

Thin Slice Sampling of Video Footage for Mother/Child Interaction: Application to Single Cases

Deborah M. James · Meghana B. Wadnerkar ·
Christa Lam-Cassettari · Sujin Kang · Anna L. Telling

© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract The purpose was to test the reliability of short samples of parent/child interaction for use in single-subject research. Four variable pairs of mother/child behaviour were coded for seven mother/child play sessions. Each session lasted 20 min and 18 min of the session was behaviourally coded using frame-by-frame analysis. The co-occurrence of the mother/child behaviours within a given time window was computed and an odds ratio was calculated for the co-occurrence of the targeted behaviours. The play session was divided into shorter segments (3, 6 and 9 min) and odds ratios of the variable pairs from the shorter segments were compared to the odds ratios from the entire session. Segments of 3 and 6 min did not yield the same pattern of results as the entire session. In single-subject research, evidence of the reliability of the time segment for behavioural coding should be reported in the methods section of original research manuscripts.

Keywords Video · Mother child interaction · Microanalysis · Contingency · Sequential Analysis · Thin slice sampling

Behavioural coding of live or recorded observations is often used as a tool to capture the intricacies of human interaction

D. M. James (✉)
Health, Community and Education Studies, Coach Lane Campus,
Northumbria University,
Coach Lane, Benton,
Newcastle upon Tyne NE7 7XA, UK
e-mail: deborah.james@northumbria.ac.uk

D. M. James · M. B. Wadnerkar · C. Lam-Cassettari · S. Kang ·
A. L. Telling
NIHR Biomedical Research Unit in Hearing,
University of Nottingham,
Nottingham, UK

in a variety of fields for experimental and clinical purposes. It is an accepted practice to code a proportion of the observations or ‘thin-slices’ of interaction as a representation of the whole interaction. The term ‘thin slicing’ was coined by Ambady and Rosenthal (1992, 1993). They were interested in exploring the phenomena that in real-life situations people could make accurate assessments of other people’s personality characteristics based on only a few moments of observing or engaging in interpersonal interaction. Their meta-analysis (Ambady and Rosenthal 1992) showed that the ability to classify people according to closed set character traits was not affected by the length of the sample of interaction that was observed and predictions based on samples of less than 5 min in duration yielded significant effect sizes based on the accuracy of classification of character.

Thin-slicing can be analogues to brief functional assessments that are often used in psychiatry. Brief assessments are known to be reliable, but the results they provide can be limited to behaviours that occur more frequently (Derby et al. 1992) with some studies showing lower accuracy from brief assessments when compared to extended assessments (Roger et al. 1990). Reliability of behavioural observation could also be related to the frequency of occurrence of behaviour so the duration of the observation window could have an impact on reliability. We conducted a review of research studies that used micro-analytic methods to code interaction behaviour between parents and children. There were 38 published studies that matched our search criteria in the previous 20 years. Eighteen of those studies used observation windows of 5 min or less for coding and only 12 studies contained a justification for the duration of the observation window (see Table 1). Of these studies, the monograph of Jaffe and colleagues (2001) provided the highest degree of detail and they cite the meta-analysis by

Table 1 Micro-analytic studies of parent–child interaction

Authors	Year	Outcomes measured	Time slice	Justification for time slice duration
Beebe, et al.,	2007	Mother–infant self- and interactive contingencies	First 2.5 usable mins	Followed previous research
Beebe, et al.,	2010	Mother–infant interaction at 4 months	First 2.5 usable mins	Followed previous research
Bornstein	2002	Continuous versus time-sampled observations of mother–infant interaction	Whole video versus time-sampled 15, 30 and 45 s slices	Followed previous research
Feldman	2006	Mother–infant social synchrony in face-to-face play	Middle 3×1 min	Followed their own previous research
Feldman, Greenbaum & Yirmiya	1999	Effect of affect-synchrony on infant self-control	First 3 min, (the initial minute was excluded as an adjustment period)	Pilot study, $n=5$
Feldman, Greenbaum, Yirmiya & Mayes	1996	Relationship between interactive rhythms and cognitive development	First 3 min, (the initial minute was excluded as an adjustment period)	Pilot study, $n=5$
Friedman, Beebe, Jaffe, Ross & Triggs	2010	Microanalysis of infant vocal affect with depressed mothers	First usable 2.5 min	Followed previous research
Jaffe	2001	Rhythmic coupling and bi-directional coordination in mother–infant interaction	Time-sampled 12×1 min	Followed previous research
Ray & Tickle-Degnan	2004	Coding scales for video-taped mother–infant interactions	Time-sampled 5×60s and 5×30 s	Pilot study
Rodriguez et al.,	2010	Effect of Healthy Families New York in promoting parenting competency	M duration of task interval	To suit study design and aims
Vigil	2002	Cultural variation in infant attention regulation	Middle 5 min	Previous research and to suit study design and aims
Weiss, Wilson, Hertenstein & Campos	2000	Tactile contact during mother–infant interactions	Middle + final 5 min	Followed previous research

Ambady and Rosenthal (1992) as the basis for their justification of ‘thin slice’ sampling. This raises the question of how thin the slice of interaction needs to be in order to accurately represent the parent/child interaction. This is especially pertinent in the context of mother/child play where the frequency of indicative behaviour is likely to vary widely with context. Prior research suggests that data based on samples of parent/child interaction selected from the beginning of a play session is more representative than samples from the middle or end of the play session (e.g., Beebe et al. 2010; Friedman et al. 2010), but these conclusions were drawn from largely impressionistic analysis rather than comparison of behavioral coding from different portions of the play session.

