Caroline de Weerth and Paul van Geert

Transitions and variability in infant behaviour

Abstract

An attempt was made at validating van de Rijt-Plooij & Plooij's (1992) theory that infants go through 10 strictly timed periods of emotional instability in their first 15 months of life. No evidence was found to support this theory. On the contrary, our findings point to important individual differences in development and no common pattern of emotional instability.

Furthermore, evidence was found of important intra-individual variability in the studied behaviours, both on a weekly and on a daily basis. The importance and possible factors underlying this variability are discussed.

Introduction

A transition in development has been defined as a qualitative change that is both large and rapid (Fischer et al., 1984). Many of the papers presented in this volume deal with this non-linear phenomenon. However, in our study a different approach was followed. Transitions have often been found to be associated with instability, disorganization and/or regression in behaviour (Woollacott & Sveistrup, 1992; Freedland & Bertenthal, 1994), and in the present study it is this very *instability* of behaviour that was used in an attempt to determine the presence of transitions in infantile development. Beyond this, the study also looks into non-linear processes at the macro-level of behaviour, by studying the important intra-individual variability that exists in emotion-related behaviours. The starting point of this approach lies in a claim by Van de Rijt-Plooij & Plooij (1992) that there exist 10 periods of instability (which they called 'regressions'), in the first twenty months of life. These authors performed a longitudinal study of 15 mother-infant pairs, and found that in the first 15 months of life there were narrowly timed periods in which the infants showed signs of emotional instability. These periods took place around: 5, 8, 12, 17, 26, 30, 36, 44, 52 and 62 weeks of life, and according to van de Rijt-Plooij & Plooij, they preceded developmental transitions in the cognitive, motor, social and language areas. They based their conclusions on maternal questionnaires which the mothers completed on a weekly basis and returned by post, and on naturalistic monthly home observations of two of the mother-infant pairs.

The criteria that the authors used to determine weeks of emotional instability or 'regression' in the questionnaire data was the following. Before the infant had reached

the age of 2 months, the mothers had to have reported both increases in crying/fussing and in the time spent in body contact with their infants; after 2 months of age, both these behaviours had to be present and also one or more of the following list of problem behaviours: a decrease in the amount of sleep, fear of other people, childish behaviour, a decrease in the amount of food intake in one meal, problems changing/dressing, a decrease in activity, a peak in cuddling the mother and/or a peak in cuddling objects. It was by using these criteria that van de Rijt-Plooij & Plooij determined 10 strictly timed periods of emotional instability in the first 15 months of life of their study group of infants. A very high accordance was found in the reports of all the mothers: while there were weeks in which none of the mothers' reports fulfilled the criteria for 'regression' weeks, there were others in which as many as 80-100% of the mothers' reports indicated 'regression' weeks. The authors also found that the monthly objective recordings of the behaviour of two of the mothers with their infants corroborated the questionnaire data, but recommended the ethological observations be used on a weekly basis in future research on the phenomenon, so that none of the 'regression' periods could be missed.

Strictly timed periods of emotional instability

The present project was set up with the goal of validating the pattern of narrowly timed 'regression' periods found by van de Rijt-Plooij & Plooij (1992). Evidence for ensuing transitions in general development would also be sought for.

The design of the project was to perform a longitudinal study (0-15 months of age) of 4 mother-infant pairs (2 male and 2 female normally developing infants) from which three types of data were obtained. First of all the mothers filled in weekly reports: a 24-hours a day logbook of their infants' hours of sleep and a questionnaire on their infants' behaviour. This questionnaire was very similar to the one used in the Plooijs' original study, was focussed on emotional instability and problem behaviour, and was the basis of a weekly interview of the mother. Second, the sleeping behaviour of the infant was automatically registered 24-hours a day, by placing sensors next to the crib. The infant's movements, crying and the presence of an adult could in this way be registered. This data has so far not been analysed and will therefore not be discussed further in this paper. Finally, naturalistic home observations of each mother-infant pair took place on a weekly basis. These observations were all performed by the first author, after receiving an extensive training in the observation technique by H.H.C. van de Rijt-Plooij. The observations took place on a fixed day of the week for each infant, started at 9.00, stopped temporarily if the infant slept, and went on until 3 hours of data had been collected. The observed behaviours were directly fed into a laptop computer and consisted of contact/distance between the mother and her infant (7 subcategories), crying, fretting/fussing, smiling and qualitatively new behaviours displayed by the infant. Contact/distance and crying were scored continuously, and fretting/fussing and smiling were scored with interval sampling (for a detailed description of the study see: de Weerth & van Geert, in press).

