

Temperamental differences in infants and young children

Ciba Foundation symposium 89

1982

Pitman
London

Temperamental differences in infants and young children

The Ciba Foundation is an international scientific and educational charity. It was established in 1947 by the Swiss chemical and pharmaceutical company of CIBA Limited—now CIBA-GEIGY Limited. The Foundation operates independently in London under English trust law.

The Ciba Foundation exists to promote international cooperation in medical and chemical research. It organizes international multidisciplinary meetings on topics that seem ready for discussion by a small group of research workers. The papers and discussions are published in the Ciba Foundation series.

The Foundation organizes many other meetings, maintains a library which is open to graduates in science or medicine who are visiting or working in London, and provides an information service for scientists. The Ciba Foundation also functions as a centre where scientists from any part of the world may stay during working visits to London.

Temperamental differences in infants and young children

Ciba Foundation symposium 89

1982

Pitman
London

© Ciba Foundation 1982

ISBN 0 272 79653 0

Published in June 1982 by Pitman Books Ltd, London. Distributed in North America by CIBA Pharmaceutical Company (Medical Education Administration), Summit, NJ 07006, USA.

Suggested series entry for library catalogues:
Ciba Foundation symposia.

Ciba Foundation symposium 89
x + 310 pages, 7 figures, 42 tables

British Library Cataloguing in publication data:
Temperamental differences in infants and young
children.—(Ciba Foundation symposium series; 89)
1. Temperament—Congresses 2. Child psychology
—Congresses
I. Porter, Ruth II. Collins, GERALYN M.
III. Series
155.4'18 BF811

ISBN 0-272-79653-0

Text set in 10/12 pt Linotron 202 Times, printed and bound
in Great Britain at The Pitman Press, Bath

Contents

Symposium on Temperamental differences in infants and young children held at the Ciba Foundation, London, 22–24 September 1981

Editors: Ruth Porter (Organizer) and GERALYN M. COLLINS

- MICHAEL RUTTER (*Chairman*) Temperament: concepts, issues and problems 1
Discussion 16
- THOMAS F. McNEIL and INGER PERSSON-BLENNOW Temperament questionnaires in clinical research 20
Discussion 31
- JIM STEVENSON and PHILIP GRAHAM Temperament: a consideration of concepts and methods 36
Discussion 46
- J. STEVENSON-HINDE and A. E. SIMPSON Temperament and relationships 51
Discussion 62
- R. A. HINDE, D. F. EASTON, R. E. MELLER and A. M. TAMPLIN Temperamental characteristics of 3–4-year-olds and mother-child interaction 66
Discussion 80
- JUDY DUNN and CAROL KENDRICK Temperamental differences, family relationships, and young children's response to change within the family 87
Discussion 101
- GENERAL DISCUSSION I Methods of assessment of temperament 106,
Categorization of temperament 111, Cross-cultural studies 113

- RONALD S. WILSON *Intrinsic determinants of temperament* 121
Discussion 135
- ANNE MARI TORGERSEN *Influence of genetic factors on temperament development in early childhood* 141
Discussion 148
- ROBERT PLOMIN *Behavioural genetics and temperament* 155
Discussion 163
- ALEXANDER THOMAS and STELLA CHESS *Temperament and follow-up to adulthood* 168
Discussion 173
- M. BERGER *Personality development and temperament* 176
Discussion 187
- WILLIAM B. CAREY *Clinical use of temperament data in paediatrics* 191
Discussion 202
- RICHARD Q. BELL and MARY F. WALDROP *Temperament and minor physical anomalies* 206
Discussion 216
- S. N. WOLKIND and W. De SALIS *Infant temperament, maternal mental state and child behavioural problems* 221
Discussion 234
- MATTI O. HUTTUNEN and GÖTE NYMAN *On the continuity, change and clinical value of infant temperament in a prospective epidemiological study* 240
Discussion 247
- I. KOLVIN, A. R. NICOL, R. F. GARSIDE, K. A. DAY and E. G. TWEDDLE *Temperamental patterns in aggressive boys* 252
Discussion 265
- BARBARA K. KEOGH *Children's temperament and teachers' decisions* 269
Discussion 279

GENERAL DISCUSSION II Different approaches to the study of temperament and its disorders 286, Practical uses of temperament research 287, Measurement and categorization of temperament 289, Developmental issues 291, Temperament and family interactions 292

MICHAEL RUTTER Chairman's closing remarks 294

Index of contributors 299

Subject index 301

Participants

- J. E. BATES Department of Psychology, Indiana University, Bloomington, Indiana 47405, USA
- R. Q. BELL Department of Psychology, University of Virginia, Charlottesville, Virginia 22901, USA
- M. BERGER Department of Child Psychiatry, St George's Hospital, Blackshaw Road, London SW17 0QT, UK
- W. B. CAREY 319 West Front Street, Media, Pennsylvania 19063, USA (and Department of Pediatrics, University of Pennsylvania Medical School, Philadelphia, USA)
- S. CHESS Millhauser Laboratories, Department of Psychiatry, New York University Medical Center, School of Medicine, 550 First Avenue, New York, New York 10016, USA
- J. F. DUNN MRC Unit on the Development and Integration of Behaviour, University Sub-Department of Animal Behaviour, Madingley, Cambridge CB3 8AA, UK
- D. W. FULKER Animal Psychology Laboratory, Institute of Psychiatry, The Bethlem Royal Hospital, Monks Orchard Road, Beckenham, Kent BR3 3BX, UK
- P. J. GRAHAM Department of Child Psychiatry, Institute of Child Health, The Hospital for Sick Children, Great Ormond Street, London WC1N 3JH, UK
- R. A. HINDE MRC Unit on the Development and Integration of Behaviour, University Sub-Department of Animal Behaviour, Madingley, Cambridge CB3 8AA, UK

