# Temperament and Physiological Responses: Relationships Among Temperament Ratings, Galvanic Skin Response, and Heart Rate in Preschool Children

by

Donna M. Von Bargen

A thesis
presented to the University of Manitoba
in partial fulfillment of the
requirements for the degree of
Ph.D.
in
Psychology
June, 1986

ISBN 0-315-34017-7

# TEMPERAMENT AND PHYSIOLOGICAL RESPONSES: RELATIONSHIPS AMONG TEMPERAMENT RATINGS, GALVANIC SKIN RESPONSE, AND HEART RATE IN PRESCHOOL CHILDREN

BY

DONNA M. VON BARGEN

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

ABSTRACT

The concept of temperament has received much empirical interest in recent years, but has yet to be defined satisfactorily. Part of this definitional difficulty stems from the fact that while most assume that temperament has a constitutional basis, the validity of this assumption has not received much empirical attention. The primary goal of the present research was to explore possible physiological correlates of the temperament ratings as a way of validating this constitutional assumption. Other goals were to provide further information regarding the validity of three temperament measures by investigating relationships among them, and to investigate a possible age effect on the uniformity of responses from different physiological systems. Ninety-six preschoolers were rated by their parents on three temperament scales. Children´s physiological reactivity and habituation of galvanic skin response (GSR) and heart rate (HR) were also assessed in response to four stimuli (two auditory and two visual), each presented for six repetitions.

Canonical correlations between the measures of HR and GSR reactivity (first and maximum responses to each stimulus) revealed no significant relationships with temperament scores. Canonical correlations between the GSR and HR

habituation measures (the difference between the averages of the first three and the last three responses) and temperament scores indicated a relationship between temperament scores and children´s habituation scores for the last stimulus presented. Exploration of these results with multiple regression analyses revealed that the relationship with temperament scores existed only for GSR scores, not for HR scores. The temperament scores which were most strongly related to GSR habituation primarily assessed either reactivity/adaptation to novel stimuli or attention/activity level. Children whose GSR responses indicated less habituation were rated as being more shy, less approaching to new people or situations, slower to adapt to new situations, and reacting more to new foods. Children who habituated less were also rated as being more active, having shorter attention spans, and as being less regular in biological functioning.

The relationship between physiological measures and temperament scores was also explored using multivariate repeated measures ANOVA. For this analysis children were categorized into difficult, intermediate-high, intermediate-low or easy diagnostic cluster groups. Again, results were significant only for GSR scores and indicated that difficult children were more reactive to experimental stimuli than were children in the other groups. Difficult

children were also less likely to habituate to experimental stimuli than were easy children. These results were similar both within and across stimulus blocks. The group of easy children demonstrated the clearest habituation pattern.

Results generally indicated that HR and GSR measures were not related to each other. In addition, children´s age had no effect on the magnitude of the correlation between HR and GSR scores. This suggests that Lacey and Lacey´s (1958) response specificity theory may generalize to preschool aged children as well as to the adults and 10 year olds they studied.

Factor analysis of the 19 subscale scores from the three temperament measures resulted in five factors, accounting for 69% of the variability. These five factors were labeled emotionality, shyness/sociability, attention, activity level, and threshold of responsivity/distractability. Factors similar to these have emerged from numerous studies, and it seems that when researchers achieve a generally accepted definition of temperament which includes the subcategories of the phenomenon, most of these five factors will be represented in some form.

Overall, the results of this research indicated that physiological correlates of temperament ratings do exist for the GSR but not for the HR measures used. In particular, finding that difficult children´s GSR responses are larger

initially and that these initial GSR responses do not show a smooth habituation pattern is both theoretically and clinically important. Future research will be needed to explore whether HR results were not found because they do not exist or because they were obscured by methodological problems. Research will also be needed to explain why relationships between GSR scores and temperament ratings were found only for the last of the four stimuli and only when children were grouped according to their NYLS diagnostic cluster score. One possibility is that experimental power was reduced by the large number of variables included, and future investigators will be able to use the results of this research to reduce the number of variables in their projects.

Table of Contents

Temperament and Physiological Responses: Relationships Among Temperament Ratings, Galvanic Skin Response, and Heart Rate in Preschool Children

List of Tables

List of Figures

ACKNOWLEDGEMENTS

When I started university, I was a not-very-happily married 18 year old with a 6 month old boy. I took evenings and Saturday morning classes, and majored in psychology because high school guidance counselor seemed the most attractive of the careers modeled for me. The part I did not like about her job was the achievement and interest testing. I was tired of school and additionally stressed by the separation from my husband by the time I completed my B.A. in 1975, and not highly motivated to go on to graduate school immediately. However, I had become more sophisticated and realized that graduate school was necessary if I wanted to be a therapist. Knowing how difficult it was to get into the clinical psychology program, I made a deal with myself to apply. If accepted I would continue but take a break and reapply in a year if not accepted.

I was accepted, and gradually as I worked towards my Master´s degree I realized that "just" a Master´s degree would not let me do what I wanted either. If I wanted to be a therapist and not primarily a psychometrician, I had to get a Ph.D. This required a radical change in my self-concept. I thought of myself as a "kid from the sticks," and my family is not academically oriented. I was by far exceeding my and their goals for me by getting a B.A., never mind an M.A. It

took a lot of prodding from my mentor, Marvin Brodsky, before I began thinking of myself as someone who could be a doctor.

Well, here it is. This thesis is the last requirement for my Ph.D. I've completed all the courses, research, practica, and internship as well. The time, hard work and sacrifices are worth it to be able to do what I want to do. During my first practicum I recall feeling like "a fish in water" when doing therapy. One of the central values in my life is for people to achieve as much of their potential as possible. I had a lot of help achieving mine and now I can do the same for others.

There are so many who have helped me during the past 16 years that a complete listing and description of their good deeds would be longer than this dissertation. Consequently, this list is necessarily incomplete.

I figure that this feeling-oriented profession entitles me to some mushy sentimentality. I am married to the most wonderful man in the world. K. Michael Dresel and I have worked our way through graduate school together. Michael completed his Ph.D. in November, 1984. We have always maintained that doctorate degrees should be awarded jointly to spouses, and now that mine is done I feel we have each earned two Ph.D.'s . I want him to know how glad I am he put up with my inviting a talkative (even during movies) friend

along on our first date; moved to Wichita, Kansas without a job so I could complete my internship there; and has put up with my dissertation hysteria for longer than I care to remember.

My son, Richard Paul Dresel, is now 16. One more year of high school and he is off to university himself. He knows what it is like to have ordinary life put on hold while Michael and I prepare for exams, write term papers, and do research. During these difficult periods he was as helpful and considerate as one can expect any child to be. At less pressured times his presence helped me avoid graduate school tunnel vision and remember there was an enjoyable world outside Duff Roblin. Paul will receive a major graduation present in recognition of his sacrifices and contributions when we are both finally done.

Both my parents, Margaret and Myron (Red) Von Bargen, and Michael´s parents, Peter and Anita Dresel, have helped me through the ordeal called graduate school. It was the security of my parent´s love for me that enabled me to tackle university as well as motherhood. My interest in understanding other people stems directly from my mother who is endlessly fascinated with the people around her. Michael´s parents have believed in me and supported me as long as I have known them. Even when I got defensive about

their inquiries of how my thesis was going (because I was having trouble with the avoidance part of my approach-avoidance conflict), I knew they cared.

Others have been adopted by me in nurturing, parental roles. The current buzzword for people in positions like this is "mentor," but this term seems insufficient to describe the amount of support received from these individuals and the way I feel towards them.

Pat and Marvin Brodsky were particularly supportive and encouraging from the difficult period surrounding my separation and divorce until my marriage to Michael. I'll always remember the visits to their cottage, and I mourn the end of their marriage almost as much as I would if they were my own parents.

Bob Atkinson was my supervisor during the three years I was at the Wichita Guidance Center. I learned a lot from him, not the least of which was how to integrate the different parts of my life more completely. A typical task for women these days is how to have satisfying careers and families without in essence working two jobs (and enjoying neither due to exhaustion). Funny how we ended up discussing this while he was supervising my therapy with a multiple personality client . . . .

Joe and Helen Brewer are adopted grandparents. Both in

our Unitarian Universalist Church and the Guidance Center, they are revered by many. Joe was the director of WGC for over 30 years, as well as Bob´s supervisor, so the lineage from him to Bob to myself is clear. Joe and Helen listened to trials and tribulations as I struggled to finish my dissertation long distance from Wichita.

Warren Eaton deserves special thanks for teaching me, step by step, how to do research. Without his careful guidance during my Master´s thesis work I would not have known how to carry out this doctoral project so independently. I also want Warren to know how grateful I am that he was willing to replace another committee member who resigned from the committee at the 11th hour, leaving me in an extremely difficult position.

My gratitude to the following. I could always count on my friends and colleagues from the Wichita Guidance Center for support, especially Barbara Williams, Debi Trimmell-Martin, Dick and Karen Moore, Rita Goss, Sandra Mahoney, Ruby Brown, Kirk Miller, Nancy Milner, Mike McGreevy. Fellow and sister graduate students and other university folk who need to be mentioned are: Joy Thompson and Wayne Olin, Wenda and Greg Dickens, Janet Wright, Ranen and Louella Sinha, Chris and Lorne Sexton, Barry Spinner,

Debi Shelton, Phil Abrami and Valerie Turner, Bayley Rayter, Dave Dush, Gabrielle D´Amato, Wendy Josephson and Bob MacIllwraith, Dave Kalb, Ron Kruk, Henry Lee, Richard and Gerry Walsh. Scott Winning has gratiously set time aside in his schedule to help me with typing. Diane Benoit and Jill Pletcher also helped with typing. Bob Knapp and Charlie Birdsall provided me with computer time at The Wichita State University. St. Francis Hospital let me use their electrocardiograph monitor. Betty Gordon and all the teachers, parents, and children from the Friends´ University Preschool made this study possible.

Donna Von Bargen

Winnipeg

September, 1986.

# TEMPERAMENT AND PHYSIOLOGICAL RESPONSES: RELATIONSHIPS AMONG TEMPERAMENT RATINGS, GALVANIC SKIN RESPONSE, AND HEART RATE IN PRESCHOOL CHILDREN

Hippocrates thought that individual differences in health and emotional reactivity were determined by proportions of blood, phlegm, and black and yellow bile in the body (Klein, 1970). This ancient theory illustrates a basic assumption made about temperament: that individual differences in behavioral style are related to underlying physiological differences. Most modern definitions of temperament also assume this physiological or constitutional basis. For example, Allport (1937) used the following definition to differentiate temperament from personality:

> Temperament refers to the characteristic phenomena of an individual´s emotional nature, including his susceptibility to emotional stimulation, his customary strength and speed of response, the quality of his prevailing mood, and all peculiarities of fluctuation and intensity in mood, these phenomena being regarded as dependent on constitutional make-up, and therefore largely hereditary in origin. (1937, p. 54)

Much of the current interest in temperament stems from clinical rather than theoretical or experimental interests,

and investigators of this type tend to sidestep the assumption of constitutional basis. They define temperament as "behavioral style," which refers to the "how rather than the what (abilities and content) or the why (motivation) of behavior" (Thomas & Chess, 1977, p. 9). With the exception of this clinical school, all definitions offered for temperament clearly assume a physiological or constitutional basis for observed behavioral differences.

For example, Plomin (1982) stated that the following definition, offered at the 1980 Temperament Research Symposium, was acceptable to most attendees:

> Temperament involves those dimensions of personality that are largely genetic or constitutional in origin, exist in most ages and in most societies, show some consistency across situations, and are relatively stable, at least within major developmental areas. (Plomin, 1981, p. 269)

Another recent definition even more clearly stated that temperament is related to underlying physiological differences. Rothbart and Derryberry (1981) stated:

> we will define temperament as "constitutional differences in reactivity and self-regulation," with "constitutional" seen as the relatively enduring biological makeup of the organism

> influenced over time by heredity, maturation, and experience. By "reactivity" we refer to the characteristics of the individual's reaction to changes in the environment, as reflected in somatic, endocrine, and autonomic nervous systems. By "self-regulation" we mean the processes functioning to modulate this reactivity, e.g., attentional and behavioral patterns of approach and avoidance. (p. 37)

This definition is indicative of continued acceptance of this basic assumption. Unfortunately, while researchers generally agree on the importance of this assumption, empirical evidence supporting it stems almost without exception from heritability studies. While this body of research indicated that temperament has a genetic component, there is almost no research on how this genetic component is manifested physiologically. The present research represented an attempt to fill this gap by exploring some potential physiological correlates of children's temperament ratings. Before describing in detail the specific research area and problem studied in this dissertation, a general review of the child temperament literature is presented.

## A Review of Child Temperament Research

The term "temperament" is derived from temperare, a Latin verb meaning to mix, blend or regulate (Carey, 1981). Early

researchers proposed relationships between temperament and physiology (Kretschmer, 1925; Sheldon & Stevens, 1946). This early work on body type was primarily completed with adult subjects, and was criticized for limitations of theory, and for methodological confounds (Diamond, 1957). There was also some interest in individual differences in infant behavior (Gesell & Ames, 1937; Shirley, 1931, 1933a,b). These projects tended to be isolated attempts to counter the prevailing psychological theory of the time, environmental determinism.

In the early 1950's, Thomas, Chess, Birch and their colleagues began the New York Longitudinal Study (NYLS) with the goal of providing a more integrated approach to the study of temperament. As parents and as clinicians, the NYLS group had noticed two facts which prompted their research. The first was that great individual differences were observable from birth on, and the second was the inability to predict or account for pathology from a strictly environmentalist stance. They adopted a view, now taken for granted, that stresses environment-individual interactions or transactions (Sameroff, 1975; Sameroff and Chandler, 1975) in which neither the environment nor the individual alone determines developmental outcomes. Rather, the consonance or "goodness-of-fit" between the two is what determines adaptive or maladaptive functioning.

Since the first publications of this group in the early 1960´s (Chess, Thomas, Birch & Herzig, 1960; Thomas, Chess, Birch, Hertzig & Korn, 1963), there has been a proliferation of research on temperament, replicating the NYLS findings of a relationship between temperament and developmental problems, and exploring such areas as heritability and effects on caretakers. The main contribution of Thomas and Chess was that they focused on the effects of individual differences at a time when the popular view held that children contributed little or nothing to their own development. They also attempted to provide a comprehensive description of the main dimensions of temperament. At present there are many contrasting views of temperament, although the NYLS conceptualization still predominates.

The rest of this review has been organized into several major areas. First, temperament measures and the reliability of these measures are described. Then measures are considered in terms of their stability over time and their validity. Next, the research investigating the factor structure of temperament measures is presented. Finally, studies demonstrating the heritability of temperament and constitutional correlates of temperament are reviewed.

## Description and Reliability of Temperament Measures

The NYLS group originally used a semi-structured parental interview which assessed a wide variety of the child´s recent

behavior in environmental context. Based on "an inductive content analysis" of data from 22 children (Thomas, Chess and Birch, 1968), nine categories of temperament were established as listed below.

1. Approach/withdrawal is the child´s typical response (positive or negative) to new stimuli such as food, objects, or people.
2. Adaptability represents the ease with which the child adjusts to new or changed situations, regardless of the initial reaction.
3. Quality of mood reflects whether the child is predominantly positive or negative in behavior.
4. Intensity of reaction refers to the strength of the child´s responses to stimuli. For example, when happy, some children smile while others laugh uproariously.
5. Rhythmicity or regularity reflects how predictable the child is in areas of biological functioning such as sleep, feeding and elimination patterns.
6. Activity level is the amount of spontaneous movement in the child´s behavior and the daily proportion of active to inactive periods.
7. Threshold of responsiveness indicates how intense stimuli must be to evoke a response

from the child. For example, does the child react to very soft sounds or to small changes in temperature?