After correcting the infants' ages for gestational age, the observational data was plotted individually against the age in weeks. The resulting graphs immediately

revealed general trends in most of the time series (i.e. crying and time spent in body contact decrease with age, while smiling increases), with peaks and dips superimposed on these trends (de Weerth & van Geert, 1998).

Different analytical techniques were used to investigate whether the observational data from the present project validated the 10-regressions pattern described by van de Rijt-Plooij & Plooij (1992). Due to lack of space, a short description will be given of each one, together with a general conclusion about the results. For more details on the methods and results, and for a general discussion of the study, see de Weerth & van Geert (1998).

With the object of investigating whether the results of this study were consistent with those of the Plooij study, several measures were derived from the data of their study. The one most used was to define 'regression' weeks as those in which 60-100% of the mothers had reported behaviour that fulfilled the 'regression' criteria. The patterns of 'regression' weeks that were determined in the present study with the methods described below, were thus compared to the pattern determined by the 60-100% criterion of the earlier study, namely weeks: 5, 8, 9, 12, 15-19, 24-26, 29, 30, 34-37, 42-45, 51-54, 61-63.

Different measures were used for determining the 'regression' weeks in this study. For the questionnaire data, the maternal report had to fulfill the same criteria used by the Plooijs in order for a week to be called a 'regression' week. For the observational data, peaks in both body contact *and* in crying (duration and/or frequency) or fretting/fussing were determined, but these variables were also used independently to determine 'regressions'. In this way the analysis of the data acquired flexibility and the chances of validating the 10-regressions pattern increased. It was for the same reason that several analytical methods were applied to the observational data: each method had its advantages and disadvantages, and by using all of them the chances of not being able to validate the phenomenon due to the analytical method were greatly reduced.

In the time series obtained with the observational data, a *weighted moving average* method was used in order to compare each data point with its neighbouring points and in this way detect the most important peaks in the data series[1]. The pattern of peaks thus obtained was then compared to the pattern of regression weeks of the Plooij study, in order to see whether the peaks fell in the hypothesized regression weeks. This method gives more weight to sharp, one-week 'regression' periods than to longer periods in which the values are elevated.

The data series were then detrended by using non-linear polynomial regression functions[2]. The residuals obtained in this manner were used for the remaining analyses.

[1] A peak occurred in week t if x_t was clearly higher than expected when compared to the values of the surrounding weeks. In order to determine whether this was the case, a variable (R_t) was created which took into account all four surrounding weeks but gave more weight to the two closest neighbouring weeks: $R_t = x_t - (x_{t-2} + 2x_{t-1} + 2x_{t+1} + x_{t+2})/6$. Thus, R_t was a weighted moving average and the largest R_t's indicated the highest peaks in the data.

[2] Most of the graphs of the observed behaviours displayed important non-linear trends in the data. Because we were interested in deviations from local averages and not in general, long-term trends, the data were detrended as follows. Second degree polynomial curves ($P_t = a + b\,t + ct^2$) were fitted onto the data series, because such curves eliminate linear and eventually also quadratic trends. The residuals were then calculated as the difference between each observational data point (O_t) and its corresponding polynomial curve point (P_t).