- C.-C. HSU Children's Mental Health Center, Department of Psychiatry, National Taiwan University Hospital, No. 1, Chang-Te Street, Taipei 100, Taiwan
- M. O. HUTTUNEN Department of Psychiatry, University of Helsinki, Helsinki, Finland
- B. K. KEOGH Moore Hall, Department of Education, University of California at Los Angeles, 405 Hilgard, Los Angeles, California 90024, USA
- G. A. KOHNSTAMM Department of Developmental Psychology, University of Leiden, Hooigracht 15, 2312 KM Leiden, The Netherlands
- I. KOLVIN Nuffield Psychology and Psychiatry Unit, Fleming Memorial Hospital, Great North Road, Newcastle Upon Tyne NE2 3AX, UK
- T. F. McNEILL Department of Psychiatry, University of Lund, Malmö Allmänna Sjukhus, S-214 01 Malmö, Sweden
- R. PLOMIN Institute for Behavioural Genetics, University of Colorado, Campus Box 447, Boulder, Colorado 80309, USA
- R. O. ROBINSON Newcomen Centre, Guy's Hospital, St Thomas Street, London SE1 9RT, UK
- M. RUTTER Department of Child and Adolescent Psychiatry, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, UK
- J. E. STEVENSON Department of Human Biology and Health, University of Surrey, Guildford, Surrey GU2 5XH, UK
- J. STEVENSON-HINDE MRC Unit on the Development and Integration of Behaviour, University Sub-Department of Animal Behaviour, Madingley, Cambridge CB3 8AA, UK
- A. THOMAS Millhauser Laboratories, Department of Psychiatry, New York University Medical Center, School of Medicine, 550 First Avenue, New York, New York 10016, USA
- A. M. TORGERSEN Department of Child Psychiatry, avd.18, Ullevål Hospital, Oslo 1, Norway

J. S. WERRY Department of Psychiatry, School of Medicine, University of Auckland P. B., Auckland, New Zealand

R. S. WILSON Department of Pediatrics, Child Development Unit, Health Sciences Center, University of Louisville, Louisville, Kentucky 40292, USA

S. N. WOLKIND Family Research Unit, The London Hospital Medical College, Turner Street, London E1 2AD, UK

Temperament: concepts, issues and problems

MICHAEL RUTTER

Department of Child and Adolescent Psychiatry, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, UK

Abstract There are marked individual differences in children's temperamental styles—differences thought to be constitutionally determined in part. The importance of temperamental features is evident in their links with various forms of psychopathology and in their effects on the manner in which other people respond to the child. For these and other reasons it has rightly come to be accepted that greater attention needs to be paid to temperamental issues in consideration of the processes of development, children's responses to stress situations, and the genesis of emotional, behavioural and learning disorders. However, major conceptual, methodological and theoretical problems remain. Problems of measurement are considered in terms of the relativity of measures, whether or not to take social context into account, the functional equivalence of measures at different ages, the circumstances to use in assessing temperament, the choice of measuring instrument and the categorization of temperamental features. The issues involved in the meaning of temperamental differences are discussed with respect to consistency, developmental change, genetic influences, brain damage and mental retardation, sex differences and the mechanisms by which temperamental variables exert their effects.

The general concept of *temperament* goes back to at least mediaeval times, when it was used to refer to a person's mental disposition, as determined by the combination of the four cardinal humours. Its usage today retains much of that emphasis. That is, the term usually implies reference to the basic elements of behavioural functioning rather than to complex or idiosyncratic aspects of a person's emotional or social style; the concept applies to those elements that show substantial consistency over time and over space; and it is assumed that, to a considerable extent, the elements are constitutionally determined. Most notably, temperament refers to a preponderant style in *how* an individual does things or how he or she responds to people and to situations, rather than to *what* the individual does (i.e. the content of behaviour), or to *why* he or she does it (i.e. motivation), or to the behavioural

1982 Temperamental differences in infants and young children. Pitman Books Ltd, London (Ciba Foundation symposium 89) p 1-19

capacities or abilities that he or she manifests (Thomas & Chess 1977). Temperamental characteristics are abstractions, rather than directly observable discrete behaviours (Rutter et al 1964), and the questions with which we need to start therefore concern the empirical evidence that justifies the abstraction and shows its utility.

Although the concepts and terms are centuries old, the scientific study of temperamental attributes began much more recently. Gesell's (1937) analysis of film records of children—to assess characteristics such as activity level or energy output, adaptability, and liveliness of emotional expression—constituted one of the earliest studies. He concluded that . . . 'certain fundamental traits of individuality, whatever their origin, exist early, persist late and assert themselves under varying environmental conditions'. This view contrasted starkly with the prevailing popular impression that all young babies were much alike and that individual differences in early childhood were of little importance. Over the next two decades, a steady trickle of studies indicated important individual differences in a wide range of behavioural functions in infancy and early childhood. Then, a most important stimulus to research on temperament came from the New York longitudinal study started by S. Chess, A. Thomas and H. G. Birch in the mid-1950s. One of their early papers (Chess et al 1959) included the statement . . . 'We believe that the data indicate that the individual specific reaction pattern appears in the first few months of life, persists in a stable form thereafter, and significantly influences the nature of the child's response to all environmental events, including child care practices'. In subsequent years they somewhat modified their claims regarding temporal consistency, but the findings from their longitudinal study (see Thomas & Chess 1977) have otherwise generally supported their initial hypotheses on the importance of individual differences in temperamental characteristics. Moreover, as shown by a variety of recent reviews (Rutter 1977, Dunn 1980, Keogh & Pullis 1980, Bates 1980)—as well as by the papers in this symposium—the claims and findings from the New York study stimulated a mass of research by other investigators. We are consequently in a much stronger position now than we were 20 years ago to assess the concept of temperament. In reviewing what we know, we need to sharpen the hypotheses and clarify the issues in order to consider the utility of the current notions, and to identify the empirical questions and theoretical problems that remain to be tackled.