8. Distractibility is the ease with which the child can be diverted from ongoing behavior.

9. Attention span and persistence are two related categories, with attention span being the length of time the child will pursue an activity and persistence being how much the child will pursue an activity in the face of obstacles.

Subjectively, the NYLS group noted that the first five of these categories tended to cluster together, and from this qualitative analysis they formed three temperament syndromes. The clustering of the dimensions of approach responses to new stimuli, adaptability to change, positive mood, mild or moderate intensity of mood, and rhythmicity of biological functioning defines the "easy" child. The "difficult" child has the reverse pattern, and the "slow-to-warm-up" child is similar to the "difficult" child except that adaptations are slow, responses are mildly rather than intensely negative, and the child has more regularity in biological functioning.

Two studies (Thomas et al., 1963; Birch, Thomas, Chess & Hertzig, 1962) have reported that parent-observer reliability of the NYLS interview is significant at either $p < .05$ or $p <$

.01. The actual correlations were not reported in these publications, but Bates (1980) estimated them to be between .33 and .46. The actual correlation for inter-observer reliability is also not reported, but Thomas et al. (1968) indicated that it was significant at $p < .05$.

One of the reasons that the NYLS conceptualization of temperament still predominates is that a pediatrician, William Carey, developed a practical and convenient measure to assess the nine temperament categories (Carey, 1970; 1972a and b, 1973). Carey (1973) also extended the syndrome classification of the NYLS group by categorizing children as difficult, slow-to-warm-up, intermediate-high, intermediate-low, and easy on the basis of standardized scores. Carey´s measure, the Infant Temperament Questionnaire (ITQ), was developed for infants from 4-8 months and has since been revised to improve psychometric qualities (Carey & McDevitt, 1978a). Age appropriate versions of the ITQ have been developed for 1-3 year olds (the Toddler Temperament Scale, TTS; Fullard, McDevitt & Carey, 1978), for 3-7 year olds (the Behavioral Style Questionnaire, BSQ; McDevitt & Carey, 1978), and for 8-12 year olds (the Middle Childhood Temperament Questionnaire, MCTQ; Hegvik, McDevitt & Carey, 1981). Thus, convenient and age-appropriate measures of the nine NYLS categories exist for children from 4 months to 12 years.

As is apparent here and in the following paragraphs, the reader of this literature can get lost in an alphabet soup of abbreviations. There are two particularly confusing sets of abbreviations. The set composed of Carey´s ITQ, Rothbart´s IBQ, and Bates´ ICQ will be accompanied by their author´s name to aid discrimination. The second set, comprised of the NYLS´s Teacher Temperament Questionnaire (TTQ) and Carey´s Toddler Temperament Scale (TTS) have similar abbreviations. They will be referred to as the teacher TTQ and the toddler TTS, even though this is redundant.

Because the ITQ has been revised specifically to improve its psychometric qualities, only the reliability of the Revised Infant Temperament Questionnaire (RITQ) will be presented here. Carey and McDevitt (1978a) report that inter-item correlations ranged from .49 to .71 for the nine scales (median r = .57, and total score r = .83). Test-retest reliability over an average of 25 days ranged from .66 to .81 (median r = .75, and overall r = .86). Carey (1981) reported inter-parent correlations ranging from .32 to .72, with a median of .54 on the nine categories. Mother-father agreement was highest when families had just returned from vacation, and lowest when the father´s job limited his contact with the child.

The Swedish questionnaire for 6 month olds based on the NYLS system (Persson-Blennow & McNeil, 1979) is quite similar

to Carey´s ITQ. Internal consistency was based on significant discrimination between the highest and lowest quartiles of the scale. Throughout this section, internal coefficients are alpha coefficients. Retest reliability across 2 to 3 weeks ranged from .53 to .86 for 8 of the nine dimensions.

Fullard et al. (1978) reported that internal consistency of the toddler TTS ranged from .53 to .86 for the nine scales with overall internal consistency at .85 (median = .70). Test-retest reliability ranged from .69 to .89, with a median of .81 and a total r of .88. Garcia-Coll, Kagan, and Reznick (1984) had parents of 117 toddlers, aged 21-22 months, complete the approach-withdrawal scale of the toddler TTS. Mother´s and father´s ratings were reliably correlated ( r = .66).

The 3- to 7-year-old BSQ´s internal consistency ranged from .47 to .80, with a median value of .70 (McDevitt & Carey, 1978). The same authors reported that test-retest reliability of the BSQ over a one-month period ranged from .67 to .94, median r = .81, and total r = .89.

Hegvik et al. (1981) reported that internal consistency of the MCTQ (for 8-12 year olds) ranged from .71 to .87. Retest reliability over a 75-day interval ranged from .80 to .93 (Hegvik et al., 1981).

Other measures assessing the nine NYLS dimensions include

the Teacher Temperament Questionnaire (TTQ; Thomas & Chess, 1977), and a Q-sort measure (Pederson, Zaslow, Cain, Anderson, and Thomas, 1980). As well, the NYLS interview has been translated into Norwegian by Torgerson and Kringlin (1978).

Bates and his colleagues (Bates, Freeland & Lounsbury, 1979) have chosen to focus on infant difficultness rather than separate dimensions of temperament with their measure, the Infant Characteristics Questionnaire (ICQ). The ICQ is used for 5-12 month olds, and Lee (1981) has reported a revision for toddlers, the Child Characteristics Questionnaire (CCQ). Bates et al. (1979) report internal consistencies of .79 for the primary factor of fussy-difficult, .75 for the second factor of unadaptability, .50 for unpredictability, and .39 for the dull factor. Test-retest reliability over a one-month interval was reported for 112 subjects and ranged between .47 and .70 for the four factors. Inter-parental ratings correlated in the .38 to .61 range.

Rothbart (1981) has attempted an alternative delineation of temperament dimensions. She selected 11 dimensions to study from those of the NYLS group and other researchers (Diamond, 1957; Escalona, 1968, 1973; Escalona & Heider, 1959; Korner, 1971; Shirley, 1931, 1933a,b) and retained six which fulfilled conceptual and item analysis criteria,

forming the Infant Behavior Questionnaire (IBQ). The six dimensions retained were activity level, smiling and laughter, fear, distress to limitations, soothability and duration of orienting. Internal consistency was controlled by retaining only items which correlated at least .20 with dimension scores (Rothbart, 1981). Inter-rater reliability (mothers with fathers or with babysitters) ranged from .45 to .69. Goldsmith and East (1981) administered Rothbart's IBQ to parents of 8 1/2 month old infants (N = 174). Parents' ratings correlated as low as .14 for the smiling-laughter and .16 for the persistence subscales, and ranged between .20 an .43 for the other four subscales. Mothers and fathers completed ratings one week apart, and this time lag may have lowered inter-rater reliabilities.

In contrast to the above research, Buss and Plomin (1975) approached temperament from a more theoretical stance, specifying five criteria for determining which dimensions of behavior to retain as temperament. The most important of these, that the dimension have a genetic component, is central to the other four. The others are that the dimension show stability during childhood, retention into maturity, have adaptive value, and be present in our animal forebearers. Based on these criteria, they chose four dimensions of temperament (emotionality, activity, sociability and impulsivity) and developed the EASI

questionnaire to measure them. More recently Buss and Plomin (1984) deleted impulsivity, changing the EASI to the EAS. In their earlier book Buss & Plomin (1975) had concluded that support for impulsivity as a temperament was equivocal, but had included it in their measure. When research since 1975 still indicated only equivocal support for this dimension, impulsivity was deleted.

Only test-retest reliability information is available for the EAS; inter-rater reliability and internal consistency data comes from research using the EASI. Buss and Plomin (1984) reported that retest reliabilities over one week for 31 children (average age was 3.6 years) were .72 for emotionality, .80 for activity, and .58 for sociability/shyness. All internal consistencies of the EASI were reported by Buss and Plomin (1975) to be above .30, with most above .50. Inter-rater reliabilities between mothers and teachers ranged from .34 to .51 (Corsini & Doyle, 1981).

To examine the relationship between the EASI and the NYLS system, Rowe and Plomin (1977) created six questions for each of the NYLS dimensions, and administered these and the EASI to the same sample. Factor analysis of the data yielded yet another measure, the Colorado Childhood Temperament Inventory (CCTI), which is applicable for children 1-6 years of age. This measure assesses six factors of temperament: sociability, emotionality, activity, attention

span-persistence, reaction-to-food and soothability. Internal consistency for the six scales of the CCTI ranged from .37 to .88. Retest reliability assessed at a one-week interval ranged from .72 to .80 for four of the scales with lower correlations for sociability and soothability (.58 and .43, respectively). There is no inter-rater reliability information available for this measure.

Summary and Comment. The above measures all seem to have adequate internal consistency and test-retest reliability with reliabilities increasing with the age of the child. Data on inter-rater reliability is not available for all measures, but what is available is generally in the .40 to .60 range. While low, these seem adequate given the problems obtaining independent raters with comparable experience with the child. Statistical controls for varying degrees of contact would be advisable given Carey's (1981) observation that reliability of inter-parental ratings varied with the amount of contact the father had with the child.

A promising approach in the search for higher temperament reliabilities is suggested by Epstein (1979, 1980). He reported that it was possible to demonstrate consistency of individual behavior if one obtained an adequate sample of that behavior. Epstein compared the usual approach of administering one test in the lab and then looking for relationships at other times or in other

situations to expecting a test containing only one item to be reliable. He demonstrated the utility of repeated assessments with adult subjects. When repeated measures have been taken of children, results are equally promising. Morgan and Camp (1981) studied 56-month-old infants' vocalizations in response to their mother's attempts to get them to vocalize. If data for one day was included, the reliability for infants' vocalizations was low (r = .22 to .23), while mothers' reliability with an equal amount of data ranged from .64 to .76. If two days' data were used, maternal reliability improved to .83, but five days' data were required to increase infants' reliability to .62.

Eaton (1983) has also demonstrated the importance of multiple assessments with children aged 42 to 62 months. Eaton assessed activity level using mechanical recording devices called actometers. Different actometers were used on different days so that increased reliability over actometers also reflected increased reliability over days. Reliability increased from .33 for one day's data to .50 for two, .60 for three, .67 for four, and .76 for five days' data (values estimated from Eaton's Figure 1). It is impossible to obtain stable or valid ratings if our measures are not reliable, and it may be that the research reviewed suffers from low reliability of measures because children's temperament has typically been assesseed only once.

Stability of Temperament

In this section I will first review studies using the NYLS formulation, then studies using other measures. This review will be followed by a consideration of factors which affect temperament stability.

There is only one study of newborns which utilized a standardized temperament measure. Katcher (1977) observed newborns for two hours on the second and fourth day of life as well as interviewing the nurses on these days, and rated their temperament on these bases. The relationships between these ratings and those from the NYLS interview administered at three months were not impressive, possibly due to the low reliability which was reported both between observers and between sessions in the newborn period.

Pain (1979) administered the ITQ at 3, 6, and 9 months and reported significant correlations for five of the nine NYLS dimensions (approach, mood, intensity, distractibility, and persistence) from 3-9 months (r's ranged from .36 to .54, median r = .50). Correlations across the three-month periods from 3-6 and 6-9 months were somewhat higher, and of these, 13 out of 18 were significant (ranging from .37 to .71, median of .63). Wilhoit (1976) also administered the ITQ at 3, 6, and 9 months, and the NYLS interview at 3 and 9 months. She concluded that children did show stability from 3 to 9 months (figures were not reported) and that the Carey scale

was not as sensitive to this stability as was the NYLS interview.

Two studies assessed stability from infancy to toddlers, using the RITQ and the toddler TTS. Martin and Wachs (1981) administered the RITQ at 6 months and the toddler TTS at 12 months and found significant correlations across that period for four of the nine dimensions: adaptability ($r = .57$), mood ($r = .46$), rhythmicity ($r = .38$), and activity ($r = .49$). McDevitt and Carey (1981) tested 115 infants at 4-8 months of age with the RITQ, and then administered the toddler TTS an average of 17 months later. Results indicated significant stability in all nine dimensions.

Carey and McDevitt (1978b) assessed the stability of temperament ratings from 4-8 months (with the RITQ) to 3-7 years (with the BSQ). These data are reported in terms of the temperament clusters (difficult, slow-to-warm-up, intermediate-high, intermediate-low, and easy) rather than dimensions. Stability was most pronounced from infancy to 3-5 years, with 39% remaining in the same category. Twenty percent remained in the same category from infancy to the 5-7 year range, and 30% were in the same category overall. Approximately 40% of the two difficult groups (difficult and intermediate-high) retained that status from 4-8 months to 3-7 years, and a similar proportion (39%) of easy children remained easy. Thus, a larger proportion of children at the

extremes shifted towards the mean than stayed at the extremes. Twenty-six percent of the difficult children retained that classification. Carey and McDevitt (1978b) reported that stability of difficult or intermediate-high status was predicted by very negative mood and by high activity level scores. Children who changed from other categories to difficult or intermediate-high status were more likely to have been very withdrawing from new situations.

It is likely that results were influenced by regression to the mean effects. This refers to the fact that whenever one chooses extreme groups to study, it is likely that at retesting many of the subjects in the group will no longer test as extreme. Thus, care must be taken against false inferences that this change is due to treatment or experimental effects. Hegvik et al. (1981) have extended these results to the 8-12 year old range. Children were first assessed when they were 3-7 years old with the BSQ, and then retested at 8-12 years with the MCTQ. Correlations across the interval of 4.4 years were significant for all nine categories (correlations ranged from .21 to .59).

Billman and McDevitt (1980) administered the BSQ to mothers of 34-64 month old children. The same children were assessed by nursery school teachers five months later with the teacher TTQ. Stability was observed over the five-month period with significant correlations for all categories but

mood. Correlations ranged from .18 for activity to .47 for approach/withdrawal (distractibility r = .26, adaptability r = .38, intensity r = .40, and persistence r = .41). Threshold was not observed adequately in the nursery and so was not scored, and the teacher TTQ does not include a scale for rhythmicity. These stability figures are slightly lower than others observed, and this may have occurred because ratings were made by different individuals in different settings.

Stability data is only available for one of the other measures of temperament, Rothbart's IBQ. The IBQ was completed by mothers at 3, 6, 9, and 12 months, and inter-age correlations for activity ranged from .48 to .78, while those for smiling/laughter ranged from .57 to .81. Duration of orienting and soothability also demonstrated stability across the 3-12 month period though not all correlations were significant. Fear and distress to limitations were stable from 6-12 months but not from the three-month assessments.

In summary, the measures reviewed above have demonstrated adequate stability during childhood. The Carey scales seem stable from 6 months to 12 years, and Rothbart's scale demonstrates consistency within the first year. Some studies which assess behavior similar to temperament, but which did not use temperament measures, will be reviewed next. In some cases these provide valuable information about

what factors lead to stability or instability of individual differences.

It is possible to observe stability from newborn to later periods with other measures. Dempsey (1977) administered Brazelton's Neonatal Behavior Assessment Scale, (NBAS; Brazelton, 1973) on the first two days and again at one month, and found a correlation of .78 between the two measures of the infant's best responses (with agreement coded as within two points on the nine-point rating scale). The discrepancy score between the infants' best and characteristic responses demonstrated very high (.91) stability. Powell (1974) administered an early version of the Brazelton-Cambridge Newborn Scales (Brazelton & Freedman, 1971) to newborns and retested infants at 4 and 6 months with the IBR. He reported correlations of .67 and .64 between newborn responsivity and 4 month and 6 month IBR, respectively. Birns, Barten and Bridger (1969) reported significant rank order correlations across the first four months for ratings of irritability, tension, sensitivity, and soothability. Activity level ratings were stable across the period if the neonatal ratings were excluded (instability in the first few days may have been due to anesthesia effects on the neonate).