Pearson's correlations were calculated between the residuals of the data series and the population curve of the Plooij study (i.e., the exact percentage of infants that fulfilled the 'regression' criteria for each week of development). This method served to investigate whether the data of the present study followed that of the earlier study in a general way.

The amount of *positive residuals* in the hypothesized 'regression' weeks was analysed in order to see whether it was relatively greater than the amount in the remaining observation weeks. This method underlines increases in the studied behaviours without giving importance to the magnitude of the increases (because only the positive or negative sign of the residual are taken into account). The *means of the residuals* in the hypothesized 'regression' weeks were compared to those in the remaining weeks, with the expectation that they would be higher in the first case. This method emphasizes both subtle and large increases in the observed behaviours. These last two methods are adequate for detecting the 10-regressions pattern if it were not characterized solely by sharp (i.e. one week) increases in the behaviours.

They can detect longer increases (i.e. two or more weeks) and general subtle increases, and even mixtures of sharp and longer increases.

The results of the analyses were disappointing. Although all four infants showed temporary elevations in crying, fretting/fussing and in the time spent in body contact with their mothers, and although all the mothers reported 'difficult' weeks which fulfilled the 'regression' criteria of the earlier study, none of the analytical methods showed that the data of this study followed the 'regression' pattern reported by the Plooijs. This was so notwithstanding the fact that the criteria had been made more flexible and that an array of analytical methods had been used, and it was true for both the maternal questionnaire data and the observational data of the mother-infant pairs. Interestingly, the present maternal reports were confirmed by the behavioural observations: the means of the residuals were consistently higher in the weeks in which the mothers' reports fulfilled the 'regression' criteria than in those in which they did not. Therefore, although indications of periods of emotional instability were found, no evidence was obtained that these periods were strictly timed and followed a general pattern for all the infants.

Thus, the conclusion of the present study is that our results fail to support the 10-regression periods pattern proposed by van de Rijt-Plooij & Plooij (1992).

Intra-individual variability: day-to-day changes

As we have seen, the data of the four infants who had been closely followed for a 15-month period failed to show a common pattern of periods of emotional instability. Moreover, except for a few general developmental trends, the data displayed an important variability or instability. The observed levels of each behaviour often varied greatly from one week to another. The fact that the intra-individual variability of the data was the rule rather than the exception, made it worthy of closer inspection.

The Plooijs used the week as the time unit to describe increases in crying/fussing and body contact. However, the mothers of the present study had often reported important

day-to-day changes in their infant's behaviour. Furthermore, there are indications from the literature that behaviours such as crying/fussing and difficult behaviour in general vary greatly on a daily basis and that this is an intrinsic feature of normal infant behaviour (St. James-Roberts & Halil, 1991; St. James-Roberts & Wolke, 1988; Barr, 1990; Rebelsky & Black, 1972). In order to investigate whether the behaviours we had observed on a weekly basis also varied importantly on a day-to-day basis, an exploratory study was implemented. The object was to compare the variability within a week to that between weeks. If a week constituted an adequate time unit in which to measure these behaviours, then the variability between weeks would be expected to be greater or equal to that between days of the same week. However, based on our maternal reports and on the above-mentioned literature on variability, we expected the contrary to be true. Specifically, the mean differences between the days of the same week were expected to be equal or greater than the mean differences between the means of different weeks.

The exploratory study was a longitudinal study of one mother-infant pair, which took place when the infant was around 52 weeks of age. This infant was selected with basically the same criteria as the remaining four infants of this research project, and was a normally developing female infant. The data was obtained with the same type of behavioural observations described above, except that this infant was observed 3 times a week (Mondays, Wednesdays and Fridays) on 4 consecutive weeks.

An example of the plots resulting from the data can be seen in Figure 1. In it the frequency of crying is plotted against the weeks of observation, with one curve for each day of the week. In this way it is possible to visualize the variations within a week and also to see what the effects had been had this infant been observed once a week on a fixed day, as the remaining four infants of this study. The variability within weeks immediately strikes the eye, together with the fact that different results would have been obtained had this infant been observed on Mondays, Wednesdays or Fridays.