The utility of the concept of temperament

The empirical evidence that points to the utility of the concept of temperament may be considered under three main headings: (a) the demonstration of

behavioural individuality; (b) the findings on constitutional determinants; and (c) the predictive power of temperamental measures for children's development of disorder and for children's responses to potentially stressful situations.

Behavioural individuality

The question of behavioural individuality can be quickly dealt with in view of the entirely consistent and extensive data showing that, far from all infants being the same, babies and young children differ strikingly in their behavioural characteristics (Buss & Plomin 1975, Rutter 1977, Bates 1980, Dunn 1980, Keogh & Pullis 1980). This has been shown, for example, for features as varied as activity level, autonomic reactivity, fussiness or irritability, soothability, visual alertness, regularity of sleep-wake patterns, adaptability to change and social responsiveness. Moreover, numerous investigators have shown that these characteristics can be measured reliably by questionnaire, interview, and by observational and mechanical means.

Constitutional basis

The question of the 'constitutional' basis of these individual differences raises more complex matters. The utility of the concept of temperament does not depend on the demonstration that it has its origins in genetically determined or any other kind of constitutional factors. If consistent individual differences in temperamental style were found, and these differences reliably predicted how children would respond to stressful situations, or how their course of later personality development would proceed, or whether they were likely to develop psychiatric problems, it would not matter if it were shown also that a child's temperament was largely shaped by early life experiences. In that case, temperament would still be viewed as a relatively enduring individual characteristic that reflected the personal qualities that a child brought to any new experience he or she encountered. The behavioural features would be 'constitutional' in the sense that they reflected an intrinsic aspect of personal functioning, even though their origins were experiential rather than genetic. Nevertheless, most workers have considered it important to claim and to demonstrate that temperament has a substantial genetic component.

The evidence on this point is limited and not free from difficulties. Even so, the empirical findings from twin studies (see Buss & Plomin 1975, Torgersen & Kringlen 1978, Matheny 1980, Goldsmith & Gottesman 1981) are reasonably consistent in showing that genetic factors play a significant part in

individual variability for at least some temperamental features. This is most evident, perhaps, for activity level and task orientation, but it applies in varying degree to many other characteristics.

Another key aspect of the 'constitutional' notion concerns the consistency of temperamental functioning over time and space. If a feature is to be considered 'constitutional', it might be expected that the individual should exhibit the attribute in similar fashion across a range of different situations and over lengthy periods of time. In fact, this supposition is not quite so self-evident as it appears at first sight (a point to which I will return), but the evidence on consistency is still of some relevance. Several longitudinal studies have examined the question of temporal stability (Buss & Plomin 1975, Thomas & Chess 1977, Moss & Susman 1980) and found that children's temperamental style shows substantial consistency over periods of several months up to a year or so, but that correlations extending over several years between the pre-school period and middle childhood are generally quite low. The question of consistency across situations is at least as important, but there is surprisingly little evidence on the matter. There is moderate, but not high, agreement between measures based on parental reports, teacher reports and direct observations (Dunn & Kendrick 1980, Billman & McDevitt 1980, Buss et al 1980). The overall evidence therefore suggests that temperamental characteristics show some features associated with 'constitutional' variables.

Prediction to other aspects of functioning

The evidence considered so far partly substantiates the notion that the abstraction of temperament has some validity, but it does not attest to its utility. That is shown by the power of temperamental measures as predictors of how children are likely to respond in various situations. It is that power that provides the main reason for regarding temperament as a crucial variable for any research in the field of developmental psychopathology.

The relevant evidence falls under three main headings. Firstly, it has been shown that children with amodal temperamental characteristics on such variables as irregularity of functioning, negative mood, non-adaptability, high emotional intensity, and low fastidiousness have a substantially increased risk of developing emotional or behavioural disorders during the next few years (see reviews by Rutter 1977, Keogh & Pullis 1980). This has been found in groups of children that might be considered both low risk (Rutter et al 1964, Thomas & Chess 1977) and high risk (Graham et al 1973) in terms of their psychosocial situation. There is also evidence linking temperamental attributes with scholastic performance (see Keogh & Pullis 1980).

Secondly, it has been found that a child's temperamental features consti-

tute an important predictor of how that child will respond to the birth of a sibling. Dunn et al (1981) found that high emotional intensity and a tendency towards negative mood were the two temperamental features associated with adverse reactions. Little is known about the importance of temperamental variables with respect to children's responses to other potentially stressful life situations or to changes of circumstances requiring adaptation of some kind, but there are various pointers to their probable relevance (see Rutter 1981a).

Thirdly, several naturalistic and experimental studies have shown that the behavioural characteristics of children have an important effect in determining how other people respond to them (see Rutter 1977, Dunn 1980). Children with different temperamental features *elicit* different behaviours from those with whom they interact. For example, in our study of families with a mentally ill parent, we found that children with adverse temperamental characteristics were twice as likely as other children to be the target of parental criticism (Rutter 1978). In contrast, easily adaptable children tended to be protected even in a stressful home environment, precisely because much of the hostility and discord was focused on other members of the family. Similarly, it has been found that weak malnourished children receive less parental attention than well-fed children and, probably because they elicit more caretaking, highly active babies are less likely to show developmental retardation in a depriving institutional environment (see Rutter 1977). Dunn found, in addition, that children's reactivity in infancy was linked with maternal responsiveness to them a year later (see Dunn 1980).