Halverson and Martin (1981) culled an activity-irritability factor of child behavior from the

Bethesda Longitudinal Study data. In this study approximately 200 families were followed from marriage until the first child was 3 years old. Children were assessed at birth, 3 months, and 3 years of age on the activity-irritability factor. Mothers were scored in a similar manner on their confidence/competence from the time of the marriage until the child´s third birthday and were divided into groups on the basis of whether their feelings of competency were stable or unstable over four time periods: newlywed, pregnancy, postpartum, and preschool. Halverson and Martin found that children of mothers with stable confidence/competence were unstable in their activity-irritability, with correlations near zero between time periods. Conversely, children of mothers with unstable confidence/competence were highly stable in their activity-irritability over time (figures unreported).

Kronstadt, Oberklaid, Ferb and Swartz (1979) had mothers complete a questionnaire assessing their infant´s sleep, feeding, and crying behavior at 5 weeks, and at 3 and 6 months. On the basis of these questionnaires infants were categorized as "difficult" or "average". From 5 weeks to 6 months, 86% were stable in this categorization. Ninety-one percent were stable from 5 weeks to 3 months, and 92% were stable from 3 to 6 months.

Matheny, Wilson, Dolan and Krantz (1981) report slightly

different stability information which was culled from the data of the Louisville Twin Study. In interviews mothers discussed twins' behavior and then indicated if the twins were similar or different. If different, mothers were asked to rank the twins. Sixteen behaviors were selected and the stability of intra-pair rankings of these behaviors was examined from 6 months to 6 years. Matheny et al. found significant rank stability for three behaviors across the entire age span: cuddling, activity, and smiling. As well, several other behaviors were stable across shorter age spans.

Krackow, Kopp and Vaughn (1981) assessed children's attention at 12, 18, 24, and 30 months. Attention correlated .81 for the shortest period from 12 to 18 months, .64 from 12 to 24 months, and .50 for the longest period from 12 to 30 months.

Feiring and Lewis (1980) report the least encouraging stability data, using behavioral observations of the temperament dimension of activity and vigor, and concluded that their results "revealed almost no individual stability" (p. 72) with the exception of motor activity from 13 to 25 months ($r = .33$). However, these investigators used questionable techniques to study temperament, and different measures at different ages. A measure of floor squares traversed during a 15-minute free-play period was used to assess motor activity. This measure has been criticized

(Cromwell, Baumeister & Hawkins, 1963) because it is possible for a child to traverse several squares by walking slowly around the room while another child may play vigorously with a toy and stay within one or two squares. Secondly, the observation period was only 15 minutes in length. Third, stability from 13 or 25 months to 44 months could not be assessed because measures of motor activity changed from squares traversed at 13 and 25 months to nursery school behavior at 44 months.

Chamberlin (1977) found similarly disappointing results when he attempted to predict across time and settings. Mothers were interviewed when their child was two years old and again when she/he was 4 or 5 years old. Teachers rated the 4-5 year olds twice on a behavior checklist at a six-month interval. Factor analyses resulted in the same three factors for all assessments (aggressive-resistant, inhibited, and high activity) and data were examined for continuity of these factors. No consistent evidence for stability was found between home behavior and school behavior, but there was stability for home behavior from 2 to 4 years and from 2 to 5 years, and for school behavior from 4 to 5 years. In the home, seven of eight inter-age correlations (ranging from .31 to .73) of the aggressive-resistant factors were significant. Stability for the withdrawing factor was evident for females (three out of

four r´s were significant, ranging from .37 to .57) but not males. Correlations within the school setting on these factors indicated more stability for males (all correlations were significant, range .34 - .61) than females (four of nine r´s significant, range .25 - .46).

Summary and Comments. The previous research has demonstrated moderate stability in temperament, as well as change over time. As noted earlier, scores on the Carey scales have remained stable from 6 months to 12 years of age, and Rothbart´s IBQ has demonstrated stability across the first year. The present section concerns variables which influence stability and the directions of these effects. For example, as noted earlier, Carey and McDevitt (1978b) noticed that stability of difficult or intermediate-high status was predicted by very negative mood and high activity level. Withdrawal from new situations predicted change from other categories to difficult or intermediate-high status.

Cameron (1977, 1978) reanalyzed data from the original NYLS study and examined the relationship between parental characteristics and changes in temperament over time. This study suggested that specific changes in temperament were associated with parental treatment. Three clusters of parental behavior (disapproval, intolerance, and rejection; conflict within and between parents; and inconsistent discipline) were related to decreases in children´s

adaptability. Other parental behaviors, such as strictness, were related to positive behavior at one age but negative behaviors at another.

Dunn (1979) reviewed three issues central to the continuity of individual differences. These are: the problem of assigning meaning to behavior, the relationship between individual differences in children and in their caretakers, and the effect of the nature of the sample on predictability. The first issue relates to the question of what we can infer from observed behavior. Dunn noted that such dissimilar behaviors as sucking differences in newborns have been related to differences in reaction patterns in the third and fourth years (Bell, Weller & Waldrop, 1971; Yang & Halverson, 1976; Dunn & Richards, 1977). Thus, stability of temperament ratings may be decreased because one particular behavior may reflect the dimension of interest at one age but not at another.

Impressive predictability from infancy measures can be obtained if one can discover which behaviors reflect the dimension of interest at different ages. Roe, McLure and Roe (1982) extended an earlier report (Roe, 1978) which indicated that higher vocal responsiveness to the mother versus a stranger in three-month-old infants predicted verbal-cognitive functioning at 3 to 5 years. Roe et al. (1982) retested the same children at 12 years of age and

indicated that 3-month DVR predicted 12 year old verbal performance on the WISC-R at the .85 level. Similarly, DVR at 3 months predicted Wide Range Achievement Test scores at 12 years (r=.86 for arithmetic and .65 for reading) and Peabody Picture Vocabulary Test scores (r=.61). These are impressive figures in an area which has difficulty finding stability over much shorter periods.

The second issue discussed by Dunn (1979) concerns the issue of isolating continuity in the child's individual differences from continuity of caretaker differences and continuity of the relationship between the two. Of course, this isolation is impossible given the transactional relationship between the individual and the environment with each affecting the other over time. Thus, when we do find discontinuity in temperament we do not know whether to attribute this to the individual, the caretaker, or their relationship. In the example presented by Dunn, newborn non-nutritive sucking was correlated with maternal affectionate interest during feeding, and both of these predicted the frequency of baby's vocal demands and mother's responsiveness to vocalizations at 14 months (Dunn, 1977). Dunn preferred to interpret these results as stability of the relationship rather than stability of either individual involved. Extending this notion to temperament stability, Dunn reported preliminary results of research by Tsitsikas

(1979) which indicated a difference in predictability of 6 and 9 month temperament ratings from newborn NBAS scores. In this case stability depended on whether the infant was or was not adopted in the intervening period. When the infants were adopted there was no relationship, but when they remained in the baby centre there was stability of temperament from newborn to 6 and 9 months (relationship size unspecified).

Finally, Dunn´s third point deals with sample characteristics which affect predictability. Predictability is often found for subjects at the extremes of behavioral ranges when it is not found for those closer to the mean. For example, in Kagan´s (1971) research, boys who were either very slow or very fast in tempo were impressively stable from 8 to 36 months. Escalona (1973) discusses this issue with reference to her earlier research (Escalona & Heider, 1959).

> We found, as others have, that by and large infant behavior characteristics are not maintained over time. Ratings of activity level, movement qualities, irritability, and the like tended to converge toward the mean, or else children shifted their position on the continuum. However, and this is potentially important, certain children did carry into the preschool years behavior attributes that had been noted when they were as young as 16 weeks. This minority of self-consistent children

> had one thing in common: all had shown significant developmental deviation or outright maladjustment. Whether their difficulties were due to repeated serious illness, surgery and hospitalization, or to the impact of grave conflict and unhappiness within the family made no difference. It is when development and adaptation proceed stressfully and not entirely successfully that certain early patterns are more likely to remain intact and continue to be present in unaltered form at later ages. (Escalona, 1973, pp. 146-7)

This comment supports the stability of infant maladjustment when the environment is stressful rather than supportive. One could also predict stability for infants at the positive end of the continuum if their environments are supportive.

Stability of extreme scores also relates to a consideration of the amount of variability in the measure. Researchers often select their samples to be as homogeneous as possible on control dimensions in order to decrease error variability. However, in a field where dimensions such as socio-economic status and education of parents (which are usually controlled) can have such large effects on individual differences, researchers may unwittingly be restricting the range of differences, and therefore decreasing the size of

correlation coefficients (McNemar, 1969). For example, McGehee, Eckerman, and Gross (1981) noted that researchers often form groups of subjects with or without perinatal complications, and then examine the relationship between group membership and some measure of later outcome. Thus, the range of one measure usually used is limited to the presence or absence of perinatal complications. Rather than rating their sample of premature, low birth weight infants on this variable, McGehee et al. looked for variability within the group on a measure of responsiveness to social stimuli. The amount of en-face gaze and the amount of time the infant maintained a quiet, alert state was assessed while the baby was socially stimulated. Correlations of these measures with the Bayley administered at six months were very high, .93 for gaze time and .98 for duration of quiet, alert state. In addition to changing their measure, it is likely that McGehee et al. capitalized on increased variability. It is also interesting that their measure was similar to that used with success by Roe (1978; Roe et al., 1982). Both seem to involve elements of attention and sociability. The techniques suggested in this section for assessing stability will have to be tried empirically before psychology as a field concludes that children s individual differences are not stable over time.

Validity of Maternal Temperament Ratings

An issue which must be addressed before presenting validity data is the distinction between inter-observer reliability and external validity. Reliability concerns the consistency of test results over time and observers. External validity refers to the generalizability of results. To assess external validity, researchers usually vary subject characteristics such as age or gender, but external validity can also be construed as generalization across experimental contexts such as examiner gender. Thus, agreement between parents, or between parents and observers, falls into both reliability and validity domains. Three recent reviews (Bates, 1980; Carey, 1981; and Rothbart and Derryberry, 1982) have all considered mother-father agreement of their child's temperament as external validity data. Because these concordances are not overly impressive, Bates questions the existence of temperament.

Although the concordance of parental ratings falls into both reliability and validity domains, this information will not be included in the following validity section because mother-father agreement is of questionable value due to confounds. Bates (1980) pointed out that parents inevitably confound each other's ratings even if they complete the scales independently because they have been communicating to each other about the child since her/his birth. Thus

parental concordances are likely to be spuriously raised due to shared perceptions of the child. On the other hand, Carey (1981) reported preliminary analyses of recent data in which the amount of contact the father had with the child affected results. The highest inter-parent correlations occurred in two cases where the family had just returned from their vacation while the lowest occurred in cases where the fathers had extremely little contact with the child due to their jobs. Whether parental agreement is spuriously increased because of inter-parental communication or decreased due to low paternal contact with the child has not been shown empirically, but until this issue is settled it seems premature to question the existence of any factor on the basis of this data. Agreement of independent raters who have had equal amounts of experience with the child would be preferable for both reliability and validity considerations. When this is not possible, researchers are advised to assess the extent of paternal contact with the child.

Relationships between scales. When researchers in this area develop a new temperament scale, they validate it by examining the relationship between their measure and the most predominant measure currently in use. Carey (1970, 1973) was the first researcher to translate the NYLS interview into questionnaire form. Consequently, he validated his measure by comparing its results to those of the NYLS interview. In

turn, all infant scales developed since that time have been validated by comparison with Carey's measure. Now that numerous measures exist, it would be interesting to know how they relate to other new measures in addition to Carey's. One benefit of such research could be the development of measures acceptable to most researchers in the area which, if widely adopted, would greatly increase the ability to relate studies to one another. Setting this issue aside, this section will discuss the relationships between Carey s (1970, 1973) ITQ and other measures. Following sections will examine the relationships between ratings and observations, and between ratings and objective measurements.

Thomas and Chess (1977) report that the ITQ identified proportions of difficult, easy and slow-to-warm-up infants that were not statistically different from those identified by the NYLS interview. Wilhoit (1976) also reported that ratings obtained by the ITQ and NYLS interview on 24 children at 3 and 9 months were similar.

Carey (1972b) standardized his measure on 200 subjects, and while there are differences in mean scores between the ITQ and the NYLS interview, both measures agree on the profile of the average 4 to 8 month old. Sameroff et al. (1982) administered the ITQ to mothers in Rochester, New York. Mean scores from this study are in almost perfect agreement with the means of Carey's standardization sample.

Further, consistent differences in mean scores between the ITQ and the NYLS interview such that ITQ scores are closer to the mode have been attributed to superiority of the interview method in circumventing maternal biases to present their babies positively. The existence of such a bias was supported by the fact that when mothers give a global rating of their child on a temperament dimension, about 25% minimize their infant´s difficulty.

Campbell (1979) examined the relationship between Carey´s ITQ and Bates et al.´s (1979) ICQ. He related a Difficulty Index, consisting of ITQ adaptability, mood and rhythmicity, to the fussy and unadaptability factors of Bates´ ICQ. The Difficulty Index correlated at low but significant levels with these ICQ factors (r=.22 for fussy-difficult and r=.32 for unadaptability). In addition, Bates et al. predicted six relationships between individual Carey ITQ scales and Bates ICQ factors on the basis of content overlap, and all were confirmed though correlations were low. For example, the ICQ fussy-difficult factor correlated -.25 with low intensity and .39 with negative mood.

Two studies have assessed the relationship between the Brazelton NBAS and the Carey ITQ (Sostek & Anders, 1977) or RITQ (Martin & Wachs, 1981). Sostek and Anders (1977) administered the NBAS at eight days and Carey´s ITQ at two weeks. The NBAS social-interactive scale correlated with the

Carey activity (-.25), intensity (.39), and distractibility (.53) scales. The motoric-interactive scale correlated with the intensity (.55), mood (-.41), and distractibility (.62) scales. State control on the NBAS related to adaptability (.33), mood (.28), and distractibility (.50). The total NBAS score related to intensity (.40) and distractibility (.65).

Martin and Wachs (1981) assessed infants at slightly later ages (NBAS at one month and the RITQ at six months). The NBAS social interactive scale correlated significantly with rhythmicity (.31), approach/withdrawal (.29), adaptability (.33) and mood (.33). The motoric-interactive score correlated with rhythmicity (.34), approach/withdrawal (.33), and adaptability (.31). State control related significantly to activity (.39), approach/withdrawal (.40), adaptability (.34), and intensity (.29). The differences between relationships found by the two studies could be due to many factors, including age differences. The fact that newborn behavior is related to ITQ scores at 2 weeks and at 6 months suggests that the ITQ and the NBAS assess similar underlying dimensions of infant behavior. It also suggests that the baby who is behaviorally organized at one month will be rated as easy at six months.

The only study of the relationships between different measures which did not include the Carey ITQ was completed by Rowe and Plomin (1977). These researchers did however relate

their measure, the EASI, to the NYLS system using items they created. Factor analysis of these items resulted in seven factors, and the relationships between these factor scores and EASI scores were examined. EASI emotionality correlated most with the factors of stubbornness (.45), reactivity (.34), and soothability (-.43). EASI impulsivity correlated significantly with attention span/persistence (-.56), stubbornness (.43), reactivity (.20), and soothability (-.26). The two sociability scales correlated .64.

Relationships between scales and observations. Six studies have been completed in which ratings of the child's temperament were correlated with independent observations of his/her behavior (Bates et al, 1979; Billman & McDevitt, 1980; Rothbart, 1986; Sarrett, 1976; Sameroff et al, 1982; Scholom & Koller, 1980). In the Bates et al (1979) study, observations of the infant's behavior in the home were submitted to factor analysis. The Difficulty Index for the Carey ITQ correlated zero with the first factor (observed fussing), .16 with the second factor (negative versus fun) and -.25 with the third factor (unsoothability). It was predicted that the fussiness and unsoothability factors would be correlated with ITQ mood and intensity and this was confirmed for fussiness ($r = .22$ for both mood and intensity), but not for unsoothability (correlations with mood and intensity were -.06 and .09, respectively).