In the wider arena of behavioural research detailed methodology on the type of sampling has been reported. In a comparison of continuous sampling versus timed-sampling techniques, Bornstein (2002) showed that the patterns of the data derived from the two alternative methods were identical, or identical enough to render any difference to fall

within acceptable limits for researchers. The two methods in Bornstein’s study require different amounts of resource. Time-sampling requires less resource than continuous sampling. The resource differential between coding methods highlights the way that the practicalities required for micro-analytic behavioural coding interacts with the choice of sampling methodology. In clinical settings the use of video footage for assessment purposes is increasing (Lord et al. 1989, 2000; Tait et al. 2001) and at the same time it is likely that the resources available for micro analysis of that footage is decreasing. Clinical scientists designing new applications using micro-analysis of behavioural data might be inclined towards thin-slice coding since it is quick and there is no evidence to question its reliability from experimental research.

The interaction between parents and children carries critical determinants for infant socio-cognitive development (Belsky and Fearon 2002; Bernier et al. 2001; Biringen et al. 2005; Bornstein 1985; Feldman and Eidelman 2009; Trevarthen and Aitken 2001). Maternal behaviours such as

the expression of affect (Moore et al. 1997), eye-gaze towards the child (Stern 1983), touch (Tronick 1995), and vocalisation (Hsu and Lavelli 2005) have been found to be predictive of a range of infant outcomes (Feldman et al. 1999, 2004b; Gable and Isabella 1992; Hsu and Fogel 2001; Hsu et al. 2000; Hsu and Lavelli 2005; Moore et al. 1997; Murray and Hornbaker 1997; Treyvaud et al. 2009). Interactional parental behaviour has been found to vary according to culture (Feldman et al. 2006; Hsu and Lavelli 2005), presence of maternal postnatal depression (Milgrom et al. 2004) and the status of the sibling/s within the family unit (Feldman et al. 2004a; Moore et al. 1997). Research has shown that it is not just the type and frequency of the parents' early interactional behaviour that is important, but the contingency of parental behaviours is also crucial for the child's development (Murray and Trevarthen 1985).

Contingency of parental behaviour can be understood as the sequence and timing of parent's behaviours within which the parent responds to the child's cues. It is considered to be a property of a wider parenting construct referred to as maternal responsiveness or maternal sensitivity (Shin et al. 2006). In the same way that interactional behaviour varies with context, contingency of parental behaviour is also subject to variation. In experimental research depression in mothers has been found to influence the contingency of the mother's responses during interaction exchanges with infants. The measurement of the contingency of behaviours during dyadic interaction is time intensive, but its current status within theoretical models of development (Jaffe et al. 2001) and its use in the design and evaluation of intervention programmes for parents (Rodriguez et al. 2010) means that the coding of contingency during interaction will continue for some time to come.

In the implementation of an exploratory phase 1 clinical trial of a parent intervention programme in the context of childhood hearing impairment, we chose to use a measure of parental contingency to evaluate change before and after the intervention. This led us to question, how much parent/child interaction we needed to code to get a reliable and representative measurement of contingency. The reliability of the data was particularly important since we planned to use the measure to assess change over time in single families. The analogous application of Ambady and Rosenthal's study to the field of developmental psychology might be articulated as such; parental contingency is an observable and measurable expressed behaviour; could brief observations of this behaviour be used to accurately judge the trait of parental responsiveness.

In the larger study for which we were defining our methodology, our intention was to test an experimental hypothesis which stated that the parent intervention in the clinical trial would enhance parental responsiveness. This hypothesis drove the need to develop a robust measure of

parental contingency which examined the sequential occurrence of the parents' behaviour in response to the child's. The dependent variable was for use in a phase 1 clinical trial with single-subject design and there was a need to test the reliability of the measurement for use with single cases. It was not enough to adopt methods that had been derived from pooled data from experimental research. So there remained a question as to whether the coding of a thin slice (5 min or less) of behaviour would yield results similar to that provided from a complete play session of 20 min duration. The practical question of how much data we needed to code to get a reliable result meant that we compared three partial segments that differed in duration, one time slice that was below 5 min duration (3 min segments—referred to as *short duration*), one time slice that was above 5 min duration (6 min—referred to as *medium duration*) and one time slice that represented half of the entire play session (9 min—referred to as *long duration*). The main experimental hypothesis of the study is that short slices of behavioural data (less than 5 min in duration) taken from naturalistic free play between mothers and children is not as reliable as longer slices of behavioural data (medium duration and long duration). A secondary experimental hypothesis of the study is that the frequency of behaviour, based on a count from the population of the entire free-play session, interacts with the reliability of data from the differing time samples. Low frequency contingency pairs will be less reliable in all the samples, but particularly in the short samples compared to higher frequency contingency pairs. A final experimental hypothesis of the study is that results based on the samples taken from the beginning of the entire play session will be more reliable than samples taken at the end of the play session. This pattern will be observed for all the sample durations (short, medium and long).

The analysis will be based on assessment of the likelihood ratios of behaviours of mother and child co-occurring within a specific time window on both case-by-case basis and grouped data. This study is based on a large scale intervention study looking at mother-child interaction and child speech development where the child is Deaf/deaf/hard of hearing. As part of a larger assessment session testing a range of psychological and acoustic speech measures associated with child development and parenting, a mother and her Deaf/deaf/hard of hearing child were video recorded while engaged in a naturalistic free play interaction.

Method

Participants

This study uses data from four mother-child dyads who were taking part in the aforementioned intervention study.

There were two boys (ages 3;10 and 6;10 years) and two girls (ages 5;3 and 1;11 years). All mothers were hearing and all children were congenitally profoundly deaf. Two children had bilateral cochlear implants and were pre-lingual. One child (girl 5;3 years) had bilateral cochlear implants, and had some level of speech. One child had a unilateral cochlear implant and was pre-lingual with severe learning difficulties (boy aged 6;10 years).