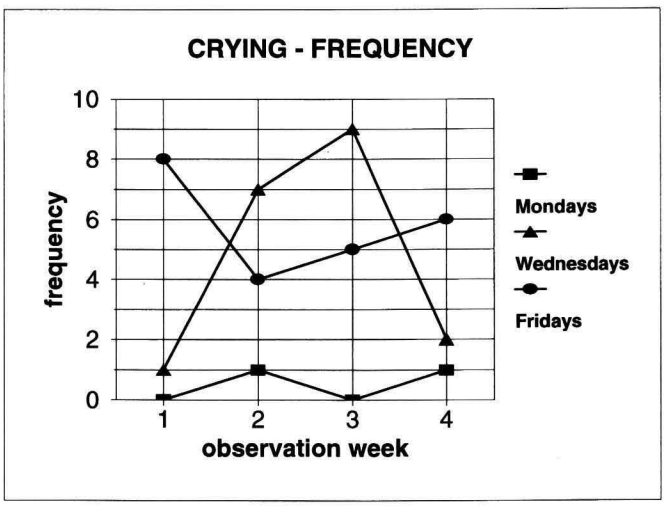


Fig. 1. Variability between and within weeks for the scores of frequency of crying.

That the variability within weeks is important is supported by the comparison of mean differences of scores. The average differences between the mean behavioural scores of the weeks are: for total time spent in body contact with mother (total contact), 3.67; for duration of crying, 0.79; for frequency of crying, 1.00; for fretting/fussing, 3.69 and for smiling, 8.80.

The average differences in scores between days of the same week are greater than the differences between the weeks, in 18 of the 20 cases (the exceptions are week 4 for duration of crying and week 2 for smiling). Also, all the general averages for within week differences (i.e. the means for the four weeks together), are greater than the mean differences between weeks: for total contact, 8.65; for duration of crying, 1.23; for frequency of crying, 4.67; for fretting/fussing, 9.66 and for smiling, 10.65. Again, this points at the day-to-day variability being as great or greater than the variability between weeks.

Notwithstanding the exploratory nature of this study, the results are consistent enough to allow the conclusion that there are indications of important day-to-day variability in the behaviours studied: total time spent in body contact with the mother, crying, fretting/fussing and smiling. For a detailed report on this study see de Weerth & van Geert (in prep.).

Intra-individual variability: week-to-week changes

Based on the above, the question arose whether it was possible to obtain a representative measure of an infant's time spent in body contact, crying, fretting/fussing and smiling, and of the developmental changes therein, based on a certain observation frequency. Longitudinal studies often consist of 2 to 6 or even 8 observation points which are often also separated from each other by weeks or months. Were an infant's behaviour to vary greatly on a day-to-day basis, then a sole observation for an infant at a given age or in a given condition need not be representative of the infant's everyday behaviour at that age or in those conditions. Studies making use of few data points would be subject to random influences on the data due to infants' variable behaviour. Moreover, researchers from different areas of infant development have repeatedly reported instability of infant behaviour over different observation or test sessions and even within a session (Canfield et al., 1995; Dittrichova et al., 1992; Wachs et al., 1990; Worobey, 1986; Isabella et al., 1985; Peters-Martin & Wachs, 1984). This instability, which consists of a lack of consistency in individual patterns, tends to decrease towards the end of the first year (Snidman et al, 1995; Denham et al., 1995; Matheny, 1986; St. James-Roberts & Wolke, 1988). After this age, inter-individual differences become more stable and established.