Rothbart (1986) assessed 46 infant´s temperament longitudinally at 3, 6, and 9 months of age. Mothers completed the IBQ, and observations of infant and mother were completed at each age. Significant correlations were found between maternal ratings and observations of 5 temperamental dimensions at 3 months (activity level, smiling and laughter, distress to limitations, fear, and vocal activity). At 6 and 9 months, four of these five scales were significantly related to observations.

Sarett (1976) reported that difficult babies cried more than easy babies and received five times as much maternal behavior. Sameroff et al. (1982) generally found that predicted relationships between ITQ and observed behavior for two to three hours in the home and the Bayley IBR were supported, but that relationships were small, ranging from .14 to .20 in absolute size. Sameroff et al. used the original ITQ rather than the revised ITQ. Since internal consistencies for the ITQ were much lower than those for the RITQ, these low correlations might be due to increased error variability in the ITQ.

Scholom and Koller (1980) used canonical correlation analysis to examine the relationship between parental temperament ratings and the infant´s interactions with the mother and father. Infant behavior was related to temperament ratings, and the same pattern of relationships was found for

mothers and fathers. Infant attention to the environment was related to the temperament dimension of activity level (.38 for mothers and .39 for fathers) and intensity (.49 for mothers and .51 for fathers). Approach was related to attachment behavior (-.35 and -.30) and intensity related to purposive behavior (r = .43 and .28).

Billman and McDevitt (1980) had mothers of 34 to 64 month old children rate their children on the BSQ, and then observed the child in the nursery school for 20 two-minute periods over a three to five month period. After this period observers rated the child's temperament again, using the teacher TTQ. Relationships were observed between children's categorizations as easy or difficult and observed behavior. The more difficult the child, the more she/he jumped (.19), wrestled (.18), hit (.23), pushed (.19), and hit and beat (.27) other children. This connection between temperament and aggressive behavior was also observed for adolescents by Olweus (1980), although temperament in this study was rated from a partly retrospective interview in terms of activity, and calmness versus hotheadedness. Although temperament only correlated .14 with later aggression, an indirect relationship with mother's permissiveness was observed (r=.36), indicating that predictability is improved when both temperamental and environmental influences are assessed.

Eaton (1983) compared parental ratings of their

preschoolers on the CCTI with teacher rankings of activity level and with actometer scores. He found that the activity scale on the CCTI correlated .75 with the objective actometer scores. Actometer scores also correlated significantly with CCTI sociability (.54) and reaction-to-food (-.44) scores.

Summary. There is a definite need to clarify the concept and measurement of temperament. Researchers who do not agree about the specific definition of the concept create measures which represent different operational definitions of temperament. Many of these measures are not subjected to strict and extensive methodological analysis before they are published. As noted in the section on the topic, most have been validated by comparison with the ITQ, a measure which has been revised because it had low internal consistencies. Sameroff et al. (1982) pointed out that none of the relationships with the ITQ scales which had the lowest internal consistency were significant. Until measures are improved, and until researchers agree on common measures, it seems likely that research results will continue to show the same pattern - i.e. low to moderate correlation coefficients which occur in predicted directions often enough for some researchers to continue work in the area at the same time as other researchers are questioning the existence of the phenomenon. It is hoped that future research with more reliable measures can improve this situation.

Temperament Factor Structure

This section will address the issue of temperament dimensionality with information gleaned from factor analytic studies. This material will be evaluated with several points in mind. The first is that you only get out of a factor analysis what you put in. For example, Graham, Rutter and George (1973) asked mothers about their child's reaction to being dirty, and, not surprisingly, factor analysis of this data resulted in a fastidiousness dimension. The second consideration is that analyses of the same measure may differ, depending on whether individual items or dimension scores are submitted to the analysis. Principal components factor analysis is most frequently used, and a third consideration is that this technique yields orthogonal factors when we do not know if temperament dimensions are in fact orthogonal. Some research has used factoring techniques which do not produce orthogonal dimensions. For example, Cameron (1978) used cluster analysis, which groups items or individuals with theoretically-chosen pivot variables. A drawback of many studies using this approach is that the rationale for pivot variable choice is not specified.

The fourth consideration concerns the degree to which investigators concern themselves with clinical issues. It will become obvious that some researchers describe factors that define a "difficultness" dimension while others focus on

more discrete behaviors such as the quality of mood, activity level or attention. The difference between these approaches lies in their emphasis. The researchers who focus on "difficulty" of the child (Bates, 1979; Thomas & Chess, 1977) are more interested in the impact of the child on the parents. What behaviors of the child make her/him difficult to raise? Other researchers approach temperament from a more theoretical stance and focus on the child more than on her/his impact on caregivers (e.g., Buss & Plomin, 1975). This last point relates to the difficulty of isolating individual from contextual variability. It is generally accepted that it is not possible to isolate the two, but investigators differ in the extent to which they focus on the child or on his/her relationship with the environment.

Factor Analyses of NYLS Dimensions. Thomas and Chess (1977) indicated that factor analysis of the nine NYLS dimensions had found three factors, but only the first of these was described in any detail. This factor, which was stable from ages one to five years, was composed of the approach/withdrawal, adaptability, mood, and intensity subscales. The concordance between the subscales included in this factor and the subscales chosen to identify the easy-difficult diagnostic clusters by the NYLS researchers (only rhythmicity is missing) adds to the validity of this classification. An earlier report (Thomas, Chess, & Birch,

1968) indicated that the other two factors were interpretable as rhythmicity/threshold and activity level/distractability.

Difficulty-easy factors have emerged from several other studies. Maurer, Cadoret and Cain (1980) and McDevitt and Carey (1978) both used person cluster analysis and replicated the NYLS easy, slow-to-warm-up, and difficult clusters. Cameron's (1978) results were similar even though he used a slightly different analysis technique. Four clusters (approach, adaptability, mood, and persistence) were identified which, when combined, differed only from the NYLS factor A by including persistence rather than intensity.

It is interesting to note that while the NYLS group found a difficulty factor, they did not classify subjects along this dimension of variability. Instead they chose to identify the two poles as separate easy and difficult syndromes. This has lead some researchers to think of these syndromes as separate phenomena rather than as the two end points of one dimension, thereby increasing the difficulty of interpreting research results. Differences in analysis approaches contribute to this error in interpreting results. If scores on certain variables intercorrelate to form a factor, this occurs because individuals are rated similarly on these variables. Thus, factor analysis results in difficult-easy dimensions while cluster analysis results in separate difficult and easy groups. Misinterpretation of the

relationship between cluster analysis and factor analysis creates unnecessary confusion in this area of research.

Three studies have reported factor analyses of infancy data, all using the unrevised version of Carey´s ITQ. Pain (1979) included the nine NYLS dimension scores, Bayley scales, and observations of mother-child interaction in her analysis. The first two factors included the observational data. The third factor was interpreted as object versus mother involvement and included activity level/persistence. The fourth factor, called easiness, was very similar to the NYLS factor A, including approach, adaptability, and mood. Other factors were child vocalness, fussiness, reactivity, rhythmicity, and Bayley scores.

Scholom, Zucker and Stollak (1979) assessed 6-month-old temperament retrospectively by having parents of 3 to 4 year old children complete the ITQ jointly. The nine dimension scores were submitted for principal components factor analysis, resulting in three factors. The first factor, called mood, includes approach, adaptability, mood and threshold. This factor is also very similar to the first NYLS factor. The second factor was defined by regularity and persistence, and was called consistency. Energy, the third factor, was composed of activity level, intensity of reaction and distractibility.

The third of these studies performed a factor analysis on

individual items rather than on the nine subscale scores. Rapoport, Pandoni, Renfield, Lake and Ziegler (1977) assessed temperament at five months and found five factors. The first included aspects of contented, adaptable behavior; the second was an activity/vigor factor; the third concerned negative responses to people and crying; the fourth factor was called mildness and the fifth, regularity.

Two studies are difficult to compare with others in this section, either because different items were used or because setting and raters were different. Rowe and Plomin (1977) performed a principal components factor analysis of 54 items they created to assess the nine NYLS dimensions while developing the CCTI. They found seven factors: reaction-to-food, attention span-persistence, sociability, stubbornness, sleep rhythmicity, reactivity, and soothability. Sobesky, List, Holden, and Braucht (1981) administered the teacher TTQ to teachers of 400 children aged 3 to 6 years. Principal components analysis resulted in five factors which were labelled energetic (activity), responsive (threshold), focused, outgoing and placid.

Factor analyses of other measures. Bates et al. (1979) completed a factor analysis of the ICQ and reported four factors, fussy-difficult, unadaptable, dull (containing activity level and social responsiveness items), and unpredictability. Lee (1981) has reported that the first

factor of the CCQ, a version of the ICQ for older children, is also fussy-difficult ( other factors were not mentioned).

McInery and Chamberlin (1978) administered a 70-item checklist to mothers of two-year-olds. Three factors resulted from principal components analysis of this data: aggressive-resistant (temper, talks back, stubborn, disobedient); dependent-inhibited (shy, clings, cries easily, fearful in new situations); and friendly-outgoing (likes to be held, cheerful, friendly, talkative).

Bayley´s (1969) Infant Behavior Record (IBR) is a rating by the examiner of the child´s behavior observed during assessment. Several of these behavior ratings are of temperament dimensions such as emotional tone, attention span, reactivity and activity, so that this measure is often used as an indication of temperament. Goldsmith and Gottesman (1981) have recently analyzed ratings very similar to the IBR from the Collaborative Perinatal Project. Principal components analyses were carried out on these ratings from three ages: 8 months, 4 years, and 7 years. At eight months two factors, vigorous persistent activity and person interest, were found. The analysis of the 4-year-old data yielded three factors: task persistence, spontaneous activity and irritability. Four factors were found at seven years: active adjustment, fearfulness, task persistence and cooperation. It must be noted that these ratings took place

during cognitive and motoric assessment in the laboratory so that children´s behavior may have been affected by these situational constraints. As noted at the beginning of this section, what comes out of factor analysis depends on what goes into it. Specifically, the result of a task persistence factor at 4 and 7 years, while undoubtedly related to an attention dimension may also reflect the fact that ratings occurred in a testing situation.

Another group of researchers (Garside, Birch, Scott, Chambers, Kolvin, Tweddler & Barber, 1975) developed an interview to assess 15 categories of behavior. They concluded that children could be described in reference to six main dimensions: withdrawal and poor adaptability, assertiveness, activity and intensity, mood, distractibility and attention span, and irregularity. However, this interpretation is another example of the confusion between factor analysis and cluster analysis. The first factor was a bipolar withdrawal versus assertiveness dimension, which was incorrectly interpreted as two separate dimensions.

Matheny and Dolan (1980) developed 23 rating scales for several temperamental characteristics and other aspects of development, and asked mothers of 7 to 10 year olds to complete it. Principal components analysis of these data resulted in six factors: compliant morality, applied cognitive, sociability, emotionality,

assertiveness/impatientness and activity/distractibility.

The more theoretical approach to temperament of Buss and Plomin (1975) resulted in a measure, the EASI, designed to assess the four dimensions of behavior which had satisfied their temperament criteria. Rowe and Plomin (1977) completed a factor analysis of the EASI, and found the expected four factors (emotionality, activity, sociability and impulsivity). The same paper describes the formation of the CCTI from a factor analysis of the nine NYLS dimensions with the four EASI factors. The CCTI includes the five highest loading items from each of the six factors resulting from this analysis. These factors are: emotionality, activity, sociability, attention span-persistence, reaction-to-food, and soothability. The last two of these factors have been questioned. Rowe and Plomin (1977) concluded that soothability was not independent from emotionality ($r = -.42$); and suggested it might not be a separate dimension. Plomin & Rowe (1977) found no support for the heritability of the reaction-to-food dimension.

A combination of quantitative and qualitative approaches to determining the factor structure of a set of variables was used by Matheny, Wilson, Dolan and Krantz (1981). They selected 16 behavioral variables which represented temperament differences between twins across the first six years of life. For the six month old data, contingency

correlations were computed between all pairs of these variables, and the resultant matrices of correlations were examined for clusters. Two were identified on a rational basis. These were called temperamentality (temper frequency, crying, irritability, activity, piercing cry, resisting sleep and low attention span) and sociability (smiling, accepting people, and cuddling). Data for later periods were organized around these two clusters, with a few behaviors changing relationships with others or moving from one cluster to the other. For example, at the early ages cuddling clustered with sociability, but after 24 months cuddling gradually became associated with the temperamental cluster because it occurred when parents soothed upsets.

Summary. The striking aspect of this research is the frequency with which a factor very similar to the NYLS factor A (i.e. difficultness/easiness) results from analyses. This factor included the approach, adaptability, mood, and intensity subscales. When researchers submit the nine subscale scores of an NYLS measure to factor analysis, a factor including the approach, adaptability and mood scales generally emerges. The intensity scale is sometimes replaced by rhythmicity, but otherwise results are very consistent.

A problem with this area is that researchers do not agree on the number or type of subcategories to be included as temperament. For example, the NYLS researchers identified

nine subscales, and these have since been accepted and perpetuated by many researchers. Other researches have developed measures with varying numbers and types of subscales. It seems important for researchers in this area to agree on a definition regarding which categores of behavior are and are not included as temperament. Once such an agreement has been reached, researchers can proceed to the question of how to divide behaviors into categories or dimensions of temperament. Decisions will include where divisions should be made and how broad or narrow these subcategories of behavior should be. For example, the NYLS researchers identified nine dimensions while Buss and Plomin (1984) identified only three.

In addition to the above observation that a difficultness-easyness factor generally results when the nine NYLS subscales are submitted to factor analysis, what can this review tell us about the subcategories of temperament? Due to the wide range in type of items/observations used, and to the lack of agreement on definitional issues, any conclusions drawn from these studies must be tentative. Other factors which seems to emerge with some regularity include some representation of attention, and reactions to people or sociability. These factors are not always labeled as such. For example, other terms for factors involving attention include activity level/persistence, task

persistence, focused, and activity/distractibility. Since there is clearly a relationship between attention and activity level, researchers will have to decide whether or not these categories belong together, and, if they do, which is predominant. Factors reflecting sociability have a similar variety of labels.

Heritability

The results of twin studies have been questioned on the grounds that identical twins are treated more similarly than are same-sex fraternal twins, and therefore that treatment differences and not genetic differences account for greater concordances between identical versus same-sex fraternal twins. Because the twin method is so important to developmental research this issue, called the equal environments assumption, has been addressed directly. Loehlin and Nichols (1976) assessed differences in treatment of identical and fraternal twins and were able to identify a few differences in treatment, such as identical twins being dressed alike more often. However, a second question addressed in this study was whether the observed differential treatment affected behavior. The answer to this question was clearly negative. There were no significant correlations between environmental treatment variables and behavior. This finding has been replicated by Matheny, Wilson and Dolan (1976), and Plomin, Willerman and Loehlin (1976). Scarr

(1968) has also found support for the equal environments assumption. She found twins mistaken for dizygotic were as similar as other monozygotic twins, while dizygotic twins mis-identified as monozygotic are no more similar than other dizygotic twins. While the equal environments assumption seems valid, there are other variables which seem to influence twin concordances.

What is expected on a theoretical basis is that, if a behavior has an inherited component, monozygotic twins should be more concordant on this behavior than dizygotic twins. However, the differences observed should not be too large. Dizygotic twins still share approximately half their genes, the same womb, and the same environment. When the difference between monozygotic and dizygotic concordances is too large, or when the dizygotic concordances are too low, one suspects that raters are emphasizing the similarities between the monozygotic twins (assimilation bias) and/or the differences between the dizygotic twins (contrast bias). Thus, while it seems reasonable to conclude that twin research can give an indication about inherited components of behavior, it would be foolish to place too much faith in these figures in terms of absolute amounts of variability accounted for by either genetics or environment without controlling for rater biases. The main question to be addressed to this section of the literature is whether temperament dimensions have an

inherited component, and not the absolute size of that component. There are primarily four groups of researchers investigating the heritability of temperament: the Colorado group (at the Institute for Behavioral Genetics),those involved with the Louisville Twin Study (Matheny, Dolan, Wilson and colleagues), those using the NYLS conceptualization, and others who use family more than twin techniques.