For the two girls, their father and sibling were also present in the room to keep the sessions as representative of the family as possible. However, the presence of additional interactive partners made the behavioural coding more complex than traditional face-to-face microanalytic studies. For coding purposes only interaction between the mother and her deaf child was coded. Two families had one pre-intervention visit each and two families had two visits each (pre and post intervention baseline) following the same procedures (family 1—one visit, family 2—three visits, family 3—two visits and family 4—one visit). Since the focus of this study is the methodology of choosing time segments for behavioural analysis, the intervention status of the families was not considered as part of the analysis. This gave a total of seven video cases for analyses in this paper.

Informed, signed consent was obtained from the parent/caregiver prior to starting the study. The study has been reviewed and approved by the Derbyshire Research Ethics Committee and Nottingham University Hospitals NHS Trust Research and Development department. All families were reported to be in good health at the time of filming.

Recording and Editing

Families were video recorded during an unstructured¹ free play interaction in the family lab at NIHR National Biomedical Research Unit in Hearing. The lab resembles a family lounge and consists of two lounges, a child's desk and chair, and a variety of toys to support play sessions. To minimise the obtrusiveness of the equipment, recordings were taken via three wall mounted hard-drive video cameras that were disguised with fluffy toys so that only the lens was showing. The mothers were instructed to play and spend time with their child as they normally would at home, using any of the toys available for 20–25 min. They were allowed to take a break if required. Following the test session, each camera file was converted into a 20 min movie in MPEG4 format to ensure each camera angle was 25 frames per second.

¹ Unlike a face-to-face interaction or structured play the mother–child pair is free to choose their play activity and move about in the room.

Behavioural Coding

Video recordings were coded as a continuous variable in INTERACT software (Mangold 2008) to extract the frequency, onset, offset and duration of each mother–child dyads' gaze and verbal behaviours. Videos were coded frame-by-frame in the format 25 frames per second. Coding was shared by two coders. Each mother–child dyad was coded independently for each behaviour, giving four codes per dyad: eye gaze of the child towards the mother coded as 'child look'; child's vocalisations coded as 'child verbal'; eye gaze of the mother towards the child coded as 'mother look' and mother's speech coded as 'mother verbal'. These behaviours were selected in line with the hypotheses for the larger study which are centred on maternal attunement as measured by eye gaze and speech. The larger study seeks to understand the effect of an intervention on mother's contingent responses to the child. Hence maternal behaviours are defined as 'Target' behaviours and child behaviours are defined as 'Given' behaviours (Bakeman and Quera 2011b). A third coder blind coded 20% of randomly selected footage from each 18 min file for each of the seven cases to calculate inter rater agreement. This equated to 3.6 continuous minutes per video. There was a good agreement between the coders as indicated by the mean $\kappa=0.85$ for overall child behaviours (child look $\kappa=0.96$ and child verbal $\kappa=0.92$) and for overall mother behaviours $\kappa=0.76$ (mother look $\kappa=0.97$ and mother verbal $\kappa=0.95$).

To account for the time taken by the participants to settle in the activity the first and last 1 min was discounted from the 20 min free play video while coding, thus leaving 18 min of free play to code. In the first instance the whole 18 min was coded for each of the seven videos. Each 18 min code file was then divided into six slices that were 3 min in duration (first 3 min, followed by the second 3 min etc.). In addition the 18 min of coding was divided into two segments of 9 min duration and three segment of 6 min duration. Each coded segment is referred to as a dataset in the analysis and results section.

Data Extraction

The coding software INTERACT stores data sets as tab-delimited text files (.ACT files). The .ACT files were extracted in ActSds using 1 s time units (Bakeman and Quera 2009) and converted into Sequential Data Interchange Standard (SDS), a format used for sequential data analysis (Bakeman and Quera 1992, 2008). These .SDS files were opened and modified for time-event sequential analysis in Generalized Sequential Querier software (Bakeman and Quera 1995, 2008, 2011a). If required, this data can then be transferred into any standard statistical programme

for further analysis or modelling. See Bakeman and Quera (2011b) for details on sequential analysis.

Data Analyses

In the first instance, observed raw frequencies were computed on the raw data for the child looking and verbal behaviours and mother's looking and verbal behaviours for each of the time slices (i.e., 18 min, 9 min, 6 min and 3 min) for each of the seven video cases in GSEQ.

In order to analyse whether the mother's looking and verbal behaviours (Target behaviours) were contingent within 3 s of the child's looking and vocalisations (Given behaviours), the WINDOW command in GSEQ was used to define four new codes: One coded the onset (first unit) second of the Given behaviour and the following 2 s as a Given window for the child's looking and vocalisations separately, while the other two coded just the onset of the mother's looking and verbal behaviours separately (Bakeman and Quera 2011b). The four behaviours with modified time windows provided the opportunity to perform sequential analysis in GSEQ and were tallied in a 2×2 table for each behaviour pair (child looking with mother looking, child looking with mother verbal, child verbal with mother looking, and child verbal with mother verbal) for each of the time slices (i.e., 18 min, 9 min, 6 min and 3 min) for each of the seven video cases. Contingency indices (Joint Frequency, and Logg odds ratio) were computed for these (see Chorney et al. 2010 for further explanation of methodology for sequential analysis).

Results from the raw data and sequential analysis were transferred into SPSS (SPSS 2007) and Microsoft Excel (2007). Results were analysed at a descriptive level using mean and standard errors of raw data frequencies, joint frequencies and contingency indexes (Logg odds) for each behaviour pair across the seven cases to examine whether the time slices provided different results from that derived from the whole 18 min free play session for each behaviour pair. Due to limited number of cases (seven cases) a modelling approach such as Repeated Measures ANOVA or Generalized Linear Mixed Models with Logg odds (since distribution of data with odds ratio was skewed) could not be considered. In any case, our focus on the reliability for individual case data meant that we were largely concerned with pattern analysis on a case-by-case basis.