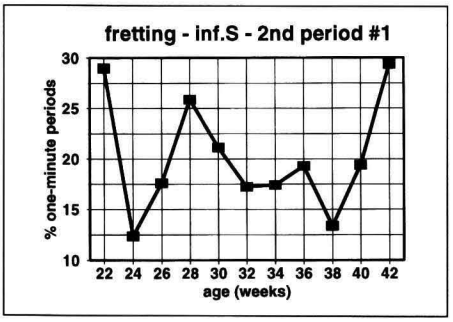
In order to answer our question, a second study was performed on the data of the four infants who had been followed weekly for 15 months. The question was instrumentalized in the following hypothesis: due to day-to-day variability, significant differences will be found in developmental trends between two alternating series of 2-

weekly observations (i.e. a series of even weeks and one of odd weeks) within one and the same infant. These differences will be greater for young infants and tend to disappear with age, as the infant's behaviour stabilizes.

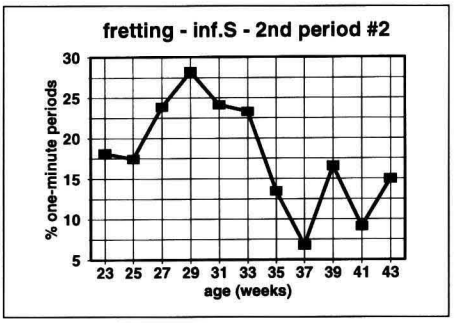
This hypothesis was expected to apply for total contact, crying and fretting/fussing. In the case of smiling, the early variability was not expected due to the fact that early on the infants are still in the process of learning to smile. These four infants had a mean of 32.3% of intervals in which they smiled during an observation session (range: 28.7-38.4%), but it took them around 10 weeks (range: 5-13 weeks) to reach a minimum level of 10% of intervals with smiles. Therefore, because the behaviour was not totally established in the beginning, it is not unreasonable to hypothesize that the variability in smiling will not be very important then.

With the object of approximating the observation length and number of data points of most longitudinal studies, and also of studying the age-related differences in variability, the 15-month time series were divided into 3 shorter series: 0-5 months, 5-10 months and 10-15 months. Multiple regressions were used to analyse the differences in developmental trends (specifically in means, linear components and quadratic components, and also overall or general differences). A meta analytic technique was used to combine the p-values of the four individual infants into one general p-value[3].

a



b



Figs. 2a and b. Alternate 2-weekly data series for fretting/fussing for a female infant between the ages of 5 and 10 months.

Figures 2a and 2b show an example of the curves that are obtained when plotting the alternate series of 2-weekly data points. In this case, fretting for a female infant between 5 and 10 months of age, the curves show apparent differences in developmental trends. However, the significant differences for the variables studied, total contact, duration of crying, fretting/fussing and smiling, can best be seen in Table I. In this table the combined p-values for the four infants are presented separately for

[3] The question we would like to answer is not whether the trends differ for each individual infant, but whether they differ for the population from which the four infants were drawn. In order to obtain an estimate of the difference at the population level, a meta analytic technique was employed based on the individual p-values of the four infants. This was necessary in order to test the hypotheses of equal mean, linear, quadratic and general effects, for the population from which the infants were selected. The meta analytic procedure used is called 'the chi-square method' (Glass, McGaw and Smith, 1981, p. 99), and it combines the p-values of the infants into one overall p-value.

each of the data series elements studied. In the first two periods significant differences were found between one or more elements of the alternate series of 2-weekly observations: in the 0 to 5 months period, for total contact and crying, and in the 5 to 10 months period, for crying, fretting and smiling. The 10 to 15 months period did not show significant differences between the data series. Further analysis of the data revealed that outlying data points were not the systematic cause of differences between curves, nor were the differences due to only one particular infant's data.

Table I. Significant differences found between alternate data series. P-values are combined for all 4 infants (meta analysis). '-' means p >.20.

