. *American Journal of Clinical Nutrition, 101*(1), 7–15.

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology, 46*(5), 423–429.

Caton, S. J., Blundell, P., Ahern, S. M., Nekitsing, C., Olsen, A., & Moller, P. ... (2014) Learning to eat vegetables in early life: The role of timing, age and individual eating traits, (2014). *PLoS ONE, 9*(5).

Cooke, L. J., Wardle, J., Gibson, E. L., Sapochnik, M., Sheiham, A., & Lawson, M. (2004). Demographic, familial and trait predictors of fruit and vegetable consumption by pre-school children. *Public Health Nutrition, 7*(2), 295–302.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.

Dearden, K. A., Hilton, S., Bentley, M. E., Caulfield, L. E., Wilde, C., Ha, P. B., & Marsh, D. (2009). Caregiver verbal encouragement increases food acceptance among Vietnamese toddlers. *The Journal of Nutrition, 139*(7), 1387–1392.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., & Oster, H. (1979). Facial expressions of emotion. *Annual Review of Psychology, 30*(1), 527–554.

Fildes, A., & Cooke, L. (2012). The munch bunch: Healthy habits start at weaning. *Journal of Family Health Care, 22*, 30–32.

Forestell, C. A., & Mennella, J. A. (2007). Early determinants of fruit and vegetable acceptance. *Pediatrics, 120*(6), 1247–1254.

Forestell, C. A., & Mennella, J. A. (2012). More than just a pretty face. The relationship between infant's temperament, food acceptance, and mothers' perceptions of their enjoyment of food. *Appetite, 58*(3), 1136–1142.

George, D., & Mallery, P. (2003). *SPSS for windows step by step: A simple guide and reference* (4th ed.). Boston: Allyn & Bacon.

Gibson, E. L., Wardle, J., & Watts, C. J. (1998). Fruit and vegetable consumption, nutritional knowledge and beliefs in mothers and children. *Appetite, 31*(2), 205–228.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*(1), 23–24.

Harris, G., Thomas, A., & Booth, D. A. (1990). Development of salt taste in infancy. *Developmental Psychology, 26*(4), 534–538.

Hetherington, M. M., Schwartz, C., Madrelle, J., Croden, F., Nekitsing, C., Vereijken, C. M. J. L., & Weenen, H. (2015). A step-by-step introduction to vegetables at the beginning of complementary feeding: The effects of early and repeated exposure. *Appetite, 84*, 280–290.

Hodges, E. A., Hughes, S. O., Hopkinson, J., & Fisher, J. O. (2008). Maternal decisions about the initiation and termination of infant feeding. *Appetite, 50*, 333–339.

Kaiser, H. F. (1960). An index of factorial simplicity. *Psychometrica, 39*, 31–36.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Maier, A., Chabanet, C., Schaal, B., Issanchou, S., & Leathwood, P. (2007). Effects of repeated exposure on acceptance of initially disliked vegetables in 7-month old infants. *Food Quality and Preference, 18*(8), 1023–1032.

Maier, A. S., Chabanet, C., Schaal, B., Leathwood, P. D., & Issanchou, S. N. (2008). Breastfeeding and experience with variety early in weaning increase infants' acceptance of new foods for up to two months. *Clinical Nutrition, 27*(6), 849–857.

McGraw, K. O., & Wong, S. P. (1996). Inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30–46.

Mennella, J., & Beauchamp, G. (1997). Mothers' milk enhances the acceptance of cereal during weaning. *Pediatric Research, 41*(188–192).

Mennella, J. A., Forestell, C. A., Morgan, L. K., & Beauchamp, G. K. (2009). Early milk feeding influences taste acceptance and liking during infancy. *American Journal of Clinical Nutrition, 90*(3), 780S–788S.

Mennella, J. A., Jagnow, C. P., & Beauchamp, G. K. (2001). Prenatal and postnatal flavor learning by human infants. *Pediatrics, 107*(6), e88.

NDDOH (2000). North Dakota Department of Health (www.health.state.nd.us): Infant care, baby facts sheets: Feeding cues. Available from: <http://www.ndhealth.gov/familyhealth/MCH/babyfacts/FeedingCues.pdf>.

Nekitsing, C., Madrelle, J., Barends, C., de Graaf, C., Morgan, S., Parrot, H., Weenen, H., Hetherington, M. M. (in press). Application and validation of the Feeding Infants: Behaviour and Facial Expression Coding System (FIBFECS) to assess liking and wanting in infants at the time of complementary feeding. Food Quality and Preference. http://dx.doi.org/10.1016/j.foodqual.2015.09.012.

Olsen, A., Ritz, C., Kraaij, L. W., & Møller, P. (2012). Children's liking and intake of vegetables: A school-based intervention study. *Food Quality and Preference, 23*(2), 90–98.

Oster, H. (2004). *Baby FACS: Facial action coding system for infants and young children*. New York: New York University.

Rosenstein, D., & Oster, H. (1988). Differential facial responses to four basic tastes in newborns. *Child Development, 59*(6), 1555–1568.

Schwartz, C., Issanchou, S., & Nicklaus, S. (2009). Developmental changes in the acceptance of the five basic tastes in the first year of life. *British Journal of Nutrition, 102*(9), 1375–1385.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing reliability. *Psychological Bulletin, 86*(420–428).

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy, 85*(3), 257–268.

Soussignan, R., Schaal, B., Marlier, L., & Jiang, T. (1997). Facial and autonomic responses to biological and artificial olfactory stimuli in human neonates: Re-examining early hedonic discrimination of odors. *Physiology & Behavior, 62*(4), 745–758.

Steiner, J. E., Glaser, D., Hawilo, M. E., & Berridge, K. C. (2001). Comparative expression of hedonic impact: Affective reactions to taste by human infants and other primates. *Neuroscience & Biobehavioral Reviews, 25*(1), 53–74.

Steiner, J. E. (1977). Facial expressions of the neonate infant indicating the hedonics of food-related chemical stimuli. In J. M. Weiffenbach (Ed.), *Taste and development: The genesis of sweet preference*. Washington, DC: U.S. Government Printing Office.

Sullivan, S. A., & Birch, L. L. (1994). Infant dietary experience and acceptance of solid foods. *Pediatrics, 93*(2), 271–277.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine, 37*(5), 360–363.

WIC (Women Infant Children) (2007). WIC infant nutrition and feeding guide. WIC learning online: Infant hunger and satiety cues. Available: <http://www.nal.usda.gov/wicworks/WIC_Learning_Online/support/job_aids/cues.pdf>.

Zeinstra, G. G., Koelen, M. A., Colindres, D., Kok, F. J., & de Graaf, C. (2009). Facial expressions in school-aged children are a good indicator of 'dislikes', but not of 'likes'. *Food Quality and Preference, 20*(8), 620–624.