Results

Contingencies were computed of mother's behaviour in response to the child's behaviour. Mother behaviours were re-coded in GSEQ to follow the child behaviours within a 3 s time window. Due to the small sample size and odds

ratio being skewed, the Logg odds ratio was computed as the contingency indices. The odds ratio is a measure of effect size and gives the direction of association. The Logg odds ratio is the natural logarithm of the odds ratio and varies from negative infinity to positive infinity, with zero as no effect. A Logg odds ratio of greater than zero indicates that the *target* behaviour is more likely in the presence of the *given* behaviour and a value smaller than zero indicates the opposite (Bakeman and Quera 2011b). For example the parent is more likely to speak (target) within the three second time window after the child has made a verbalisation (given). The data from all seven cases were analysed using the Logg odds ratio and are presented as single cases and as grouped data.

The data for the single cases are presented in Table 2. To assess the reliability of the short samples all results for each variable pair that differed in Logg odds ratio across the zero boundary were identified. Any value in the opposite direction of the Logg odds was defined as a discrepancy. If the Logg odds value was positive, based on analysis from the entire play session for a particular dyad, then a negative value was considered as discrepant and vice versa. These discrepancies are highlighted in Table 2. In all seven cases the shorter segments (3 or 6 min) yielded Logg odds ratios that differed from the Logg odds derived from the whole play session across the zero boundary. Values for the longer duration of 9 min, especially for the first 9 min, closely matched values for 18 min.

The second research question concerned the reliability of low frequency pairs. Low frequency contingency pairs were identified using descriptive statistics from the data from the entire play session. The lowest frequency pair for each dyad is marked with an asterisk in Table 2. The data suggest that there was a higher likelihood of the low frequency contingency pair being uncomputable and this is especially the case in the shorter segments (labelled as 'undef'). However, other pairs, that were not low in frequency were also left undefined in the computer analysis, again, most notably in the short segments.

The final question concerned the time of the play session when the segment was sampled (beginning of session, middle or end). The data from the single case analysis is inconclusive in this regard. There is no evident difference in the degree of discrepancy when the segments were sampled from the beginning, middle or end of the play session from the data in Table 2.

We set out to test the primary research question with the grouped data from the seven dyads. This was done by comparing the mean of the Logg odds ratios from the short segments with the mean and standard deviation of the entire play session. Z-scores were computed for the Logg odds derived for each variable pair from each of the short segments. The z-score is an indication of the probability of

Table 2 Logg odds ratio for the case by case data for each time slice (3 min, 6 min, 9 min) and entire play session (18 min). Pair1: Child look & mum look; Pair 2: Child look & mum verbal; Pair 3: Child verbal & mum look; Pair 4: Child verbal & mum verbal. Undef: when

there were no co-occurrences of contingent child and mother behaviours the Logg odds could not be computed and are termed as undefined. Low frequency variable pairs are marked *; High frequency variable pairs are marked in bold in the final column

#	Variable pairs	Time segments												
		3min-1	3min-2	3min-3	3min-4	3min-5	3min-6	6min-1	6min-2	6min-3	9min-1	9min-2	18mins	
1	Family1 visit one	Pair 1	-0.01	0.1	1.05	undef	-0.58	-1	0.38	1.44	0.35	0.58	0.71	0.65
		Pair 2	0.92	0.77	2	1.05	1.83	0.85	0.51	0.59	0.39	0.48	0.42	0.48
		Pair 3*	-0.01	0.1	1.05	undef	-0.58	-1	-0.11	1.22	-0.76	0.26	-0.36	0.16*
		Pair 4	0.92	0.77	2	1.05	1.83	0.85	0.82	1.48	1.43	1.16	1.22	1.17
2	Family2 visit one	Pair 1	1.24	undef	undef	0.86	1.43	1.31	-0.09	1.77	1.19	0.33	1.63	0.78
		Pair 2*	-0.43	undef	undef	-2.54	1.73	-1.89	-0.52	-0.43	-1.82	-0.64	-0.49	0.22*
		Pair 3	0.82	0.1	0.41	-1.56	-2.17	1.23	-0.02	-0.15	0.15	0.21	-0.43	0.17
		Pair 4	-0.01	0.86	-1.73	-2.37	-1.91	-1.17	0.37	-2.97	-2.01	0.22	-2.97	-0.02
3	Family2 visit two	Pair 1	undef	1.89	-1.15	undef	undef	0.34	undef	-1.11	1.24	1.54	1.6	1.61
		Pair 2*	undef	-1.35	0.97	undef	undef	0.1	undef	0.55	0.21	0.93	0.28	0.54*
		Pair 3	undef	1.05	-1.57	undef	undef	0.36	1.77	-1.83	1.21	0.32	1.64	1.14
		Pair 4	-1.21	-2.08	0.48	undef	undef	0.48	-1.45	0.57	1.17	-0.28	1.64	0.89
4	Family2 visit three	Pair 1	1.18	1.87	-1.28	1.49	undef	undef	1.49	1.22	1.85	1.27	1.85	1.42
		Pair 2*	0.67	1.06	-1.89	-0.21	undef	-2.55	0.82	-0.63	-0.59	0.58	-0.47	0.32*
		Pair 3	0.63	-0.11	0.27	-0.46	undef	undef	0.45	0.06	0.81	0.45	0.46	0.51
		Pair 4	0.25	1.18	-2.86	0.58	0.2	-0.45	0.72	0.21	-0.03	0.43	0.24	0.34
5	Family3 visit one	Pair 1	2.73	2.5	1.11	1.7	2.17	0.92	2.64	1.45	1.44	1.89	1.49	1.69
		Pair 2	0.47	0.21	0.4	0.54	-0.43	0.36	0.3	0.45	0.06	0.35	0.2	0.23
		Pair 3*	-0.65	0.54	0.29	0.28	-0.06	0.63	-0.03	0.34	0.38	0.07	0.28	0.24*
		Pair 4	-0.11	-0.05	0.16	-0.06	-0.14	0.43	-0.05	0.03	0.19	0.01	0.11	0.04
6	Family3 visit two	Pair 1*	0.52	0.26	0.25	0.5	undef	1.09	0.33	0.27	1.41	0.28	1.08	0.51*
		Pair 2	0.72	0.19	0.12	-0.62	undef	0.92	0.37	-0.14	0.94	0.34	0.41	0.36
		Pair 3	0.59	0.62	-0.22	0.51	0.04	0.24	0.74	0.15	0.12	0.43	0.29	0.37
		Pair 4	0.5	-0.22	0.37	0.01	0.23	1.01	0.23	0.18	0.6	0.27	0.45	0.34
7	Family4 visit one	Pair 1*	0.07	0.79	0.86	0.36	-0.43	undef	0.39	0.5	-0.18	0.51	0.14	0.22*
		Pair 2	-0.32	-0.04	0.16	-0.03	0.02	undef	-0.17	-0.06	0.5	0	0.26	0.11
		Pair 3	0.15	-0.32	0.47	0.77	0.33	0.83	-0.14	0.64	0.61	0.14	0.68	0.41
		Pair 4	-0.26	0	0.2	-0.05	-0.01	1.85	-0.12	0.07	0.79	0.08	0.51	0.29