1st period (0-5 months)

Difference in:	Behavior:			
	Tot.Cont.	Crying(d)	Fret/fuss	Smiling
means	0.79	0.04	(.17)	-
linear components	0.01	0.01	-	-
quadratic components	0.03	0.02	-	-
general	0.19	0.02	0.79	0.92

2nd period (5-10 months)

Difference in:	Behavior:			
	Tot.Cont.	Crying(d)	Fret/fuss	Smiling
means	-	0.30	0.05	0.03
linear components	-	0.01	0.24	0.81
quadratic components	-	0.03	0.20	0.57
general	0.99	0.04	0.11	0.19

3rd period (10-15 months):

No significant differences

The conclusion of this study is that the results support the hypothesis: 5 (first period: total contact, crying; second period: crying, fretting and smiling) out of the 7 expected variables (the five mentioned, plus fretting in the first period and total contact in the second period) showed significant differences between the series. These differences might very well be due to day-to-day variability in infant behaviour. For a detailed description of the study and its results, see de Weerth, Hoijtink & van Geert (subm.).

Discussion

As a general conclusion, our results point to the existence of a marked variability in an infant's expressions of negative and positive emotion and in the time spent in

physical contact with its mother. These behaviours vary greatly from one week to the other, and there are also indications that they vary greatly on a day-to-day basis. Taking this variability into account, it is perhaps not surprising that the general pattern of strictly timed periods of emotional instability proposed by van de Rijt-Plooij & Plooij (1992) was not found in this study population. Although both the maternal reports and the behavioural observations indicated that the infants went through 'difficult' periods of increased negative emotionality, there were important inter-individual differences in the amount and patterning of these periods.

The sample of infants studied in this project appeared to be normal, the length of the observation period was adequate and the reliability of the observer was controlled for (see de Weerth & van Geert, 1998). Because of this and of the literature that has reported day-to-day variability in infants' difficult behaviour (St. James-Roberts & Halil, 1991; St. James-Roberts & Wolke, 1988; Barr, 1990; Rebelsky & Black, 1972), we are inclined to believe that the intra-individual variability encountered in the present study was not due to unstable scores caused by external influences or methodological problems, but that it is an intrinsic characteristic of early infant behaviour.

The question which immediately arises is: Why should the studied behaviours show such marked day-to-day variability? One of the factors underlying the variability might be changes in the mother-infant relationship. This relationship is developing: both partners try out new ways of communicating with each other, and also change them over time. Maternal behaviour has been shown to change in the course of the infant's first year (Crockenberg & McCluskey, 1986), and of course, so does the behaviour of the rapidly developing infant (see de Weerth & van Geert, subm., for an analysis of changes in patterns of mother-infant behaviours). Therefore, mother and infant tune in to each other and influence each other with their moods, attitudes, developing skills, etc. Thus, apart from changes in the long run, the interaction quite possibly varies also on a daily or weekly basis (especially in such behaviours as body contact and negative vocalizations) with changes in both partners' mood and activities, and differences in the infant's abilities, experiences and needs.

Also, an infant's day-to-day variability could have a biologically adaptive value in that it helps retain the mother's attention. A young infant's main means of attracting its mother is by crying/fussing. It seems only logical that were the infant to present these behaviours in exactly the same quantities and patterns each day, the mother would become accustomed to the infant's displays and cease to pay as much attention to the infant as would be optimal for him. Day-to-day variability in behaviour would ensure the infant of continuous maternal attention and efforts to take care of his needs and to reach a stable mother-infant interaction. Thus, the mother-infant bond would strengthen and the infant would be able to develop more sophisticated ways of communicating his needs. As an infant grows older, he will depend less on negative vocalizations as his main channel of communication. Sophisticated facial expressions, gestures, sounds and words, and independent locomotion, will enrich his possibilities for interacting with his mother and communicating his needs effectively.

Our finding that the intra-individual variability decreases in the period between 10 and 15 months fits in with this view of development.