The temperamental qualities brought by a child to the interactions and situations that he or she encounters therefore play an important part in determining how that encounter proceeds and whether it is likely to result in the development of some form of maladaptive response or emotional-behavioural disturbance. The empirical evidence on those points provides ample justification for regarding the concept of temperament as theoretically important and practically useful.

Problems in measurement

The most immediate issue that bedevils the study of temperament concerns the question of how to measure temperament—not necessarily which particular instrument or measuring device to use but, rather, how to conceptualize and operationalize the characteristics to be assessed.

Relativity

For example, some problems stem from the use of relative rather than

absolute measures. Nearly all instruments (other than strictly quantitative bits of gadgetry, such as actometers or stabilimeters to assess activity level) compare children with other children on some behavioural feature. For example, the parental questionnaire developed by the New York group (Thomas & Chess 1977) includes items such as 'my child splashes hard in the bath and plays actively' or 'if my child is angry or annoyed, he gets over it quickly' or 'my child is highly sensitive to changes in the brightness or dimness of light'. Adverbs such as 'hard', 'actively', 'quickly' or 'highly' all demand a knowledge of what is the 'norm'. Almost inevitably, a person's norm consists of other children of the same age that he or she happens to know. The consequence is that little confidence can be placed on comparisons across ages or across sociocultural groups. A lack of difference may simply mean that the person who is rating the child has adjusted the ratings to a different set of norms.

On the other hand, for some purposes, a relative measure may be more useful, provided that there is a known and constant norm. Thus, age-standardized Intelligence Quotient scores are used to remove the massive effects of age. As Kagan (1980) points out: 'When text books say that children's intelligence is stable from 5 to 10 years of age, they do not mean that cognitive ability is stable; it is not. They mean that the differences in test scores among a cohort of children remain stable, despite dramatic changes in the abilities that accompany growth'.

Social context

Whether or not to take social context into account constitutes a further variant of the same issue. For example, which is the higher activity level—wandering around aimlessly on the wing in a soccer match or fidgeting and squirming vigorously in a chair in front of the television set? The first may well involve greater energy output, but that is a function of the demands of the situation rather than of how the child responds to it. However, the situation may well reflect the child's *choice* of activities that suit his or her temperamental style. Precisely how to deal with social context, within a single narrow age span, is not obvious; it is even more difficult across age periods—the content of the activities of a 14-week-old baby bear little resemblance to those of a 14-year-old adolescent.

Functional equivalence

A third question concerns the issue of functional equivalence. Crying in a six-month-old and crying in a 16-year-old, for example, are both real

reflections of emotional expression. But do they have the same functional meaning, and is it sensible to regard correlations between the two as if they were reflections of continuity or discontinuity in development? Conversely, because behaviours appear different in form, does it follow that in functional (or genotypic) terms they are dissimilar? Obviously not (see Kagan 1980, Moss & Susman 1980)—developmental changes may modify or alter the particular manner in which a characteristic is manifest. But how does one determine what is functional equivalence for temperamental features?

Situations for measurement

A further issue stems from the question of which situations or circumstances to use when attempting to assess temperament. Is it preferable to measure an individual's temperamental qualities in terms of his or her response to new situations or new demands or, rather, is it better to assess them in terms of his or her behaviour in more habitual circumstances? Of course, some characteristics (such as adaptability) can be assessed only in relation to new situations because it is the nature of the child's response to novelty that reflects the characteristic. But that does not necessarily apply to other characteristics (such as emotional intensity). On the other hand, the routine of day-to-day activities may reduce the opportunity for individuality to be shown. But how does one decide what is novel? Is an outing to Hyde Park novel because previous outings have been to St. James' Park or do the park-like qualities of each make them comparable and hence non-novel? Is the coming of a different baby-sitter a new experience, or does it matter whether the child has previously had one or 43 baby-sitters?

Measuring instrument

Having decided on the answers to these questions, one faces the further issue of what sort of measuring instrument to use—should it be direct observation, questionnaire, interview or some form of mechanical measurement? The crucial point here concerns the need to sample behaviour over a wide range of situations and over a reasonably long span of time. Obviously, it would be pointless to attempt to assess adaptability or regularity of functioning on a single episode, because the very definition of the variable requires knowledge of a person's behaviour over time.

It was this consideration that led the New York group to choose the parent as the observer. They argued that the parent represented a potential source of

extensive direct observations of the child over many situations and over a prolonged time and, hence, that if this experience could be adequately assessed it would constitute a rich and economical source of information on temperamental style. They, and those of us who have utilized interview measures (see Graham et al 1973, Thomas & Chess 1977), have relied on careful questioning about actual behaviour (rather than generalizations about supposed traits), and on detailed descriptions of specific examples in order to get an account of what the child *actually* does—an account that is as free from the biases of selective reporting or attitudinally influenced pre-judgements as skilled questioning can make it. There is no doubt that this method provides vivid and detailed accounts of temperamental style which give rise to measures that reflect individuality in the way intended.

There are, however, three main queries regarding interview assessments. The first concerns the validity of the reports. Do the differences in temperament reported by mothers simply reflect biased perceptions of their children? Dunn & Kendrick (1980) have investigated this issue, comparing mother's and observer's ratings and contrasting each with detailed observational measures. On the whole, the results are reassuring in that they indicate satisfactory validity for most variables. It seems most *unlikely* that the differences are solely a function of biased perceptions, but more data are required before firm conclusions should be drawn.

The second query concerns the fact that the mother's information is necessarily limited to situations when she has been present. With older children, this is likely to mean that she has not observed behaviour in many of the situations of most interest. With pre-school children, of course, this is not as much of a problem, provided that the mother is the main caretaker (but it would be a problem with children in full-time day-care). However, to an important extent, everyone's behaviour is relatively situation-specific (see Mischel 1979, Bem & Funder 1978, Epstein 1979). The reported behaviour of the child may therefore be influenced as much by the pressure stemming from the family interaction patterns as by any qualities intrinsic to the child. The only satisfactory solution to that problem is to tap the child's behaviour by interviewing more than one informant. The third query stems from the fact that this form of detailed interviewing with cross-questioning and the eliciting of multiple examples is immensely time-consuming. Could not the whole process be streamlined by using a standard questionnaire?