obtaining a score within a standard population. In this study, the standard population is the entire play session. The degree of variability within the standard population affects z-scores. In statistical terms, the distribution of the standard normal population is characterised by the mean and standard deviation. In a normal distribution 95% of the population lie within ± 1.96 SDs from the mean. A z-score of 2 or more (either plus or minus) can be considered to lie outside of the variance of the standard population since only 5% of the scores from the standard population would have fallen at that observed score. Therefore, for our analysis, z-scores of 2 or more were categorised as falling outside the normal distribution of the standard population.

Z-scores for each of the shorter segments were derived for each of the contingency pairs. This data are in Fig. 1. The graphs are presented separately for each time slice i.e. all parts of 3 min (Fig. 1a), all parts of 6 min (Fig. 1b) and both parts of 9 min (Fig. 1c). Here the horizontal axis at 0 represents the benchmark as derived from the mean and standard deviation from the 18 min play session. The lines represent the z-scores (standard scores) for each of the shorter segments across all four contingency pairs. Lines that are closer to zero deviate less from the Logg odds computed from the entire play session. The contingency

pairs in the 9 min segments are within one z-score of the entire play session. Some of the z-scores of the contingency pairs from the 6 min and 3 min segments fall outside one standard deviation of the mean of the entire play session. Some of the z-scores from the shorter segments (6 min and 3 min) fall well below two standard deviations (usually in a negative direction, but not always) thus indicating wide variation from the findings of the entire play session.

These data suggest that for some contingency pairs (child look_mother verbal) there was particularly wide variability when data from the short segments are compared to each other. In three of the seven cases, this contingency pair was the pair that occurred least frequently in the entire play session. With regard to the question about the position in the play session that data is sampled from, there is an indication, from observing the data in Fig. 1 that the first 9 min sampled yielded more reliable results than the 9 min segment sampled from the end of the play session.

Discussion

The findings from the single subject and group analysis presented here show that in single subject research, short