Supporting the idea that day-to-day variability is a constitutional and adaptive factor of infant behaviour, St. James-Roberts & Halil (1991) found that during the first year of development, normal infants showed an overall more *irregular* day-to-day crying pattern than infants whose parents had sought clinical assistance due to their infant's excessive crying. These authors concluded that day-to-day variability is common and is not affected by the pattern or level of crying in normal infants. Therefore, it would appear that a high level of crying, notwithstanding it showing more regularity in the day-to-day pattern, is more stressful to parents than a lower level of crying displayed with important day-to-day variations. This would mean that variability *per se* need not be as stressful to parents as one might tend to think, and supports the hypothesis that it is an intrinsic feature of normal infant behaviour, that possibly has adaptive advantages.

There are also theoretical reasons to believe intra-individual variability should be expected in the behavioural repertoire of infants. Goldsmith (1993) approaches the question of early intra-individual variability by adhering to a dynamic systems theorists view of emotions as variable rather than discrete. The instability found in individual infants' measures of temperament would be due to emotions being self-organizing processes that result from dynamic interactions among many constituent elements, such as cognitive, motor and social components (see also M.D. Lewis, 1994, 1996; van Geert, 1994). Therefore, individuals would show great variability in outcome of emotional regulation processes not only with respect to other individuals, but also over different situations and ages. As the individual grows, more 'temperamental' consistency probably appears. Studies that find increasing stability in temperamental scores as an infant develops (Matheny, 1986; Denham et al., 1995) support this idea.

A comparable dynamic systems approach to the development of mother-infant interaction patterns is followed by Fogel (1992, 1993, Fogel and Thelen, 1987). According to Fogel, the development of communication is based on the establishment of specific consensual frames, which make use of the available behavioural and emotional repertoires. At the beginning, new frames must be explored and negotiated, a process which requires as well as produces considerable variability in the resulting interactions.

Thus, assigning fixed temperamental characteristics to young infants is an apparently useless simplification of the issue, if, as the evidence appears to indicate, the most important 'temperamental' characteristic of this group is a constitutional variability of behaviours related to emotions.

A final remark can be made based on our results. As we have seen, a marked variability was found in the trajectories followed by the growth curves of individual infants. If one is interested in making generalizations on developmental processes over longer periods of time, then studying the phenomenon at the group level seems a good choice. However, the general growth curve found might not be representative of any single infant. For example, when fitting a quadratic model to

the group data, some individual curves with positive curvatures may cancel others with negative curvatures, the result being a group curve for development that is descriptively inadequate for individual development. This marks the importance of examining individual time series (which as we have seen, should also take into account intra-individual variability), before performing analyses at the group level.

Future research is needed in order to determine the extent of intra-individual variability in behavioural and physiological measures, and how generalizable it is over different areas of research and ages. Also, the exact influence that this phenomenon has on research results and conclusions should be determined, thus aiding future researchers in the design of their research projects.

References

Barr, R.G. (1990). The normal crying curve: What do we really know?. *Developmental Medicine and Child Neurology*, **32**, 356-362.

Canfield, R.L., J. Wilken, L. Schmerl and E.G. Smith (1995). Age-related change and stability of individual differences in infant saccade reaction time. *Infant Behavior and Development*, **18**, 351-358.

Crockenberg, S. and K. McCluskey, (1986). Change in maternal behavior during the baby's first year of life. *Child Development*, **57**, 746-753.

Denham, S.A., E.B. Lehman, M.H. Moser and S.L. Reeves (1995). Continuity and change in emotional components of infant temperament. *Child Study Journal*, **25**, 289-308.

Dittrichova, J., K. Paul, M. Tautermannova and J. Vondracek (1992). Individual variability in infant's early behavior. *Studia Psychologica*, **34**, 199-210.

Fischer, K.W., S.L. Pipp and D. Bullock (1984). Detecting developmental discontinuities. Methods and measurement. In R.N. Emde and R.J. Harmon (Eds.) *Continuities and Discontinuities in Development*, (pp. 95-121). New York: Plenum.

Fogel, A. (1992). Movement and communication in infancy: The social dynamics of development. *Human Movement Science*, **11**, 387-423.