Several investigators have developed questionnaires that show as much individuality as do the interview measures, with moderately satisfactory re-test reliability and with reasonable agreement between the scores on questionnaires completed by mothers and by fathers (see e.g. Persson-Blennow & McNeil 1979, Bates et al 1979). So far, so good—but there remains the major concern about whether questionnaires are more prone

than interviews to the effects of perceptual or attitudinal distortions. The data on this point are sparse indeed but Bates et al (1979) found not only a low level of agreement between questionnaire measures and observational measures but also that the questionnaire measures showed associations with maternal characteristics whereas the observational measures did not. As this group points out, their finding is open to several different interpretations, depending in part on the weight attached to the observational data. The question of possible biases in questionnaire completion remains open and the matter requires further study.

Observational measures constitute a third possibility for the measurement of temperament. In some respects, they are more objective than either interview or questionnaire measures but they are the most time-consuming of all methods and are severely limited in their scope. By their nature, they can sample only one situation at a time and only certain sorts of situations at that. Because there are important situation-specificities in people's behaviour, it is only by taking *multiple* samples of behaviour that the investigator can maximize the chances of obtaining a valid appraisal of general response dispositions or temperamental attributes (Epstein 1979). For this reason, although observational data constitute valuable checks on interview and questionnaire measures, they cannot necessarily be assumed to constitute validating criteria and it is unlikely that they will prove to be the best general means of assessing temperamental features.

Categorization of temperamental features

Several questions need to be raised here but I shall focus on just three. The first is whether to concentrate on variations within the normal range or to focus on the few individuals showing extreme patterns. In an attempt to improve the psychometric properties of temperamental measures some investigators have rejected items applying to only small proportions of the population (Persson-Blennow & McNeil 1979) but perhaps it is precisely those extremes that predict best. The possibility requires examination.

The next two questions—how far to reduce the separate temperamental variables to a smaller number of factors or summary measures, and what criterion to use in deciding this matter—are most conveniently considered together. Different investigators have given different answers. Garside et al (1975) argued that the intercorrelations between the items (as reflected in a principal components analysis) should constitute the criterion. That is a statistically tidy solution, but is it sensible? I would suggest that it is not, for three rather different reasons: (a) especially with a questionnaire the intercor-

relations are strongly influenced by the particular constructs (organization of ideas regarding behaviour) used by the rater; (b) the specific factors obtained tend to vary according to the statistical method employed and the sample studied (for example, the New York group derived factors of strikingly different composition); and (c) the fact that particular items intercorrelate with each other in a normal population may have little bearing on the grouping of items that predict best how an individual will respond to stress situations, or indeed to any other kind of outcome variable such as cognitive impairment, learning difficulties, or the development of psychiatric disorders (see e.g. Dunn et al 1981, Rutter et al 1964, Graham et al 1973, Hertzig 1982).

An alternative approach is to consider each of the many possible temperamental variables separately. This method has been followed by most investigators, at least as the first step in statistical analysis. It has paid off in so far as individual temperamental attributes have proved useful predictors for various sorts of outcome (see the last group of references cited above). On the other hand, in several investigations particular clusters or patterns of characteristics have proved much better predictors than any single characteristic.

This repeated finding has led to a third approach—namely, the development of various composite measures such as the ‘temperamental adversity’ index (Rutter 1978), the ‘difficult child’ index (Thomas Chess 1977, Hertzig 1982) and ‘temperament risk’ scores (Cameron 1978). Again, this approach has proved useful for predicting the development of emotional-behavioural problems. Nevertheless, it would be quite premature to regard any of these measures as having solved the question of how to categorize temperament. In the first place, although most of the composite measures are broadly similar, they differ in certain crucial respects and it remains uncertain which is superior. Secondly, empirical findings show that it is not always the same temperamental features that are crucial. Thus, for example, Schaffer (1966) found that high activity level was protective in a depriving institutional environment; Dunn et al (1981) found mood variables to be the best predictors of children’s responses to the birth of a sibling; and Graham et al (1973) found irregularity, lack of malleability and low fastidiousness to be the attributes associated with the development of emotional-behavioural disorder in the children of mentally ill parents. It should *not* be assumed that the categorization of temperament that proves most effective for one purpose or in one situation will be equally effective in others. Nor should it be assumed that composite scores will always be more powerful than single variables (although it is likely that this will often be so). Nevertheless, the strategy of considering the categorization of temperament in terms of the groupings that best predict different sorts of outcome does seem worthwhile.

Implications of temperamental differences

Consistency

The question of consistency, of course, extends far beyond the issue of temperamental attributes in terms of the vigorous clashes over the last dozen years between trait theories and situationism theories. On the one hand, it has been argued that people's behaviour is highly inconsistent over time and place, being largely determined by situational factors (Mischel 1968). On the other, it has been asserted that genetically determined, semi-permanent personality dispositions play a major part in ensuring that people *do* behave consistently (Eysenck 1970). It is now clear that both extremes in these views must be rejected. The importance of personality traits is shown by the great individual variation in people's responses to any one situation, but the need to invoke environmental determinants is equally evident in the extent to which any person's mode of functioning alters from situation to situation (see Bem & Funder 1978, Epstein 1979). However, it is not sufficient to regard both the traits and the situations as important; a further question concerns the extent to which the two interact predictably (see Bem & Funder 1978). It would be misleadingly limiting to regard temperament as the reflection of the degree of similarity in a person's behaviour across all situations. Rather, some of the key aspects of temperament may concern the degree to which a person can adapt or modify his or her behaviour according to different environmental demands, or a person's vulnerability to certain kinds of stressors, or a tendency to respond in an unusual way to specific environments. These various kinds of ordinal and disordinal interaction effects* have been little explored up to now but the knowledge that they may be important in relation to temperamental variables has serious implications for the way in which we both assess temperament and analyse its effects.