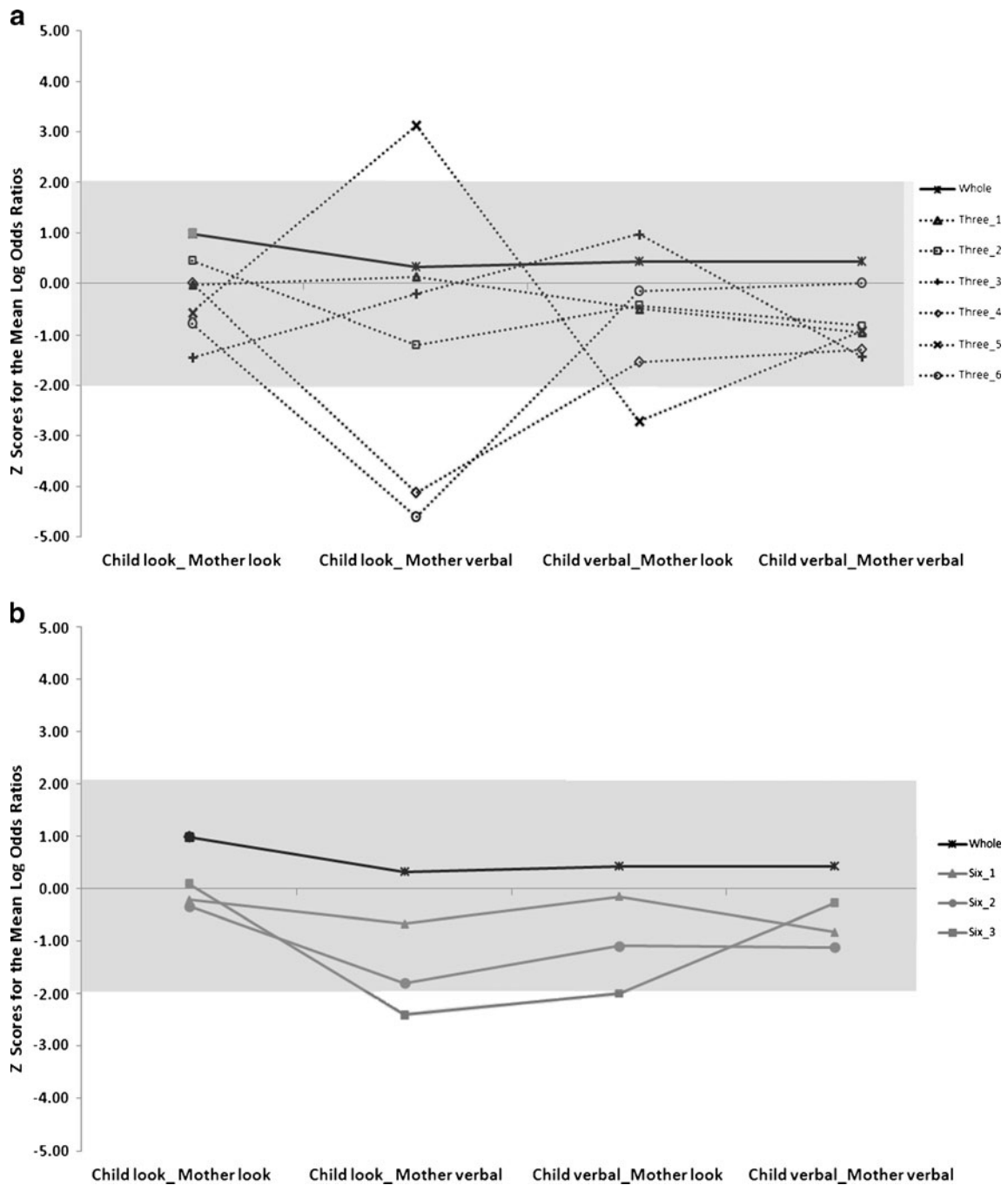


Fig. 1 a *Z-score* calculated on the mean and standard deviation of the Logg odds ratios from the entire play session (18 min), compared to each of the 3 min segments. **b** *Z-score* calculated on the mean and standard deviation of the Logg odds ratios from the entire play session

(18 min), compared to each of the 6 min segments. **c** *Z-score* calculated on the mean and standard deviation of the Logg odds ratios from the entire play session (18 min), compared to each of the 9 min segments

time slices (3 min and 6 min) produced discrepant findings compared to the findings from the whole population (18 min). The results suggest that coding based on half the entire play session yielded results that were neither discrepant nor widely varying when compared to the findings derived from the entire play session. Given the centrality of contingency to the hypothesis in the larger study for

which our methods were being defined, the discrepancy, and the direction of that discrepancy derived from the short time slices provided significant methodological justification for the design of our analysis of a larger study. If we had used 3 min segments to code data then our conclusions would have differed considerably from the findings based on an entire play session. For example, we would have often

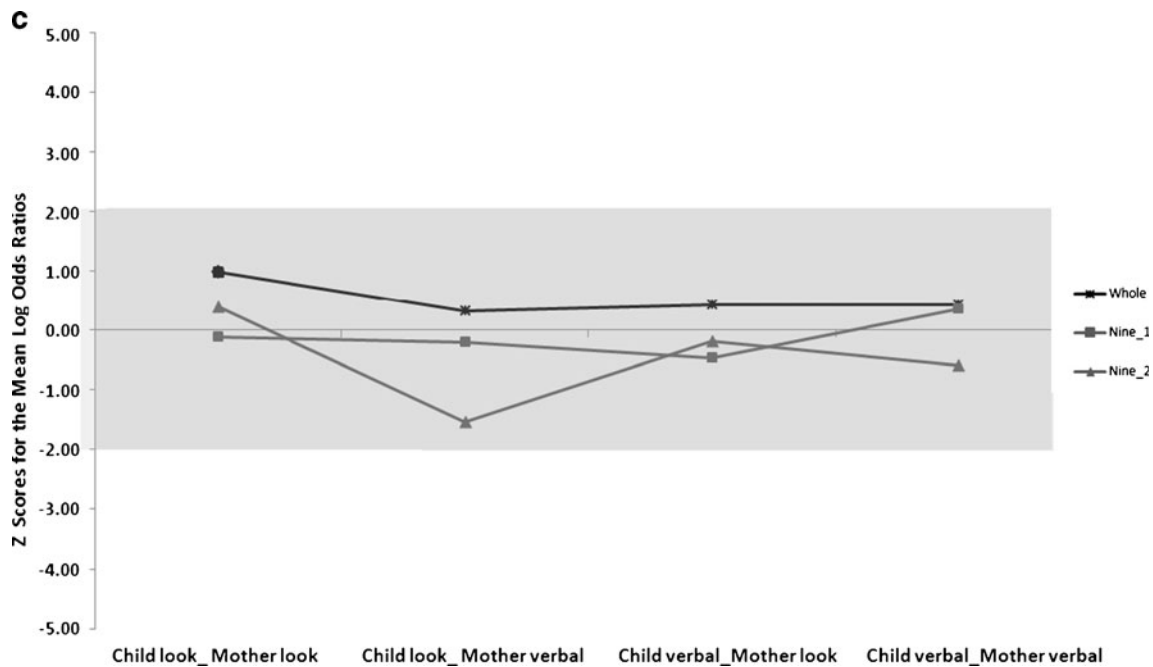


Fig. 1 (continued)

concluded that a mother was unlikely to show verbally contingent behaviour on her child's vocal behaviour where, in fact, this was not the case. Therefore, the experimental hypothesis is supported.