Fogel, A. (1993). *Developing through relationships: Origins of communication, self, and culture.* Chicago: University of Chicago Press.

Fogel, A. and E. Thelen (1987). The development of expressive and communicative action in the first year: Reinterpreting the evidence from a dynamic systems perspective. *Developmental Psychology*, **23**, 747-761.

Geert, van P. (1994). *Dynamic systems of development: change between complexity and chaos.* New York: Prentice Hall/Harvester Wheatsheaf.

Glass, G.V., B. McGaw and M.L. Smith (1981). *Meta analysis in social research.* London: Sage.

Goldsmith, H.H. (1993). Temperament: Variability in developing emotion systems. In M. Lewis and J.M. Haviland (Eds.) *Handbook of Emotions* (pp. 353-364). New York: The Guilford Press.

Freedland, R.L. and B.I. Bertenthal (1994). Developmental changes in interlimb coordination: Transition to hands-and-knees crawling. *Psychological Science*, **5**, 26-32.

Isabella, R.A., M.J. Ward and J. Belsky (1985). Convergence of multiple sources of information on infant individuality: Neonatal behavior, infant behavior, and temperament reports. *Infant Behavior and Development*, **8**, 283-291.

Lewis, M.D. (1994). Cognition-emotion feedback and the self-organization of developmental paths. Human Development, **38**, 71-102.

Lewis, M.D. (1996). Self-organising cognitive appraisals. *Cognition and emotion*, **10,** 1-25.

Matheny, A. (1986). Stability and change of infant temperament: Contributions from the infant, mother and family environment. In G.A. Kohnstamm (Ed.) *Temperament discussed. Temperament and development in infancy and childhood*, (pp. 49-58). Lisse: Swets and B.V. Zeitlinger.

Peters-Martin, P. and T.D. Wachs (1984). A longitudinal study of temperament and its correlates in the first 12 months. *Infant Behavior and Development*, **7**, 285-298.

Rebelsky, F. and R. Black (1972). Crying in infancy. *The Journal of Genetic Psychology*, **121**, 49-57.
Rijt-Plooij, H.H.C. van de and F.X. Plooij (1992). Infantile regressions: Disorganization and the onset of transition periods. *Journal of Reproductive and Infant Psychology*, **10**, 129-149.
Snidman, N., J. Kagan, L. Riordan and D.C. Shannon, (1995). Cardiac function and behavioral reactivity during infancy. *Psychophysiology*, **32**, 199-207.
St. James-Roberts, I. and D. Wolke, 1988 - Convergences and discrepancies, among mothers' and professionals' assessments of difficult neonatal behaviour. *Journal of Child Psychology and Psychiatry*, 29, 21-42.
St James-Roberts, I. and T. Halil (1991). Infant crying patterns in the first year: Normal community and clinical findings. *Journal of Child Psychology and Psychiatry*, **32**, 951-968.
Wachs, T.D., J. Morrow and E.H. Slabach (1990). Intra-individual variability in infant visual recognition memory performance: Temperamental and environmental correlates. *Infant Behavior and Development*, **13**, 397-403.
Weerth, de C. and P. van Geert (1998). Emotional instability as an indicator of strictly timed infantile developmental transitions. *British Journal of Developmental Psychology*, **16**, 15-55.
Weerth, de C. and P. van Geert (in prep.). *Does an infant's behavior vary much on a day-to-day basis?*
Weerth, de C. and P. van Geert (subm.). *Changing patterns of infant behavior and mother-infant interaction.*
Weerth, de C., H. Hoijtink and P. van Geert (subm.). *Intra-individual variability in infant behavior.*
Woollacott, M. and H., Sveistrup (1992). Changes in the sequencing and timing of muscle response coordination associated with developmental transitions in balance abilities. *Human Movement Science*, **11**, 23-36.
Worobey, J. (1986). Convergence among assessments of temperament in the first month. *Child Development*, **57**, 47-55.