Developmental change

The first reports from the New York study (Chess et al 1959) claimed that temperamental patterns were established in the first few months of life and remained stable thereafter. However, these workers' own empirical findings

*Ordinal interaction means that one variable (e.g. temperament) influences the *degree* to which a person responds to another variable (e.g. some environmental stressor) without altering the direction or type of response. In contrast, disordinal interaction means that the first variable alters the *direction* of response, so that some people react in one manner whereas others show an opposite response to the same second variable.

(Thomas & Chess 1977), as well as those of others, have been consistent in showing near-zero correlations from the first year of life to age five years onwards. There are various reasons why early infancy measures are likely to show little continuity with measures in later childhood (Rutter 1970). It should be noted especially that temperamental attributes in the first few months of life tend to show less of a genetic component than do the same attributes in later infancy or early childhood (Torgersen & Kringlen 1978, Matheny 1980); and whereas temperamental variables at age three or four years have generally shown significant associations with psychiatric risk, those in the first year of life often have not (see Rutter et al 1964, Cameron 1978). Nevertheless, temperamental attributes, even at age two, three or four years have usually been found to correlate at a very modest level ($r \approx 0.3$) with those assessed in middle childhood. Where does that leave the notion of an enduring temperamental style? Of course, life experiences of various kinds are likely to play their part in shaping temperament and these influences will, necessarily, reduce consistency over time. But the direction and degree of developmental change themselves may be genetically conditioned—as shown by the finding from the Louisville twin study (Matheny & Dolan 1975, Wilson 1977). In addition, as noted already, phenotypic expression may alter its form over the course of development so that the simple correlation of like behaviour with like behaviour may not constitute the most appropriate test for temperamental consistency.

Genetic influences

Up to now, the evidence regarding genetic influences on temperament has stemmed from a small number of twin studies, each based on rather different measures (not all of which have been entirely adequate for the assessment of temperament). However, there are particular problems in the use of twins for temperamental studies if interview or questionnaire measures are used. Because the measures are relative, because many parents like to emphasize the individuality of each twin, and because the other twin tends to constitute the main comparison for ratings, it is unlikely that the degree of difference within twin pairs (monozygotic or dizygotic) will be on the same scale as that between siblings or between unrelated children. There is a great need to utilize other genetic designs in addition to the twin method but, whatever the design, it will be important to examine the possibility that the genetic contribution concerns the overall *pattern* of temperamental functioning as much as the variation on each separate attribute.

Brain damage and mental retardation

Although some published reports suggest that brain damage or mental retardation may be associated with a distinctive pattern of temperamental functioning, there has been little systematic study of the matter (see Thomas & Chess 1977). It is apparent already that both retarded and brain-damaged children exhibit as wide a variety of temperamental characteristics as do normal children, and that few differences are found when individual temperamental features are considered. However, Hertzig's (1982) longitudinal study of low-birthweight infants has shown a significant association between abnormalities on a neurological examination and the 'difficult child' temperamental pattern. It is well established that brain damage and mental retardation carry a markedly increased psychiatric risk, and the possibility that this increased risk is due in part to an effect on the pattern of temperamental functioning warrants further study.

Sex differences

Numerous studies have shown that psychiatric disorder is substantially more common in boys than girls; also it has been found repeatedly that boys are more susceptible than girls to emotional-behavioural problems in association with family discord and disruption (Rutter 1981b). The reasons for this greater vulnerability of boys remain rather obscure but it would seem plausible that sex-linked temperamental differences might play a part. Accordingly, it seems surprising that most investigations have reported few, if any, significant sex differences in temperamental variables. However, the negative findings may be a consequence of looking at each attribute separately rather than examining composite scores associated with psychiatric risk. The question of possible sex differences in overall temperamental pattern requires further study.

Modes of operation of temperament

Perhaps the most fundamental question concerns the manner and mechanisms by which temperamental variables exert their effect. As we have seen, temperamental attributes are important predictors not only of how children respond to 'stress' situations but also of the likelihood that they will develop various types of emotional and behavioural disorder. But what do these statistical associations mean in terms of underlying mechanisms, and why are children with particular temperamental patterns more at risk than those with

other temperamental features? Cameron (1978) used the geological metaphor of temperament reflecting 'fault lines' in the emerging personality so that behavioural 'earthquakes' arise in those children with 'fault lines' who experience environmental strains. However, this analogy seems both inapt and unhelpful. Why should particular temperamental patterns be regarded as 'faulty' and in what way do they put the child at risk? Dunn (1980) and Rutter (1977) suggest that children's temperamental differences may influence development through several mechanisms—including effects of these differences on how other people respond to and interact with the child; the shaping of life experiences; the determining of what is an *effective* environment for the child; the reflection of the child's social adaptive capacity i.e. malleability and adaptability in responding to altered environmental circumstances; and the rather imprecise concept of psychological vulnerability. There is a certain amount of empirical support for each of these mechanisms but the issues remain little explored up to now. We do not know, for example, how temperamental variables relate to psychophysiological measures such as autonomic reactivity or to styles of coping with stress situations.