For the wider study for which we were developing our methodology the findings from this small set of cases is enough to generate a well rationalised approach to the sample selection in our single subject design. Clearly, the sample size to be selected is 9 min in duration and the beginning 9 min will be used since this was the only sample that yielded consistently non discrepant results from that of the population. If we were inclined to make a judgement about the parenting style of the parent according to level of responsiveness to the child, based on the shorter 3 min or 6 min observations, we would have classified her differently from the classification that we would have made based on a longer observation. Thus, from these data, in this context, the applicability of Ambady and Rosenthal's thesis, which provides a rationale for thin-slice sampling, was not supported.

A limitation of this study is the small sample size. Whilst this does not impact on the applicability of the findings from this study to our own single case study research or to others who might be using micro-analytic behavioural coding in single subject research, these findings do not have any implications for research adopting group designs. The findings presented here cannot be used by other research groups to provide a rationale for the specific window duration for sample selection of behavioural data. Participant numbers, the number of variable pairs coded, and the frequency of the

measured behaviours will all have an impact on the power to detect the significant patterns within behavioural data. The value of this research report lies in the attention that it draws to the need for researchers to develop and test their methodology for window duration decisions for clinical and experimental research using behavioural coding.

In the review of recent research in the field we were surprised to find how often no rationale on the sampling choices was provided in the study methodology. Our response to this has been to make explicit our own steps to producing robust micro-analytic methods for single-case studies. We think this study provides new knowledge which is relevant to those involved in the design or use of video footage in the applied contexts of pediatric family assessment and diagnosis. Given the constraints on resources in clinic and research, clinical scientists designing new applications using micro-analysis might be a) inclined towards thin-slice coding due resource considerations and b) draw the conclusion that thin slices of behavioural coding gives reliable indicator of a state or trait in single cases based on the group findings in experimental research. Given the potential use of clinical applications using video, such as in assessment and diagnoses and design of therapeutic interventions we urge caution in the application of thin-slice methodology and underline what Bornstein suggests, that every researcher should choose the acceptable limits of unreliability for their study. For our laboratory, making a choice based on previous research was not possible given the way in which we judged the applicability of the prior gold standard evidence on thin-slice coding to our

research. Our aim in articulating our choice making process, including the results of our test is to model what Bornstein (2002) advocated and to show the benefit in questioning the basic premise that underpins methodological choices in this field.

Acknowledgments The NIHR National Biomedical Research Unit in Hearing is funded by the National Institute for Health Research, U.K.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: a meta-analysis. *Psychological Bulletin*, *111*, 256–274.
- Ambady, N., & Rosenthal, R. (1993). Half a minute: predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, *64*, 431–441.
- Bakeman, R., & Quera, V. (1992). SDIS: a sequential data interchange standard. *Behavior Research Methods*, *24*, 554–559.
- Bakeman, R., & Quera, V. (1995). *Analyzing interaction: Sequential analysis with SDIS and GSEQ*. New York: Cambridge University Press.
- Bakeman, R., & Quera, V. (2008). ActSds and OdfSds: programs for converting INTERACT and the observer data files into SDIS timed-event sequential data file. *Behavior Research Methods*, *40*, 869–872.
- Bakeman, R., & Quera, V. (2009). ActSds (Version 2.3).
- Bakeman, R., & Quera, V. (2011a). GSEQ: Generalized Sequential Querier (Version 5.1.07).
- Bakeman, R., & Quera, V. (2011b). *Sequential analysis and observational methods for the behavioral sciences (prepublication copy)*. Cambridge: Cambridge University Press.
- Beebe, B., Jaffe, J., Buck, K., Chen, H., Cohen, P., Blatt, S., et al. (2007). Six-week postpartum maternal self-criticism and dependency and 4-month mother–infant self- and interactive contingencies. *Developmental Psychology*, *43*, 1360–1376.
- Beebe, B., Jaffe, J., Markese, S., Buck, K., Chen, H., Cohen, P., et al. (2010). The origins of 12-month attachment: a microanalysis of 4-month mother–infant interaction. *Attachment and Human Development*, *12*, 3–14.
- Belsky, J., & Fearon, R. M. P. (2002). Early attachment security, subsequent maternal sensitivity, and later child development: does continuity in development depend upon continuity of caregiving? *Attachment and Human Development*, *4*, 361–387.
- Bernier, A., Carlson, S. M., & Whipple, N. (2001). From external regulation to self-regulation: early parenting precursors of young children's executive functioning. *Child Development*, *81*, 326–339.
- Biringen, Z., Damon, J., Grigg, W., Mone, J., Pipp-Siegel, S., Skillern, S., et al. (2005). Emotional availability: differential predictions to infant attachment and kindergarten adjustment based on observation time and context. *Infant Mental Health Journal*, *26*, 295–308.
- Bornstein, M. (1985). How infant and mother jointly contribute to developing cognitive competence in the child. *Proceedings of the National Academy of Sciences of the United States of America*, *82*, 7470–7473.
- Bornstein, M. H. (2002). Measurement variability in infant and maternal behavioral assessment. *Infant Behavior and Development*, *25*, 413–432.
- Chorney, J. M., Garcia, A. M., Berlin, K. S., Bakeman, R., & Kain, Z. N. (2010). Time-window sequential analysis: an introduction for pediatric psychologists. *Journal of Pediatric Psychology*, *35*, 1061–1070.
- Derby, K. M., Wacker, D. P., Sasso, G., Steege, M., Northup, J., Cigrand, K., et al. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: a summary of 79 cases. *Journal of Applied Behavior Analysis*, *25*(3), 713–721.
- Feldman, R. (2006). From biological rhythms to social rhythms: physiological precursors of mother–infant synchrony. *Developmental Psychology*, *42*, 175–188.
- Feldman, R., & Eidelman, A. I. (2009). Biological and environmental initial conditions shape the trajectories of cognitive and social-emotional development across the first years of life. *Developmental Science*, *12*, 194–200.
- Feldman, R., Greenbaum, C. W., Yirmiya, N., & Mayes, L. C. (1996). Relations between cyclicality and regulation in mother–infant interaction at 3 and 9 months and cognition at 2 years. *Journal of Applied Developmental Psychology*, *17*, 347–365.
- Feldman, R., Greenbaum, C. W., & Yirmiya, N. (1999). Mother–infant affect synchrony as an antecedent of the emergence of self-control. *Developmental Psychology*, *35*, 223–231.
- Feldman, R., Eidelman, A. I., & Rotenberg, N. (2004a). Parenting stress, infant emotion regulation, maternal sensitivity, and the cognitive development of triplets: a model for parent and child influences in a unique ecology. *Child Development*, *75*, 1774–1791.
- Feldman, R., Keren, M., Gross-Rozval, O., & Tyano, S. (2004b). Mother–child touch patterns in infant feeding disorders: relation to maternal, child, and environmental factors. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*, 1089–1097.
- Feldman, R., Masalha, S., & Alony, D. (2006). Microregulatory patterns of family interactions: cultural pathways to toddlers' self-regulation. *Journal of Family Psychology*, *20*, 614–623.
- Friedman, D. D., Beebe, B., Jaffe, J., Ross, D., & Triggs, S. (2010). Microanalysis of 4-month infant vocal affect qualities and maternal postpartum depression. *Clinical Social Work Journal*, *38*, 8–16.
- Gable, S., & Isabella, R. A. (1992). Maternal contributions to infant regulation of arousal. *Infant Behavior and Development*, *15*, 95–107.
- Hsu, H. C., & Fogel, A. (2001). Infant vocal development in a dynamic mother–infant communication system. *Infancy*, *2*, 87–109.
- Hsu, H. C., & Lavelli, M. (2005). Perceived and observed parenting behavior in American and Italian first-time mothers across the first 3 months. *Infant Behavior and Development*, *28*, 503–518.
- Hsu, H. C., Fogel, A., & Cooper, R. (2000). Infant vocal development during the first 6 months: speech quality and melodic complexity. *Infant and Child Development*, *9*, 1–16.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. (2001). Rhythms of dialogue in infancy. *Monographs of the Society for Research in Child Development*, *66*, 1–132.
- Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., et al. (1989). Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*, *19*, 185–212.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*, 205–223.
- Mangold (2008). INTERACT (Version V9.4.2): Mangold International.
- Milgrom, J., Westley, D. T., & Gemmill, A. W. (2004). The mediating role of maternal responsiveness in some longer term effects of postnatal depression on infant development. *Infant Behavior and Development*, *27*, 443–454.