The Colorado group. Buss and Plomin (1975) asked mothers of twins (139 pairs from 1 to 9 years old) to rate each child on the EASI. Mothers rated each twin of 139 pairs (average age = 55 mo., ranging from 1 to 9 years). Correlations on the four scales of the EASI were clearly higher for identical than for fraternal twins. All differences were significant except for impulsivity in females; identical twin correlations ranged from .50 to .84, while fraternal twin correlations ranged from .00 to .20. Plomin (1976) reported data from 137 families with 2 to 6 year old twins. Using the mean of parents' ratings, the average identical twin correlation was .55, while the average correlation for the same-sex fraternal twins was -.07. The results were similar when fathers rated one twin only and mothers rated the other. Data from both of these studies were probably affected by some rater bias because fraternal twin concordances were so low.

Rowe and Plomin (1977) administered the CCTI to 91 twin pairs aged 1 to 7 years and found that correlations were significantly higher for identical than for same-sex fraternal twins (with the exception of the reaction-to-food factor). With this exception, identical twin correlations ranged from .56 to .76, while fraternal twin correlations ranged from -.27 to +.27. Large differences between identical and fraternal correlations for sociability, activity and soothability suggested that assimilation and/or contrast biases were operative.

An interesting analysis was conducted by Rowe and Plomin (1977) to assess genetic contribution to the covariation between emotionality and soothability. A correlation of -.42 between these subscales indicated that the more emotional child was also less soothable. To determine the covariance, the cross-correlation between emotionality in twin A and soothability in twin B is obtained (and vice versa, soothability in twin A and emotionality in twin B), and compared for identical and fraternal twins. The cross-correlation was -.33 for identical twins and .11 for fraternal twins. Rowe and Plomin (1977) concluded that the same gene or gene system that leads to emotionality also leads to soothability. This may also indicate that these subscales can be expected to have similar relationships to physiological processes.

The Colorado group have viewed research using rating scales such as the EASI and CCTI as preliminary to more controlled techniques. Recent work (Plomin & Rowe, 1978, 1979) has utilized behavioral observations with appropriate controls for observer biases. Plomin and Rowe were interested in infant social behavior, which they observed in the home with a sequence of mother and stranger interactions similar to Ainsworth´s strange situation (Ainsworth, Blehar, Waters & Wall, 1978). Their subjects were 46 twin pairs whose age ranged from 13 to 37 months (mean = 22.2 months). Eight behaviors of the child indicated a genetic influence, and seven of these were directed at the stranger. Also, differences between the infants´ response to the stranger and the mother were computed, and intrapair correlations for these scores were compared between identical and fraternal twins. Only difference scores in the warm-up situation, when the stranger´s strangeness was most salient, indicated heritability. Plomin and Rowe discuss these results in terms of Buss and Plomin´s (1975) quantity versus quality components of social behavior. Buss and Plomin speculated that it is the quantity component of sociability (responsiveness to strangers) which is inherited more than the quality component (responsiveness to the mother). Research by Horn, Plomin and Rosenman (1976) supports the notion that the inherited component of sociability reflected

in this research is responsiveness to novel stimuli. These investigators carried out a factor analysis of items from the California Psychological Inventory and found a factor involving only interactions with strangers which had high heritability.

The Louisville Twin Study. Matheny and Dolan (1980) presented data on 105 same-sex twin pairs who were 7 to 10 years of age. As indicated earlier, factor analysis of maternal ratings on 23 items yielded six factors (compliant morality, applied cognitive, sociability, emotionality, tough-minded, and activity-distractibility). Identical twins were more similar than fraternal twins for five of these factors (all but the tough-minded factor), and in their pattern of scores across factors. In these data fraternal twin correlations were not as low as in the Colorado research reported earlier, but maternal rating biases were still possible.

Matheny et al. (1981) also demonstrated a genetic influence for emotionality and sociability. Data in this case were maternal comparisons of 305 same-sex twin pairs. Mothers indicated at several different ages whether twins were the same or different with respect to 16 behavioral variables (e.g., attention span, cuddling, crying). If a difference was indicated, the twins were ranked. As expected, identical twins were more often than fraternal

twins reported to be the same, particularly in the first 36 months. These differences were not significant at all ages or for all variables except for attention span and responses-to-heat. As reported earlier, the 16 behaviors factored into a temperamental and a sociability cluster, and genetic influences were more apparent for the former than for the latter.

Research using the NYLS approach. Two twin studies have been conducted using the NYLS approach to temperament. Rutter, Korn and Birch (1963) reported data from the NYLS sample of 3 monozygotic, 5 dizygotic, and 26 sibling pairs. The small sample size greatly restricts confidence in these results, but some indication of genetic influence was found for all dimensions except regularity. The strongest evidence occurred for the activity, approach/withdrawal, and adaptability dimensions.

More recently Torgersen and Kringlin (1978) studied 53 same-sex twin pairs, using the NYLS interview, when the infants were two and nine months of age. At two months, monozygotic twins were more similar than dizygotic twins only for the dimensions of regularity, threshold and intensity. Differences between monozygotic versus dizygotic correlations were significant for all dimensions at nine months. Intrapair differences increased for dizygotic twins (changes were significant for activity level, approach/withdrawal and

threshold) but remained constant or decreased for monozygotic twins from two to nine months. The difference in results between two and nine months was attributed to two possible factors: pre- or perinatal effects may still have been operative at two months, causing great behavioral variability and unreliability or measures which masked temperamental differences, and/or the broader range of behavior at nine months may have made temperament differences more apparent.

Other researchers. The first two studies described in this section used twins as their subjects while the last three studies used parent-child correlations. As noted earlier, Goldsmith and Gottesman's (1981) data consisted of psychologist ratings, similar to Bayley's IBR, during batteries of cognitive and motor tests. These ratings were factor analyzed separately for each age. At eight months, the factor of vigorous activity versus psychomotor passivity demonstrated genetic influence while the responsiveness to people factor did not. At four years, the task persistence and irritability/negative mood factors both demonstrated heritability while the spontaneity/activity factor did not. At seven years active adjustment and fearful/inhibited factors seemed genetically influenced but not the task persistence or agreeable/cooperative dimensions.

Goldsmith and East (1981) presented preliminary analyses of parental temperament reports of 8 1/2 month old twins

using Rothbart´s IBQ. Genetic influences were most likely for fearfulness (identical r = .47, fraternal r = .09) and distress to limitations (identical r = .57, fraternal r = .13).

Stevenson and Lamb (1979) assessed sociability of 12 month old infants both at home (responsiveness to a stranger) and in the laboratory (with IBR ratings). Maternal sociability was assessed with observations and the HOME inventory (Caldwell & Bradley, 1983), the first scale of which assesses the mother´s responsivity. The child´s initial sociability to the stranger was significantly related to maternal emotional and verbal sociability (r´s = .34 and .37, respectively).

Malatesta and Haviland (1981) corroborate and extend these results in their study of three and six month old infants and mothers´ emotional behavior. Mothers rated their infant´s temperament using Rothbart´s IBQ, and their own temperament with the emotionality scale of the EASI, which has subscales for fear, anger and general emotionality. Maternal anger and fear were significantly correlated with infant anger (r = .35 and.30, respectively). Also, maternal emotionality was significantly related to infant expressions of joy and interest.

Sameroff, Seifer and Elias (1982) assessed maternal anxiety and psychopathology and infant temperament at four

months using Carey's ITQ. Infants were also observed in their home and in the laboratory. Maternal anxiety was related to infant's approach, adaptability, mood and rhythmicity. These are four of the five subscales used to determine difficult-easy syndrome classification. In one sample, anxiety was also related to threshold.

Summary and Comments. The studies reviewed in this section indicate quite clearly that there is genetic influence on temperament, most notably on the dimensions of emotionality, activity and sociability. For sociability, it seems that how the child responds to people in general or to strangers (the quantity component) is much more clearly inherited than how the child responds to the parent (the quality component).

The three studies which demonstrated a relationship between maternal and child temperament (Malatesta & Haviland, 1981; Sameroff et al, 1982; Stevenson & Lamb, 1979) do not answer questions concerning the direction or cause of these relationships. Do parents affect the child's temperament, do children affect parents' temperament or are both true? Some have suggested (eg. Bates, 1980; Sameroff et al, 1982) that the child's temperament scores primarily reflect an anxious mother's perception of her child. It is also possible that the relationship between maternal anxiety and child difficultness exists because the characteristics of difficult

babies (intense negative reactions to most stimuli which do not habituate easily) make mothers feel inept and anxious, regardless of the mother's qualities before the child's birth. The data supporting genetic influences in temperament suggests that both of the above possibilities are likely. For example, the fact that emotionality is genetically influenced means that more emotional mothers will tend to have more emotional babies. After the child's birth, the factors which determine whether the mother or the baby has more influence can only be discovered through longitudinal research which assesses the parents and their environment before the child's birth.

Temperament and Constitutional Measures

In the period prior to the NYLS, some researchers did attempt to relate adult temperament to physical measures. Kretschmer (1925) was one of the first psychologists to study temperamental differences in behaviour, and he proposed an association between schizophrenic and manic-depressive psychopathology and physique. Sheldon and Stevens (1946) formulated a more elaborate association, proposing three body types (endomorphs, mesomorphs and ectomorphs) and three corresponding categories of temperament (visceratonia, somatotonia, and cerebrotonia). While some research supported Sheldon's ideas, they were also criticized because it was not possible to determine whether body type existed

because of temperament or because, as children, they were fed differently and expected to behave in certain ways because of their body type. As well, no theoretical framework existed for the observed relationships which ran counter to theory (Westman & Baldwin, 1973). Eysenck (1953) has also conducted some research in this area. He found relationships between physiological responses and the two factors of neuroticism and introversion-extroversion. However in the period since the NYLS began studying children rather than adults in the 1950´s, little has been done to relate temperament scores to constitutional measures. There are only two studies which attempted to relate temperament scores to physiological measures.

Garcia-Coll et al. (1984) studied 21-22 month old infants. Parents completed the toddler TTS independently, and their ratings of the child´s approach/withdrawal were related to a measure of behavioral inhibition to novel stimuli in the laboratory (r=.54 for mothers and .49 for fathers). Withdrawing/inhibited infants showed higher HR levels and lower HR variability than did more approaching/ uninhibited infants. This pattern of the inhibited infants´ HR responses was attributed to continued vigilance in the situation. When other infants relaxed after an initial period of adjustment to the lab, their HR decreased and fluctuated with respiration. Inhibited subjects continued

their vigilance, thereby maintaining high HR levels which did not fluctuate with their breathing pattern.

Shane, Brodsky, and Brodsky (1982) had teachers rate 45 kindergarten children (aged 60-70 months) with the teacher TTQ and identified three difficult, nine slow-to-warm-up, and 14 easy children. Due to the small sample size, the slow-to-warm-up children were included with the difficult group. Subjects completed three equally difficult sections of the Matching Familiar Figures Test under three consecutive conditions of reinforcement, punishment, and reinforcement while their galvanic skin response (GSR) was recorded. GSR magnitudes discriminated significantly between the two groups. As predicted, difficult children's GSR was significantly higher during punishment than reinforcement conditions. There was also a significant treatment effect. It was predicted that difficult children would perform more poorly than easy children following punishment, and this hypothesis was supported.

The Current Research Area and Hypotheses

It is apparent from the foregoing literature review that while temperament is generally assumed to be related to underlying physiological differences, there is very little research available to validate this assumption. The discussion of "constitutional differences in reactivity and self-regulation" offered by Rothbart and Derryberry (1981,

1982) served as a springboard for this study. Given Rothbart and Derryberry´s emphasis on reactivity and self-regulation, the present author chose measures of each of these processes for this research. The autonomic nervous system (as well as the somatic and endocrine systems) was discussed by Rothbart and Derryberry as reflecting the individual´s response to the environment.

Brodsky and Brodsky (1978) have reviewed the literature (especially Woodworth & Schlosberg, 1954) which indicated that GSR reflects activation or arousal of the general nervous system. Similarly for heart rate, Graham and Clifton (1966) integrated Sokolv´s (1963) work on the orienting response with Lacey´s (1959) suggestion that HR deceleration reflected "stimulus intake." The HR deceleration response has been widely used in infant research to assess infant learning and perception (see Von Bargen, 1983, for a review). This large body of research supports the use of HR and GSR as reflecting the response of the organism to stimulation.

As indicated above, previous research (Lynn, 1966; Milner, 1970) supports the sensitivity of the autonomic nervous system to the environment. Further, Duffy (1972) summarized the research indicating that the autonomic nervous system may reflect individual differences in responsivity to environmental stimuli. For example, Wenger (1941) studied several autonomic measures, including HR and GSR, in 48

children (aged 6 to 12 years) for a period of one year. He found considerable consistency of his measures of autonomic balance over a one year period, with correlations ranging from .42 to .69. While this was not a twin study, the pattern of data for the twins in his sample suggested a genetic influence. Further, Wenger observed relationships between temperament and autonomic balance, and suggested that "children with high scores are less emotional, more controlled in behavior, and more shy than those with low scores" (p.433).

For reactivity measures, the present author chose the magnitude of HR and GSR responses to stimulation. Research supporting the sensitivity of the HR response to stimulation comes primarily from Lipton, Steinschneider and Richmond (1961). These researchers conducted extensive investigations of infants´ HR response to stimulation, and found that their sample could be subdivided into three groups on the basis of individual differences in response magnitude. Further work by the same group (Lipton, Steinschneider & Richmond, 1966) indicated that the magnitude of HR change to stimulation was stable from 2 1/2 to 5 months of age.

Further support for using the magnitude of HR change to stimulation as an indicant of reactivity comes from the following three studies. Vandenberg, Clark and Samuel (1965) found support for a genetic influence on HR responses to

startling stimuli in older children. Garcia-Coll et al. (1984) provide further support for a connection between temperament and HR. She reported that toddlers who were rated by their parents as being more withdrawing to new stimuli had higher HR levels and less HR variability when exposed to new stimuli. Von Bargen (1983), in reviewing the literature on infant HR responses, found HR change to stimulation to be one of the most stable and reliable measures. To summarize, research generally indicates that HR change to stimulation is sensitive to environmental stimulation, reflective of stable individual differences, and genetically influenced.

A similar measure was used for GSR reactivity. Support for the use of the magnitude of the GSR response to stimulation comes from Brodsky and Brodsky (1978) who indicated that this measure reflects arousal or energizing of the organism in response to stimulation. For example, Tarchanoff (1976) found that GSR was very sensitive to various types of environmental stimulation as well as various mental activities. Corah and Stern (1963) have established the reliability and stability of GSR measures. Further support for this measure comes from the study by Shane et al. (1982) in which the magnitude of GSR responses significantly differentiated difficult from easy children, with difficult children having larger responses. The GSR measure has also

been used in relation to some adult temperament measures (e.g., Eysenck & Levy, 1972; Teplov, 1972).

The next task was to choose measures for the self-regulation aspect of Rothbart and Derryberry's (1981) definition. They mention several types of self-regulation processes including approach and avoidance behaviors, self-stimulation and self-soothing, and attention. Attention has been studied extensively by examining the habituation of various physiological responses (Lewis, 1970; Lewis & Baldini, 1979). Habituation has also been used as an indication of the ability to orient to stimuli, and to form an internal representation of stimuli (Graham & Clifton, 1966; Lacey, 1959; Sokolov, 1963, 1969). Consequently, the habituations of GSR and HR responses to stimulation were chosen to assess the self-regulation aspect of Rothbart and Derryberry's definition.