Finally, we have only a limited understanding of the role of temperament in personality development. We use the term 'temperament' in childhood rather than 'personality' and we talk of the importance of temperamental variables in shaping the emerging personality, but just what does that mean? What is the relationship between temperamental attributes such as intensity of emotional expression and personality variables such as extraversion and neuroticism? Do the differences in terminology merely reflect differences in the concepts of the workers who introduced these various terms or are there basic developmental processes and changes that the different terms reflect?

Conclusions

As this brief review has emphasized, the last decade has seen a burgeoning of interest in temperament. There has been an accompanying substantial growth in our knowledge and understanding of the importance of temperamental differences. Temperament constitutes a variable of considerable predictive power in developmental psychopathology, a power with both practical and theoretical implications. But many difficult issues and problems still require resolution.

REFERENCES

- Bates JE 1980 The concept of difficult temperament. *Merrill-Palmer Q* 26:299-319
 Bates JE, Freeland CAB, Lounsbury ML 1979 Measurement of infant difficultness. *Child Dev* 50:794-803

- Bem DJ, Funder DC 1978 Predicting more of the people more of the time: assessing the personality of situations. *Psychol Rev* 81:485-501
- Billman J, McDevitt SC 1980 Convergence of parent and observer ratings of temperament with observations of peer interaction in nursery school. *Child Dev* 51:395-400
- Buss AH, Plomin RA 1975 A temperament theory of personality development. Wiley Interscience, New York
- Buss DM, Block JH, Block J 1980 Preschool activity level: personality correlates and developmental implications. *Child Dev* 51:401-408
- Cameron JR 1978 Parental treatment, children's temperament, and the risk of childhood behavioral problems. 2: Initial temperament, parental attitudes, and the incidence and form of behavioral problems. *Am J Orthopsychiatry* 48: 141-147
- Chess S, Thomas A, Birch HG 1959 Characteristics of the individual child's behavioral responses to the environment. *Am J Orthopsychiatry* 29:791-802
- Dunn J 1980 Individual differences in temperament. In: Rutter M (ed) *Scientific foundations of developmental psychiatry*. Heinemann Medical, London, p 101-109
- Dunn J, Kendrick C 1980 Studying temperament and parent-child interaction: comparison of interview and direct observation. *Dev Med Child Neurol* 22:494-496
- Dunn J, Kendrick C, MacNamee R 1981 The reaction of first born children to the birth of a sibling: mothers' reports. *J Child Psychol Psychiatry Allied Discip* 22:1-18
- Epstein S 1979 The stability of behaviour. 1: On predicting most of the people much of the time. *J Pers Soc Psychol* 37:1097-1126
- Eysenck HJ 1970 *The structure of human personality*. Methuen, London
- Garside RF, Birch HG, Scott DM, Chambers S, Kolvin I, Tweddle EG, Barber LM 1975 Dimensions of temperament in infant school children. *J Child Psychol Psychiatry Allied Discip* 16:219-232
- Gesell A 1937 Early evidences of individuality in the human infant. *Sci Monthly* 45:217-225
- Goldsmith HH, Gottesman II 1981 Origins of variation in behavioral style: a longitudinal study of temperament in young twins. *Child Dev* 52:91-103
- Graham P, Rutter M, George S 1973 Temperamental characteristics as predictors of behavior disorders in children. *Am J Orthopsychiatry* 43:328-339
- Hertzog ME 1982 Temperament and neurologic status. In: Rutter M (ed) *Behavioral syndromes of brain dysfunction in childhood*. Guilford Press, New York, in press
- Kagan J 1980 Perspectives on continuity. In: Brim OG, Kagan J (eds) *Constancy and change in human development*. Harvard University Press, Cambridge, Mass., p 26-74
- Keogh BK, Pullis ME 1980 Temperament influences on the development of exceptional children. *Adv Spec Educ* 1:239-276
- Matheny AP 1980 Bayley's infant behavior record: behavioral components and twin analyses. *Child Dev* 51:1157-1167
- Matheny AP, Dolan AB 1975 Persons, situations and time: a genetic view of behavioral change in children. *J Pers Soc Psychol* 32:1106-1110
- Mischel W 1969 *Personality and assessment*. Wiley, London
- Mischel W 1979 On the interface of cognition and personality: Beyond the person-situation debate. *Am Psychol* 34:740-754
- Moss HA, Susman EJ 1980 Longitudinal study of personality development. In: Brim OG, Kagan J (eds) *Constancy and change in human development*. Harvard University Press, Cambridge, Mass., p 530-598
- Persson-Blennow I, McNeil TF 1979 A questionnaire for measurement of temperament in six-month old infants: development and standardization. *J Child Psychol Psychiatry Allied Discip* 20:1-14
- Rutter M 1970 Psychological development—predictions from infancy. *J Child Psychol Psychiatry Allied Discip* 11:49-62

- Rutter M 1977 Individual differences. In: Rutter M, Hersov L (eds) *Child psychiatry: modern approaches*. Blackwell Scientific Publications, Oxford, p 3-21
- Rutter M 1978 Family, area and school influences in the genesis of conduct disorder. In: Hersov LA et al (eds) *Aggression and anti-social behaviour in childhood and adolescence*. Pergamon Press, Oxford, p 95-113
- Rutter M 1981a Stress, coping and development: some issues and some questions. *J Child Psychol Psychiatry Allied Discip* 22:323-356
- Rutter M 1981b Epidemiological-longitudinal approaches to the study of development. In: Collins WA (ed) *The concept of development*. (Proc 15th Minnesota Symp on Child Psychol) Erlbaum, Hillsdale, NJ, in press
- Rutter M, Birch H, Thomas A, Chess S 1964 Temperamental characteristics in infancy and the later development of behavioural disorders. *Br J Psychiatry* 110:651-661
- Schaffer HR 1966 Activity level as a constitutional determinant of infantile reaction to deprivation. *Child Dev* 37:595-602
- Thomas A, Chess S 1977 *Temperament and development*. Brunner/Mazel, New York
- Torgersen AM, Kringlen E 1978 Genetic aspects of temperamental differences in infants. *J Am Acad Child Psychol* 17:433-444
- Wilson RS 1977 Mental development in twins. In: Oliverio A (ed) *Genetics, environment and intelligence*. North-Holland Publishing, Amsterdam, p 305-334

DISCUSSION

Werry: You discussed only briefly the issue of the difference between temperament and personality. While it may be true that the study of children's temperament is relatively recent, that is not true of personality, for which sophisticated methods of analysis have long been available. What do you see as the fundamental difference between temperament and personality? Is temperament simply another name for personality when applied to children?