- Moore, G. A., Cohn, J. F., & Campbell, S. B. (1997). Mothers' affective behaviour with infant siblings: stability and change. *Developmental Psychology, 33*, 856–860.
- Murray, A. D., & Hornbaker, A. (1997). Maternal directive and facilitative interaction styles: associations with language and cognitive development of low risk and high risk toddlers. *Development and Psychopathology, 9*, 507–516.
- Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 177–197). NJ: Ablex.
- Ray, S. A., & Tickle-Degnen, L. (2004). The validity of a methodology to describe videotaped mother–toddler interactions. *OTJR Occupation, Participation and Health, 24*, 123–133.
- Rodriguez, M. L., Dumont, K., Mitchell-Herzfeld, S. D., Walden, N. J., & Greene, R. (2010). Effects of healthy families New York on the promotion of maternal parenting competencies and the prevention of harsh parenting. *Child Abuse and Neglect, 34*, 711–723.
- Roger, T., Zarcone, J., & Iwata, B. (1990, May). A comparison of methods for conducting functional analyses of self-injurious behaviour. D. Wacker (chair). *Applications of functional analysis across settings and procedures*. Symposium presented at the Annual Meeting of the Association for Behavior Analysis. Nashville, TN.
- Shin, H., Park, Y.-J., & Kim, M. J. (2006). Predictors of maternal sensitivity during the early postpartum period. *Journal of Advanced Nursing, 55*, 425–434.
- SPSS. (2007). *SPSS base 16.0 for Windows user's guide (Version 16.0.0)*. Chicago: IBM-SPSS.
- Stern, D. N. (1983). A microanalysis of mother–infant interaction. *Psychiatrie De L Enfant, 26*, 217–235.
- Tait, M., Lutman, M. E., & Nikolopoulos, T. P. (2001). Communication development in young deaf children: review of the video analysis method. *International Journal of Pediatric Otorhinolaryngology, 61*(2), 105–112.
- Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: research, theory and clinical applications. *Journal of Child Psychology and Psychiatry, 42*, 3–48.
- Treyvaud, K., Anderson, V. A., Howard, K., Bear, M., Hunt, R. W., Doyle, L. W., et al. (2009). Parenting behavior is associated with the early neurobehavioral development of very preterm children. *Pediatrics, 123*, 555–561.
- Tronick, E. Z. (1995). Touch in mother–infant interaction. In T. M. Field (Ed.), *Touch in early development* (pp. 53–65). NJ: Erlbaum.
- Vigil, D. C. (2002). Cultural variations in attention regulation: a comparative analysis of British and Chinese-immigrant populations. *International Journal of Language and Communication Disorders, 37*, 433–458.
- Weiss, S. J., Wilson, P., Hertenstein, M. J., & Campos, R. (2000). The tactile context of a mother's caregiving: implications for attachment of low birth weight infants. *Infant Behavior and Development, 23*, 91–111.