Now that physiological measures have been chosen, temperament measures must be chosen. As noted in the introduction, numerous scales and questionnaires exist to assess temperament in children of all ages. Since this research would be carried out with a preschool sample, three measures suitable for use with this age group were chosen by the present author. These were the BSQ (Carey & McDevitt, 1978) from the NYLS school, the EAS (Buss & Plomin, 1984) from the Colorado researchers, and the CCTI (Rowe & Plomin,

1977), a measure created in an attempt to integrate the first two. It was initially proposed to use the earlier form of the EAS (i.e., the EASI), but since the measure had been revised (Buss & Plomin, 1984) the most recent version was used.

Several subscales of these temperament measures seem likely, on the basis of face validity, to be related to the measures of reactivity and self-regulation. The BSQ dimensions of threshold, intensity of reaction, and approach/withdrawal all seem related to the organism's reactivity to stimulation. The EAS and CCTI dimensions of emotionality and activity, and the CCTI dimension of reaction-to-food also seem likely to be related to physiological reactivity.

There also seem to be several candidates for relationships with attentional aspects of self-regulation. Behavioral differences in attention are an established part of most, if not all temperament rating scales. For example, the BSQ and CCTI both have subscales labelled attention span/persistence. Other subscales which seem, at least on a face validity basis, to be related to self-regulation or attention are the NYLS adaptability and distractibility subscales and the CCTI soothability subscale.

Some support for expecting these particular temperament subscales to be related to physiological indicants of

reacitivity and self-regulation comes from Plomin (1982). He reviews the work of one of the most biologically oriented researchers of adult temperament, Jan Strelau (1969). Strelau has pursued Pavlov´s work on the strength, balance, and adaptability of excitation and inhibition of neural processes, primarily with adult subjects. In discussing this research, Plomin notes relationships between NYLS subscales and Strelau´s concepts, again on the basis of face validity. While not the same as the subscales chosen in this research, there is some overlap with Plomin´s choices. He sees a relationship between Strelau´s concept of "Mobility" or adaptability and the NYLS adaptability and threshold of responsiveness. I included NYLS adaptability along with other scales assessing behaviors related to attention. The main difference between the present research choices and the relationship mentioned by Plomin is in relation to the Strelau concept of "Excitation" or reactivity. Plomin (1982) sees this as related to the NYLS dimensions of activity, attention span/persistence, and distractability. The current researcher focused more on NYLS intensity of reaction, threshold, and EAS and CCTI emotionality, and included the dimensions of attention span/persistence and distractibility as measures of self-regulation.

The first aim of the present research was to provide validational support for the assumption of a constitutional

basis for temperament by investigating relationships between temperament and physiological measures. To summarize the above choices of measures, the magnitude of HR and GSR responses to stimulation were chosen as indicants of reactivity. The habituations of these same responses were chosen as indicants of self-regulation. Three temperament measures suitable for use with preschoolers were chosen: the BSQ, the EAS, and the CCTI. On the basis of face validity, subscales of each of these measures were selected as being likely to be related to the physiological measures.

The main limitation in the research relating the NYLS and Colorado approaches to temperament was that Rowe and Plomin (1977) created their own items to assess the NYLS dimensions, rather than using a measure created by NYLS researchers. Now that several measures exist for this age group, it seemed appropriate to reexamine this relationship. Therefore, a second aim of the present research was to explore the relationships among three different temperament measures and to provide validational support for them by demonstrating their relationships with the physiological measures chosen.

A third issue to be addressed was the consistency of individual physiological responsivity. Lacey and his colleagues (Lacey, 1959; Lacey & Lacey, 1958; Lacey, Bateman, & Van Lehn, 1953) suggested that people did not respond in a consistent manner to stress, but that some might respond with

HR changes while others might show increases in muscle tension or visceral secretions. Lacey´s youngest subjects were 10-year-old children, and it is possible that younger children might evidence more general reactions. Lipton, Steinschneider and Richmond (1960) concluded from their study of newborns that central activation resulted in a generalized behavioral and cardiac response. The behavioral component of this response could be controlled only by an external restraint such as swaddling. With adults, we know that internal excitement is not necessarily related to external behavior. Therefore, it seemed possible that there is a developmental change in the consistency of responding from the generalized responsiveness of infants to the more specific responding of the children and adults studied by Lacey. This possibility was addressed in the present research by comparing HR and GSR measures in a sample of children ranging in age from 3 to 6 years. It was expected that preschoolers might show consistency in their physiological responses (i.e., a stronger correlation between GSR and HR) even when 10 year old children and adults do not.

The temperament literature reveals no consistent indications of sex differences on the various measures. However, there is a large body of research indicating that males are generally more active than females. For example, Eaton and Enns (1986) have conducted a meta-analysis of

available data pertaining to activity level differences and concludes that males are more active than females. Consequently, it was expected that the activity subscales of the three temperament measures would reflect this difference.

In summary, this research was an exploration of the physiological substrates of the behavioral phenomena described as temperament. Since recent work has focused on reactivity and self-regulation, indicants of each of these were chosen for further study. For reactivity, the indicants chosen were HR and GSR responses to new stimuli. For self-regulation, the indicants chosen were the habituations of the initial HR and GSR responses. A second aim was to improve existing temperament measures by exploring their relationships to each other and to physiological processes. The third goal of this research was to investigate consistency between HR and GSR responses in a younger sample than had previously been studied. Specifically, it was predicted that:

1. The reactivity measures, HR and GSR responses to new stimuli, would be significantly related to the following set of eight temperament scores: BSQ threshold, intensity, and approach/withdrawal; EASI emotionality and activity; and CCTI emotionality, activity, and reaction-to-food.

2. The self-regulation measures, HR and GSR habituation, would be related significantly to the following set of five temperament scores: BSQ adaptability, distractibility, and attention span/persistence; and CCTI attention span/persistence and soothability.
3. There would be a significant correlation between the GSR and HR responses to new stimuli.
4. There would be a significant correlation between the GSR and HR habituation.
5. The correlation between GSR and HR responses to new stimuli would have an inverse relationship with age.
6. Children´s gender would only affect BSQ, EASI and CCTI activity scores. Other temperament and physiological scores were expected to be independent of sex effects.

It should be stressed that this is exploratoy research. The above hypotheses were formulated to address the goals of the research, but the goals extend beyond the specific nature of these hypotheses. Consequently, data will be examined for relationships other than those specified here.

## Method

### Subjects

The experimenter attended parents' meetings of a large multi-classroom university preschool (9 classrooms, serving 140 children) and described the research project to parents. Following these meetings, letters describing the goals and procedures of the study and asking for parental consent (see Appendix A) were circulated. Permission was given for 96 children to be included (a participation rate of 69%), and parents of these children were given questionnaires to complete. Parents of 90 children provided complete data on the EAS, parents of 87 children completed all items of the CCTI, and parents of 86 children completed all items of the BSQ. Of the 96 children included, five failed to complete the experiment because of equipment difficulties, refusal to complete the session, or parents not completing questionnaires. Thus, of the original 96, 91 had complete data for all measures which were included in analyses. This number is larger than the numbers completing all items of the scales because subscale scores could still be computed if only a few items had been missed. Of these 91 children, 41 (45%) were female; 50 (55%) were male. Children ranged in age from 41 to 86 months.

### Apparatus

Questionnaires used (see Appendices B, C, and D) included

the Behavioral Style Questionnaire (BSQ; Carey & McDevitt, 1978; McDevitt & Carey, 1978), the EAS (Buss & Plomin, 1984), and the Colorado Childhood Temperament Inventory (CCTI; Rowe & Plomin, 1977). Retest reliability of the BSQ over a one-month period ranged from .67 to .94 (median r = .81, McDevitt & Carey, 1978). Retest reliabilities of the EAS over a one-week period for 31 children (average age of 3.6 years) were .72 for emotionality, .80 for activity, and .58 for sociability/shyness (Buss & Plomin, 1984). Rowe and Plomin (1977) reported that retest reliability of the CCTI over one week ranged from .72 to .80 for four of the six subscales. Lower retest reliability was obtained for the sociability (r = .58) and soothability (r = .43) subscales. Inter-rater reliability is only available for the EASI (not the EAS). Corsini and Doyle (1981) reported that agreement between mothers and teachers ranged from .34 to .51. Shane et al (1982) also found relatively low parent-teacher agreement on the EASI.

GSR was monitored continuously during stimulus presentations with a portable Lafayette Psychogalvanometer (model 7609A). Two Bard Biomedical EKG lab electrodes (#160100) were attached, one to the palm and one to the top of the child´s non-preferred hand (determined by noting which hand the child used to write the first letter of his/her name immediately before testing). GSR was recorded on Lafayette

Instrument Company, Inc. rolled chart paper (#71011) at a speed of three inches per minute.

HR was monitored continuously during stimulus presentations with an ECG Monitor (manufactured by Electronics for Medicine, model IR2). Three Bard Biomedical EKG lab electrodes were placed on the subject´s chest in an inverted triangle array (left and right shoulders and left waist). HR data was recorded on General Medical Corporation ECG Recording Charts (#26-140) at a speed of 25 mm per second.

Two tones and two lights were used to assess the subjects´ reactivity and self-regulatory abilities. Each stimulus was presented six times to assess the child´s ability to habituate. For both GSR and HR, Graham (1973) reported that the largest decrement occurred between trials 1 and 2, and that asymptote levels were reached in 3 to 20 trials. Thus, some subjects were expected to reach asymptote while others were not. If sufficient children failed to reach asymptote in the six trials, the present author was concerned that the variability of the habituation measure would be decreased. However, other researchers (see Von Bargen, 1983, for a review) generally choose to present stimuli for 6 trials, and this number was adopted rather than some larger number of trials because of a concern for the session duration. Preschoolers cannot sit still in an

experimental session for long periods of time, and this research required them to sit through six presentations each of four stimuli.

Auditory stimuli were generated by an audio-oscillator delivering 500 Hz warble and 1000 Hz narrow band noises at 75 db, and were presented for 2 second durations with 20 second inter-trial intervals. The warble and narrow band noise were presented to the children via audio recordings, with the recorder placed on a desk beside them. Visual stimuli were presented by illuminating the wall in front of the subjects wth either a red or blue 35 mm slide. Lights were presented for 2 seconds by a Model 4400 Kodak Carousel slide projector, also with a 20 second inter-trial interval.

Procedure

Parents who indicated their willingness to participate were given an envelope containing copies of the three temperament questionnaires and instructions for completion. If there were two parents in the family, they were asked to complete the measures collaboratively in order to obtain a more reliable assessment of the child. Collaborative ratings were requested because Epstein (1979, 1980) suggested that reliability could be increased if one combined repeated measures of variables or measures by more than one person. Eaton (1983) applied this technique to assessment of activity level and found good reliability levels with aggregate

ratings made by several preschool teachers.

The experimenter spent one day in each classroom prior to testing to ensure that the children were all familiar with her. On a second day the experimenter and one teacher took the class in small groups to see the experimental room. The procedures were described to the children, and they were told that the experimenter would ask some of them to come to the room again later to "hear some sounds and see some slides." On a third day, the experimenter began to approach children for data-collection sessions. Any children who were unwilling to participate at this point were reapproached at another time. If they were still reluctant, they were dropped from the sample (this occurred only with one child).

When children were taken to the experimental room, they were seated at a child's table and asked to write the first letter of their name in order to determine their preferred hand. Electrodes were placed on their non-preferred hand and chest, and children were asked to sit quietly while the experimenter calibrated the psychogalvanometer. This time was also used to allow the child to adapt to the situation before administering stimuli. For GSR, this meant that the session did not begin until a resting base level, a parallel or declining recording occurred. The session did not begin until HR was also stable or declining.

A stable or declining record prior to stimulus onset

allows more reliable discrimination between baseline and response levels. For example, with GSR the response is a fairly sharp increase in level which gradually returns to baseline. If baseline levels are gradually increasing, it is conceivable that an increase in the slope of this trend could be misinterpreted as a response. The situation was somewhat different for HR. The expected response to the stimulus was a deceleration, so that a declining trend could be mistaken for a response. However, even if the rate was gradually declining, data were examined for a sharp change of slope indicating a clear response to stimulation.

Once the physiological measures had stabilized, the experimenter explained to the child the importance of looking at the wall in front of them so they could "see the special slides," listening carefully so they could "hear the special sounds," and that she would not be able to talk to them during this period.

The four stimuli (light through red and blue slides, noise and warble sounds) were presented in counterbalanced order to control for fatigue and order effects, with modality alternating (i.e., sound-light-sound-light, or the reverse). There were eight orders, and the particular order to which a child was exposed was randomly determined. Since each stimulus was repeated six times, children were exposed to 24 stimuli in all (six repetitions each of the two sounds and

two lights). For example, in one of the eight orders, children were exposed to light from the blue slide six times, the warble sound six times, light from the red slide six times, and then the noise sound six times. When a sound or light was presented, the experimenter activated an event marker to indicate which portion of the child's GSR and HR occurred after the stimulus presentation. Between each set of sound or light repetitions, the experimenter ensured that the child was comfortable and facing forward. When all four stimuli were presented, the experimenter detached all electrodes, thanked the child, allowed him/her to choose a small reward, and escorted the child back to the classroom. This entire procedure required approximately 30 - 45 minutes.

Scoring

Temperament questionnaires were scored according to published guidelines into subscale scores. This yielded 18 scores; nine scores for the BSQ, four scores for the EAS, and six scores for the CCTI. As well, diagnostic cluster scores were obtained for the BSQ.

HR and GSR recordings were scored to determine the magnitude of change from baseline values immediately following stimulus presentations. For GSR, the size of the deviation from baseline was determined by counting the number of graph lines on the chart paper from baseline to the highest point of the response curve. As indicated by Brodsky

and Brodsky (1978) this measure accurately reflects resistance/conductance changes.

For HR, baseline was calculated as the rate averaged across the three beats immediately preceding each stimulus presentation. The change from baseline was calculated as the difference between baseline and the average of the three slowest consecutive beats occurring in the ten-second period immediately following the stimulus presentation. From these GSR and HR change scores, two indicants of reactivity were computed. These were the First response (the magnitude of the response to the first presentation of each of the four stimuli) and the Maximum response (the largest response to each of the four stimuli).

Self-regulation was indicated by the child's ability to habituate to repetitions of the same stimulus. Habituation was operationally defined as the difference between the average magnitude of the first three and the last three responses to each stimulus.

To summarize, these scoring procedures resulted in the following measures. For temperament, the three scales were scored into 18 subscale scores plus the BSQ diagnostic cluster score. There were two measures for reactivity (the First response and the Maximum response) and one measure for habituation (the Difference score). Each of these three measures was computed for GSR and for HR for each of the four

stimuli, resulting in 24 physiological measures (12 each for GSR and HR).

## Results

Data analyses are presented in the following order. First, temperament scores of this sample are compared with scores of other samples (both in terms of average scores and the results of factor analyses) to help determine the generalizability of results. Second, a section labelled "Data reduction" presents the results of a factor analysis of the 19 subscale scores from the three temperament measures. The factor scores which resulted from this analysis were used in later analyses.

Next, the third, fourth and fifth sections address the main aim of this research, exploration of the relationships between temperament and physiological measures. These sections present several analyses which were not specifically included in the a priori hypotheses, but which were conducted because of the exploratory nature of this research. The third section describes the results of canonical correlations used to test the first hypothesis regarding reactivity. The fourth section is concerned with canonical correlations used to test the hypothesis regarding self-regulation. The fifth section presents a repeated measures multivariate ANOVA testing the relationship between children´s easy-difficult

diagnostic clusters and physiological measures.

After this group of three sections, the sixth section includes analyses which address the third, fourth and fifth hypotheses regarding relationships between GSR and HR reactivity and habituation, and the possible effect of age on these relationships. The seventh section addresses the hypothesis concerning gender effects. For reference purposes, Table 1 presents the correlations between all measures, and Table 2 presents the means and standard deviations of all the measures.