Rutter: I'm not sure, but it is striking that the two bodies of published work scarcely overlap or cross-reference. Temperament has tended to focus on rather specific aspects of behavioural style, rather than on general concepts such as neuroticism, extraversion and the like. Undoubtedly the two approaches should be brought together.

Werry: Many people believe still that personality is, primarily, constitutionally determined. But even if one eschews the question of aetiology, I think that temperament and personality are the same thing and that we are serving up old wine in new bottles. There is resistance to the concept of personality as a stable thing in children, who are supposed to be malleable, and thus a different term has been coined.

Thomas: One of the problems is what we mean by 'personality'. Although there is some consistency about what temperament means, for every 10

people involved in a discussion about personality one can obtain 15 definitions. There is a wide difference between the constitutional approach to personality and the psychodynamic approach, and that is where the problem lies. We ought to reach a consensus about the meaning of personality so that we can study it in relation to temperament.

Berger: I believe that the issues involved in definitions of personality and of temperament are much more complex than that, as I shall be discussing later in my paper (p 176-190).

Hsu: Cross-cultural studies sometimes give useful clues to answering the issues that Professor Rutter has raised about understanding temperament (see also p 113-119).

Rutter: There are, of course, opportunities to examine cultural differences even within the same country. Jerry Kagan's work on day care (Kagan et al 1978) showed striking temperamental differences between Chinese-American children and white American children. That result is interesting because the measures used in the two groups were the same, and one can be reasonably confident that they meant the same thing. The origins of those differences remain obscure, but I agree that cross-cultural studies are potentially very useful.

Hsu: The validity of studying sufficient numbers of comparable groups of children in different countries, with standardized instruments (i.e. methods of measurement), has already been established by Wolff (1973) as a means of highlighting the constitutional origin of neonates' reactions to alcohol. For instance, by using the same method, we may be able to see the similarities and dissimilarities between Chinese infants born and raised in the United States and in Taiwan.

Thomas: You mentioned, Professor Rutter, that boys surprisingly showed no significant differences in temperament from girls (Thomas & Chess 1977), even though the two sexes differ in rates of behavioural disorder. A reasonable hypothesis to explain this would be that even if the temperaments of boys and girls are similar, the reactions of people in the environment—parents and others—to the same temperamental characteristics in boys and in girls may be very different. This reaction may be crucial in accounting for the difference.

Rutter: Indeed it may. I raised the issue because of the striking disparity between the highly consistent differences between boys and girls, in almost any measure of psychopathology, and the contrasting responses of boys and girls to stress situations. A variety of behavioural measures also highlight striking differences between the sexes. It seems curious that, in most studies, temperamental differences between boys and girls are, on the whole, small. I remain unconvinced, however, that the different responses of people to boys and to girls is wholly a matter of cultural expectations rather than responses

'shaped' by actual differences in the *characteristics* of boys and girls. Some studies (Bell & Carver 1980, Smith & Lloyd 1978) have examined mother's responses to babies, in an attempt to differentiate the effects of cultural expectation and of the impact of sex differences in the babies' behaviour. The ingenious experimental strategy used for this purpose involved getting mothers to play with male and female babies presented in either sex-appropriate or cross-sex clothes and names. In this way it was possible to determine whether the mothers' style of interaction was most influenced by the perceived (but wrong) or the actual (but not perceived) sex of the child. The findings are not wholly consistent but they seem to demonstrate both expectational differences and differences that stem from the actual behaviour of the child. But this field has only just begun to be explored.

Wilson: In opposite-sexed twin-pairs that we have observed, there is a clear difference between the boys and the girls by the age of two years: the boys are generally more physically active, and this cannot be fully explained by parental expectation.

Hsu: In studies that have reported few sex differences, I suspect that the sample size was not large enough. In our study of children aged 3–7 years (Chen 1981) we found that in five of the nine temperamental categories described by Thomas & Chess (1977)—activity, intensity, mood, persistence, and distractibility—there were significant sex differences (see Table 1).

TABLE 1 (Hsu) Sex and temperament scores of young children

<i>Temperament categories</i>	<i>Male children (n = 997)</i>		<i>Female children (n = 934)</i>		<i>t value</i>
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	
Activity	4.06	0.83	3.74	0.87	8.27**
Rhythmicity	4.47	0.79	4.45	0.78	0.57
Approach/Withdrawal	4.50	0.86	4.44	0.83	1.56
Adaptability	4.84	0.78	4.80	0.76	1.14
Intensity	3.87	0.75	3.78	0.80	2.55*
Mood	4.76	0.62	4.83	0.64	-2.44*
Persistence	3.98	0.62	3.92	0.59	2.18*
Distractibility	4.20	0.71	4.28	0.69	-2.51*
Threshold	3.30	0.77	3.25	0.78	1.42

** $P < 0.01$; * $P < 0.05$; Student's *t*-test; degrees of freedom, 1930; $t_{0.95} = 1.96$; $t_{0.99} = 2.58$

Rutter: Sample size is certainly one vital feature, and your sample must be one of the largest available. A further issue is whether one is looking at the *relevant* variables.