---

Insert Tables 1 & 2 about here

---

Comparison of present and previous results

Scores on temperament measures from this sample conform quite closely to those from previous samples. Figure 1 depicts the correspondence of the present BSQ scores with those of McDevitt and Carey (1978), and Figure 2 compares the present CCTI scores with those of Rowe and Plomin (1977). It was not possible to present the comparison of current EAS scores with Buss and Plomin (1975 & 1984) because data are in different forms and not in comparable age groups. However, scores are similar.

---

Insert Figures 1 and 2 about here

---

Principal components factor analyses were performed on the EAS and the CCTI, separately and together, to see if current factor structures replicate earlier research on these measures. Varimax rotations were used for components with eigenvalues greater than one. Decisions regarding which factors to retain were made according to both variance accounted for and scree tests (Cattell, 1978).

Results with the 20 EAS items were five factors with eigenvalues greater than 1.00, accounting for 70% of the variability. Interpretation of these five factors was also supported by scree test (Cattell, 1978). While results generally supported Buss and Plomin´s (1975, 1984) conceptualization of emotionality, activity, and sociability/shyness as temperaments, there were some differences. The first two factors, accounting for 27% and 19% of the variance, respectively, were clearly interpretable as emotionality and activity. Items with highest loadings on these first two factors were the same as those identified by the scale´s designers, emotionality and activity. The main difference was that there were three rather than one sociability/shyness factors, suggesting that these may not be discrete temperament categories.

The third, fourth and fifth factors (11%, 8%, and 5% of the variance) were primarily composed of shyness items. The third factor included three sociability items with moderate coefficients and a shyness item with the highest coefficient, and seems closest to Buss and Plomin´s sociability/shyness dimension. The fourth and fifth factors had only shyness items with high loadings.

Factor analysis of the CCTI also resulted in similarities and differences with earlier research (Rowe & Plomin, 1977). The first factor was clearly interpretable as Sociability, while the third was clearly interpretable as Emotionality. These two factors accounted for 20% and 10% of the variability, respectively. However, there was no factor identifiable as Activity Level or Reaction-to-Food. Instead, there were three factors (#2, 5, and 7) clearly identifiable as Attention Span, each leaving highest loadings on items from the Attention Span subscale. These three factors accounted for 11%, 7%, and 4% of the variance. There were two factors with highest loadings on items from the Soothability subscale which accounted for 8% and 6% of the variablity, respectively.

When these two scales were submitted together to principal axis factor analysis, eight factors were chosen to be interpreted, based on variance accounted for (65% in total) and scree test. These factors were identified as

Emotionality (19% of variance), Sociability (13%), Attention Span (9%), Activity Level (6%), Reaction to Food (6%), a second Activity Level factor (5%), Shyness (4%), and Soothability (3%). Since the results of the principal axis factor analyses of the EAS and CCTI were fairly similar to earlier findings, factor scores were not computed for separate analysis. It was not possible to submit the BSQ to principal axis factor analysis due to insufficient subjects and the large number of items in the questionnaire.

Data reduction

The EAS, CCTI and BSQ subscale scores were computed according to published guidelines. The resulting 18 subscale scores were submitted to principal axis factor analysis, resulting in five factors with eigenvalues greater than 1, and accounting for 69% of the variability. They were labeled Emotionality (26%), Shyness/Sociability (18%), Attention Span (11%), Activity (7%), and Threshold/Distractibility (7%). The subscales which loaded highest on each factor and their factor loadings are listed in Table 3. Table 4 presents the same information in a rearranged and simplified format to ease interpretation. Scores were computed for all subjects on these five factors and used in later analyses.

---

Insert Tables 3 and 4 about here

---

Reactivity Hypothesis

The first hypothesis of this research was that a set of eight temperament scores (the "reactivity set") would be related to physiological reactivity scores (First and Maximum scores). This hypothesis was tested by two canonical correlations. The first included the reactivity set of eight temperament scores and the First HR and GSR scores. The second canonical correlation included the same set of temperament scores and the Maximum HR and GSR scores. Although not specified in the first hypothesis, the relationship between temperament and physiological reactivity was also explored with canonical correlations between physiological measures and two other sets of temperament scores: the set of all 18 temperament scores, and the set of five temperament factor scores. Also, since children were expected to be most reactive to the first of the four stimuli, the first stimulus (not the second, third, or fourth) was analyzed separately in one canonical correlation. The results of all these canonical correlations are summarized in Table 5. None of these canonical correlations was significant. HR and GSR Maximum and First scores were not significantly related to the reactivity set of

temperament scores, to the set of all 18 temperament scores, or to the set of five temperament factors.

---

Insert Table 5 about here

---

Self-regulation Hypothesis

The second hypothesis of this research was that HR and GSR habituation, operationally defined as the difference between the averages of the first three and last three responses, would be significantly related to the following set (the "habituation set") of temperament scores: BSQ adaptability, distractibility, and attention span/persistence; and CCTI attention span/persistence and soothability. This hypothesis was tested by canonical correlation analyses between two sets of physiological measures (HR & GSR Difference scores for all four stimulus presentations, and for the last stimulus alone), and three sets of temperament scores (all 18 temperament scores, the habituation set, and the five temperament factor scores). As noted, the Difference scores for the last stimulus presented were examined separately because children were expected to habituate most to this stimulus since it was presented when they were most comfortable in the situation. Separate analyses were not conducted for the other three stimuli. The results from these canonical correlations are presented in

Table 6.

---

Insert Table 6 about here

---

While all canonical correlations which included the Difference scores from four stimuli were non-significant, the analyses which included only the Difference scores from the last stimulus were marginally significant. The canonical correlation between the five temperament factor scores and HR and GSR Difference scores for the last stimulus presentation resulted in a canonical variate which accounted for .16 of the shared variability ($X^2_{(10,\ N\ =\ 85)} = 17.86$, $p = .057$). The analysis which included all 19 temperament scores resulted in a canonical variate which accounted for .34 of the shared variance ($X^2_{(38,\ N\ =\ 85)} = 52.35$, $p = .060$). The canonical correlation including the habituation set resulted in the smallest canonical variate, accounting for .15 of the shared variability, and the chi-square test of this relationship was the least significant ($X^2_{(10,\ N\ =\ 85)} = 16.54$, $p = .085$).

Follow-up Analyses. The relationships indicated by these canonical correlations were explored further by performing simultaneous and stepwise multiple regressions between HR and GSR Difference scores (separately) for the last stimulus presented, and the three sets of temperament scores: all 18

scores, the habituation set, and the five factor scores. These multiple regressions would indicate which variables contributed most to the significant canonical correlation results. As seen in Table 7, only the simultaneous regressions which included GSR Difference scores were significant or marginal. None of the regressions which included the HR Difference scores were significant. Consequently, relationships with HR scores were not explored further.

---

Insert Table 7 about here

---

Stepwise multiple regression analyses provided further clarification of the significant relationships between GSR and temperament scores. The analysis between GSR Difference scores for the last stimulus and the habituation set of temperament scores included in hypothesis 2 proceeded for three steps (see Table 8).

---

Insert Table 8 about here

---

BSQ adaptability entered on the first step and accounted for 9% of the shared variance. CCTI attention span entered second, adding 3% of variability, and BSQ distractibility entered last, adding a further 2% of variability. Overall,

the combination of these three variables accounted for 14% of the variability in GSR Difference scores for the last stimulus presented. Since large positive GSR Difference scores indicated that the child did not habituate to repeated stimulus presentations, an examination of the regression coefficients indicated that children who habituated less were slower to adapt to new situations, had shorter attention spans, and were more distractible (i.e., had higher scores on BSQ adaptability, lower scores on CCTI attention span, and higher scores on BSQ distractibility, respectively).

The analysis including all 18 temperament scores proceeded for eight steps, with variables entering (or being removed) in the following order: BSQ adaptability, BSQ activity, EAS shyness, BSQ adaptability removed (indicating that, with the activity and shyness scores included, the adaptability score no longer predicted a significant unique amount of variability), BSQ rhythmicity, BSQ approach, CCTI attention span, and CCTI reaction-to-food (see Table 9).

---

Insert Table 9 about here

---

The BSQ approach score had a small positive correlation with the GSR Difference criterion, but a large negative regression weight indicating suppression effects. Thus, the BSQ approach score was more important to the regression

equation by accounting for error in the other dependent variables than for its relationship with the GSR Difference score for the last stimulus presented.

Overall, the regression equation accounted for 26% of the variability in GSR Difference scores for the last stimulus presented. The results of this analysis indicate the optimal set of temperament scores for predicting children´s GSR Difference scores. Children who habituated less were more shy (EAS shyness), less approaching to new stimuli (BSQ approach), more active (BSQ activity), less rhythmic in biological functioning (BSQ rhythmicity), had shorter attention spans (CCTI attention span/persistence), and reacted more to new foods (CCTI reaction-to-food). Children´s Difference scores were related to other temperament subscale scores (e.g., BSQ adaptability), but these other subscale scores do not add significant unique amounts of variability to the prediction equation including the above 6 subscales.

The analysis between the five temperament factors and the GSR Difference score for the last stimulus proceeded for four steps. Children who habituated less were more distractible, more shy, more active, and more emotional (had higher scores on the Distractability, Shyness, Activity Level, and Emotionality factors). Overall, the relationship accounted for 16% of the variability in GSR habituation and was

statistically significant (see Table 10).

---

Insert Table 10 about here

---

Repeated Measures Analysis

While the first hypothesis specified eight temperament scores expected to be related to physiological indicants of reactivity, and the second hypothesis specified five temperament scores expected to be related to physiological habituation, this research was designed with a more general goal: to explore whether physiological correlates of parental ratings of temperament existed. Thus, as seen earlier, analyses have not been limited to the sets of scores specified in the first two hypotheses. For both hypotheses analyses were completed on two other sets of temperament scores: the full set of all 18 measures, and the set of 5 factors which resulted from principal components analysis of those 18 scores. Another way of exploring the relationship between physiological measures and temperament ratings was to use the BSQ diagnostic cluster categories.

Children were categorized as easy (N = 41), intermediate-low (N = 35), intermediate-high (N = 11), or difficult (N = 8), according to published criteria (McDevitt & Carey, 1978). There was only one child categorized as slow-to-warm-up, so this category was dropped. With these

groupings, it was possible to complete a multivariate repeated measures ANOVA on the physiological data. This analysis used the multivariate approach to repeated measures ANOVA to avoid the restrictive assumptions of the univariate repeated measures ANOVA.

This analysis resulted in several significant effects. However, examination of these effects revealed that the HR scores were never different in any of the multivariately significant effects. Therefore the analysis was repeated using only the GSR scores. Table 11 presents the multivariate repeated measures ANOVA results on GSR scores.

---

Insert Table 11 about here

---

Significant effects were found for the main effects of stimulus ($F_{3, 76} = 8.35$, $p < .001$) and trials ($F_{5, 74} = 17.30$, $p < .001$). The main effect for diagnostic cluster was non-significant ($F_{3, 78} = 1.97$, $p < .126$) but diagnostic cluster interacted significantly with stimulus ($F_{9, 185.11} = 2.72$, $p < .006$) and with trials ($F_{15, 204.68} = 1.95$, $p < .022$). Children's GSR responses habituated across the 4 stimuli as indicated by their response magnitude changes. As shown in Figure 3, the children's responses decreased in magnitude across the 6 presentations of each stimulus and dishabituated when the stimulus changed.

---

Insert Figure 3 about here

---

Figures 4 and 5 show the stimulus by diagnostic cluster, and trials by diagnostic cluster interactions, respectively. No other significant effects were found.

---

Insert Figures 4 and 5 about here

---

Post hoc analyses (Tukey's) revealed a significant difference between the difficult group and the other three groups on the first stimulus, and a significant difference between the difficult group and the easy group on the last two stimuli. In both cases the difficult children demonstrated larger responses than did the easy groups. There was also a tendency for the easy children to habituate on successive stimulus blocks. Both the difficult and the easy groups showed habituation across the first three stimulus blocks, and the difficult group dishabituated on the final stimulus block.

Analyses of the Consistency of Physiological Responses

The third and fourth hypotheses regarding relationships between GSR and HR were tested by examining correlations between GSR and HR Maximum, and GSR and HR Difference scores. Both simple correlations and partial correlations

(partialling out the effects of age, gender, and order of stimulus presentation) were computed. These are presented in Table 12.

---

Insert Table 12 about here

---

The only significant relationship between HR and GSR was found for the second pair of the Maximum scores. The significant negative correlation ($r_{(92)} = -.312$, $p < .01$) indicated that large HR decelerations were associated with large GSR increases. This provides only limited support for hypothesis 3 as the relationship was nonsignificant on three of the four trials.

Hypothesis 4, that GSR and HR habituation scores would be correlated, received no support as shown in Table 12. None of the relationships between HR and GSR changed when age, gender and order effects were partialled out. In other words, the relationship (or lack thereof) between HR and GSR did not depend on how old the child was, whether the child was male or female, or in what order stimuli were presented. This lack of effect of age on the relationship between GSR and HR refutes the fifth hypothesis of this research.

Gender Effects

The final hypothesis predicted that the child's gender would only affect activity scores on the temperament

measures. An examination of the correlations between gender and the 18 temperament subscale scores indicated that no significant relationship exists with CCTI activity ($r_{(90)}$ = .14), EAS activity ($r_{(90)}$ = .05), or BSQ activity ($r_{(90)}$ = .09). The only correlation with gender which was significant was that with EAS sociability ($r_{(N = 90)}$ = -.21, t = -2.02, p < .05), but the correlation with CCTI sociability was nonsignificant. The correlation between CCTI reaction-to-food and gender was marginal (r = -.187, $t_{(88)}$ = -1.78, p < .08). This hypothesis was also examined in relation to the temperament factor scores, and the only correlation which was significant was the one with the Threshold factor (r = -.21, $t_{(88)}$ = -2.024, p < .05). The correlation between the Activity Level factor and gender was marginally significant (r = .18, t = 1.78, p < .08). Taken together, the significant results indicate that girls were rated as being more sociable (but only on the EAS, not on the CCTI) and as having lower thresholds than boys. The marginal results suggest that girls were rated as reacting more to new foods and as being less active than boys.

## Discussion

The central issue of this research was whether or not any

relationship(s) existed between the physiological measures chosen and parental ratings of children´s temperament. Such a relationship has been assumed by most researchers in this area but has received little empirical attention. The issue was addressed in three separate sets of analyses. First, canonical correlations were completed between indicants of physiological reactivity (GSR and HR First and Maximum scores) and temperament scores. Second, canonical correlation analyses were carried out between temperament scores and physiological indicants of habituation (HR and GSR Difference scores). The third set of analyses involved the multivariate repeated measures ANOVA between physiological scores and the BSQ diagnostic cluster, which was used as a grouping variable.

None of the canonical correlation analyses between HR or GSR reactivity measures (First response and Maximum response) and temperament scores were significant, indicating no support for the first hypothesis from these analyses. With regards to the second hypotheses, significant relationships were found between temperament scores and the physiological indicants of habituation for the last of four stimulus blocks. Follow-up analyses indicated that these results were due only to GSR scores, and not to HR scores. Stepwise multiple regression analyses were carried out to determine which set of temperament scores best predicted the GSR

habituation scores for the last stimulus presented. All but one of the subscales included in this set assessed attention span/activity level, or initial reactions/adaptations to novel stimuli. Children whose GSR responses habituated more were rated as being less active and distractible, more attentive, less shy and emotional, and more sociable. Thus, the second hypothesis that temperament is related to physiological processes of adaptation or habituation received more support than did the first hypothesis that temperament was related to physiological reactivity. However, it is pertinent to the first hypothesis (concerning reactivity) that several subscales included in this stepwise multiple regression included indications of initial reactions to novel stimuli. The Difference score used to assess the second hypothesis was affected both by the level of the initial reaction as well as by habituation across repetitions, suggesting that these significant results may provide some indirect support for the first hypothesis.

The multivariate repeated measures ANOVA (and follow-up analyses) indicated relationships between GSR scores and BSQ diagnostic cluster categories (Difficult, Intermediate-high, Intermediate-low, and Easy). Children in the Difficult group had larger GSR responses than did children in the Easy group to the first and last stimuli. In addition to demonstrating larger GSR responses to these stimuli, Difficult children

differed from Easy children in their pattern of responses across the four stimuli. Easy children habituated gradually across the four stimuli while Difficult children showed a large decrease from stimulus 1 to stimuli 2 and 3, and then an increase in response magnitude to the fourth stimulus.

This difference in the response pattern between Difficult and Easy children was also apparent within the six repetitions to the four stimuli. Again, Difficult children´s initial responses were much larger than Easy children´s initial responses. Difficult children´s responses indicated large decreases across repetitions two and three, increases to repetitions four and five, and then a decrease again to repetion six. This erratic pattern contrasted with the smoothly decreasing habituation pattern shown by the Easy group. There was no significant relationship between HR scores and diagnostic cluster.

Several questions occur when an attempt at integration and explanation of these results is made. In reference to the first hypothesis, why were there significant relationships between physiological reactivity and temperament scores in the multivariate repeated measures ANOVA and not in the canonical correlations? Why were the ANOVA and canonical correlation results only significant for GSR scores and not for HR scores? Why were the canonical correlations between temperament scores and Difference scores

only signficant for the last of four stimulus blocks and not for all four?

In response to the first of these questions, it may be that other relationships existed but were not detected because this exploratory study included a large number of variables. The large number of variables may have decreased power so much that less robust relationships could not be detected. Tables 3 and 4 certainly indicate increasing significance of results when the number of variables included in analyses decreased. The last row in each of these tables includes a test of scores from only one of the four stimuli while those above include scores from all four stimuli. For the self-regulation hypotheses, these are the only tests which are marginally significant. None are significant in Table 1 for the reactivity hypothesis, but p levels change from a range of .50 - 1.00 to .17 - .37. Hopefully, future researchers will be able to use the results of this study to decrease the numbers of variables included in their projects. For example, since there was no effect of stimulus modality, this variable may be excluded from future experiments. A second and related response to this first question concerns the fact that the equipment available was not "state-of-the-art". This may also have increased the amount of error variance in the data, thereby reducing power and making it harder to detect differences.

A third response to this first question is that children´s scores on separate subscales may be less important (in terms of a relationship with their physiological responses) than how they vary on an important subset of temperament scores. The results suggest that the BSQ diagnostic cluster assesses an important difference in how environmental stimuli are perceived and processed by difficult versus easy children. The NYLS researchers (Thomas, Chess, Birch, Hertzig & Korn, 1963) arrived at the diagnostic cluster groupings via nonstatistical means. This analysis received some statistical validation when the first factor resulting from factor analysis of their data combined four of the five subscales included in the diagnostic cluster. As noted in the review section on factor analyses, a difficultness-easiness dimension has been found in several studies using both factor and cluster analytic techniques (eg., Bates et al, 1979; Cameron, 1978; Pain, 1980; Scholom, et al., 1979). Although the validity of a difficultness dimension has been questioned in recent papers (eg., Bates, 1980), the results of this and other research indicates that this dimension merits further research.

In response to the second question (concerning the reason(s) for results only being significant for GSR and not HR scores), it is possible that the unsophisticated equipment used had more effect on HR than on GSR scores. The latter is

a slower and cleaner response and seemed, subjectively, to reflect more clearly the extent to which stimuli were perceived and processed. It was certainly expected from earlier research that HR would have reflected stimulus perception and processing, but it was more difficult to detect this from the record. More sophisticated equipment would have been capable of transforming the beat-to-beat interval to heart rate, with extraneous variability (for example, from respiraton) removed. Not having this capability may have affected the results. Significant relationships between HR and temperament may have been found with a more sophisticated cardiac monitor, and I recommend that future researchers ensure the availability of such equipment before embarking on a similar research project.

Another response to this question concerns anticipatory responses. When scoring the data it was sometimes noticed that HR began to decrease just before the stimulus was presented. This suggests that some children anticipated the next stimulus presentation towards the end of the 20-second ITI, and responded with HR deceleration. Consequently, when the stimulus was presented, their HR showed only a very slight deceleration or a slight acceleration. A randomly varied ITI would be preferable to a constant interval in future research.

The third question concerned why relationships between

temperament and habituation were only significant for the last stimulus block and not for all four stimulus blocks. One possibility is that results may not have been sufficiently robust to emerge when all variables were included in analyses. It also seems possible that children may not have been sufficiently secure or relaxed in the situation to show habituation clearly in six trials until the last stimulus block was presented.

The most interesting results, which pertain to both the first and second hypotheses, are clearly those revealed by the multivariate repeated measures ANOVA, that difficult children´s GSR responses were different from easy children´s, both in terms of initial reactivity and habituation patterns. As this is exploratory research, interpretation of these results must be tentative. One possible interpretation is suggested in a paper by Brodsky and Brodsky (1978). Difficult children who have low thresholds and are very sensitive to stimuluation, and who give large responses when stimulated, may often feel overwhelmed by stimulation. Brodsky and Brodsky (1978) speculated that, in defense to this overstimulation, difficult children may react by raising their threshold so that they will be less sensitive to further stimulation. Thus, a child might have large responses to the first stimulus, feel overwhelmed and defensively raise his/her threshold to stimulation.

Consequently, he or she would have almost no response to the next few stimulus blocks because stimuli were too low for their now-raised threshold. By the time the last stimulus was presented, the children´s thresholds might have decreased sufficiently for them to perceive and respond again. Possible brain mechanisms for this process would include an overly sensitive reticular activating system and increased cortical inhibition following initial stimulation. Indeed, while collecting data it did seem that children whose initial GSR responses were very large would show little or no response to the next few stimuli, and then would show large responses again to the last stimulus.

Another possible explanation is that these children had short attention spans, and that once the novelty of the first stimulus had worn off, they were no longer interested in the next few stimulus blocks. They attended to other stimuli in the room. By the time the fourth stimulus was presented, room stimuli no longer interested them and they attended again to the experimental stimulus.

Whatever the explanation, it is clear that results of this study must be interpreted from a broader perspective than a simple summary of whether specific hypotheses were or were not supported. Since this was exploratory research, specific hypotheses were formulated to address whether physiological correlates of temperament measures existed.

The question underlying these specific hypotheses is more general, and this question must be addressed as well. Thus, analyses of the data resulting from this study went further than tests of specific hypotheses. Data were examined in several ways and many post hoc analyses were carried out. These analyses were valuable in the exploration of the general question, but must be interpreted cautiously due to the problem of experiment-wide error rate. Also, any answers provided to the general question from these post hoc and exploratory analyses must be tested as predictive hypotheses in future research before conclusive answers can be determined.

With these cautions in mind, the answer to the question about the existence of physiological correlates of temperament ratings must be affirmative. A relationship was demonstrated between the habituation pattern of GSR responses and temperament subscales assessing reactions to novel stimuli and attention. As well, a relationship was demonstrated between GSR responses and the NYLS diagnostic cluster scores.

The third, fourth, and fifth hypotheses, which predicted that GSR and HR indicants of reactivity and habituation would correspond, and that this correspondence would decrease with age, received no support from these data. This may indicate the correctness of Lacey´s theory of autonomic

response-stereotypy (Lacey, 1956; Lacey, 1959; Lacey, Bateman & VanLehn, 1953; Lacey & Lacey, 1958a,b). However, Duffy (1972) discusses the methodological problems which could mistakenly lead to a conclusion that individuals who are more reactive in one response area are not more reactive in others.

One of the issues raised by Levy (1972) seems particularly pertinent here. Vandenberg, Clark and Samuels (1965) found that only intense, startling stimuli revealed a hereditary influence on heart rate and breathing rate. These researchers did not specifically address the issue of intra-individual correspondence of measures, but their study suggests that the lack of concordance between HR and GSR in this study may have been affected by the stimuli used. I purposely chose mild stimuli to avoid startling the children, since I wished to assess their orienting responses and not their defensive responses. As indicated by Garcia-Coll, less approaching infants tend to show higher HR levels and decreased variability in their HR. This may have occurred in this study. If it did, it may have obscured the relationship between HR and GSR. Also, in the earlier discussion it was pointed out that HR responses sometimes seemed to decelerate in anticipation of stimuli so that when the stimulus was presented, the response may have been obscured by the anticipatory response.

Given these problems, it is not possible to draw any conclusions about correspondence between different measures of reactivity in preschool children, even though current data do not provide much support for such a correspondence (only 1 of 8 correlation coefficients was significant, and which could easily occur by chance).

The sixth hypothesis predicted a relationship between gender and parental ratings of activity level. It was surprising to find only a marginal relationship between gender and the Activity Level factor when extensive research indicates that boys have been objectively measured with actometers as being more active than girls (Eaton & Enns, 1986). These results are consistent in direction with this meta-analysis.It may have been that parents in this study were reluctant to rate males as different than females.

The five temperament factors resulting when all 19 temperament subscale scores were submitted for analysis represent supplementary information to those attempting a decision regarding the subcategories of temperament. These five factors, labeled Emotionality (which includes 3 of the 5 NYLS easyness-difficultness subscales), Shyness/Sociability, attention, Activity level, and Threshold/Distractability share marked similarities with those found by other investigators. It seems that when researchers in this area decide on different categories of temperament, their list

will probably include representations of emotionality, sociability/shyness, attention, and activity level.

Summary. The general aim of this research was fulfilled, even though specific hypotheses may not have been supported or may have received only partial support. This study was designed with the general goal of exploring physiological correlates of parental temperament ratings. The first hypothesis, which predicted a relationship between physiological reactivity and eight temperament subscales, received partial support from the multivariate repeated measures ANOVA and other analyses. Some of the temperament subscales included in significant relationships with GSR habituation assessed reactions to novel stimuli. The multivariate repeated measures ANOVA supported the reactivity hypothesis by indicating that difficult children were more reactive to initial and final stimuli, both within and across stimulus blocks, than were easy children.

The second hypothesis predicted a relationship between physiological habituation and five temperament subscales, and received moderate support. Analyses were significant,but only for the final block of stimuli, not for all four blocks. Stronger support for this hypothesis again resulted when children were grouped on the BSQ diagnostic cluster variable. Easy children showed a smooth habituation pattern, both within and across stimulus blocks. Difficult children showed

some habituation, but then dishabituated towards the end of trials or trial blocks.

Possible reasons for these mixed results were discussed, and recommendations made for future investigators (e.g., decreasing the number of variables and incorporating random inter-stimulus intervals into the design). Even though results were mixed in their support of specific hypotheses, relationships between temperament and physiological measures were discovered.

The third, fourth, and fifth hypotheses were all concerned with the correspondence between HR and GSR measures. The third predicted a significant relationship for reactivity measures, the fourth for habituation measures, and the fifth predicted a greater correspondence between GSR and HR for younger versus older children. Only one of several relationships was significant, indicating almost no support for these three hypotheses. The last hypothesis, predicting a significant relationship between gender and ratings of activity level, received marginal support.

In addition to discovering relationships between temperament measures and children´s GSR scores, the other important contribution of this study is the clarification of temperament factor structure. While not a conclusive answer to the question of how many categories of temperament exist, four of the five factors which resulted from the principal

components analysis (emotionality, shyness/sociability, attention, and activity level) are sufficiently similar to those found in many other studies to suggest the beginnings of an answer.

# REFERENCES

Ainsworth, M.S., Blehar, M.C., Waters, E., & Will, S. (1978). Patterns of Attachmant: A Psychological Study of the Strange Situation. Hillsdale, N.J.: Laurence Erlbaum Assoc.

Bates, J.E. (1980). The concept of difficult temperament. Merrill-Palmer Quarterly, 26, 299-319.

Bates, J.E., Freeland, C.A.B., & Lounsbury, M.L. (1979). Measurement of infant difficultness. Child Development, 50, 794-803.

Bell, R.Q., Weller, G.M., & Waldrop, M.F. (1971). Newborn and preschooler: Organization of behavior and relations between periods. Monographs of the Society for Research in Child Development. 36, No. 142, (1-2)

Billman, J., & McDivitt, S.C. (1980). Convergence of parent and observer ratings of temperament with observations of peer interaction in nursery school. Child Development, 51, 395-400.

Birns, B., Barten, S., & Bridges, W.H. (1969). Individual differences in temperamental characteristics of infants. Transactions of the New York Academy of Sciences, Series 2, No. 8, 1071-1082.

Bowlby, J. (1969). Attachment and Loss. Volume 1: Attachment. New York: Basic Books.

Brazelton, T.B. (1973). Neonatal Behavior Assessment Scale.

Little Club Clinics in Developmental Medicine, No. 50. London: William Heinemann Medical Books; Philadelphia: Lippincott.

Brazelton, T.B., & Freedman, D.G. (1971). Manual to accompany Newborn Behavioral Neurological Scales. In G.B.A. Stoelinger & J.J. vander Weifften (Eds.), Normal and abnormal development of brain and behavior. Leiden: Leiden University Press.

Brodsky, P., & Brodsky, M. (1978). An arousal interval scale: A psychophysical scale for GSR analysis. Perceptual and Motor Skills, 47, 747-756.

Buss, A. H., & Plomin, R. (1975). A temperament theory of personality development. Toronto: John Wiley & Sons.

Buss, A. H., & Plomin, R. (1984). Temperament: Early-developing personality traits. Hillsdale, New Jersey: L. Erlbaum Associates.

Caldwell, B., & Bradley, R. (1983). Home Observation for Measurement of the Environment. New York: Dorsey.

Cameron, J.R. (1977). Parental treatment, children´s temperament, and the risk of childhood behavior problems: 1. Relationships between parental characteristics and changes in children´s temperament over time. American Journal of Orthopsychiatry, 47, 568-576.

Cameron, J.R. (1977). Parental treatment, children´s temperament, and the risk of childhood behavior problems: 2.

Initial temperament, parental attitudes, and the incidence and form of behavior problems. American Journal of Orthopsychiatry, 48, 140-147.

Campbell, S.B.G. (1979). Mother-infant interaction as a function of maternal ratings of temperament. Child Psychiatry and Human Development, 10, 67-76.

Carey, W.B. (1970). A simplified method for measuring infant temperament. Journal of Pediatrics, 77, 188-194.

Carey, W.B. (1972). Measuring infant temperament. Journal of Pediatrics, 81, 414. (a)

Carey, W.B. (1972). Clinical applications of infant temperament measurements. Journal of Pediatrics, 81, 823-828. (b)

Carey, W. B. (1973). Measurement of infant temperament in pediatric practice. In J. C. Westman (Ed.), Individual differences in children. New York: Wiley.

Carey, W. B. (1981). The importance of temperament-environment interaction for child health and development. In M. Lewis & L. A. Rosenblum (Eds.), The Uncommon Child. New York: Plenum.

Carey, W. C., & McDevitt, S. C. (1978). Revision of the infant temperament questionnaire. Pediatrics, 61, 735-739.

Cattell, R. B. (1978). The scientific use of factor analysis in behavioral and life sciences. New York: Plenum Press.

Chamberlin, R.W. (1977). Can we identify a group of children at age 2 who are at high risk for the development of behavior or

emotional problems in kindergarten and first grade? Pediatrics, 59, 971-981.

Corsini, D. A., & Doyle, K. (1979). Temperament traits of preschool children: Across setting consistency. Paper presented at the meetings of the Society for Research on Child Development.

Dempsey, J.R. (1977). The measurement of individual differences in newborns. Unpublished Master´s thesis, University of Kansas.

Diamond, S. (1957). Personality and Temperament. New York: Harper.

Duffy, E. Activation. (1972). In N. S. Greenfield & R. A. Steinbach (Eds.), Handbook of psychophysiology. New York: Holt, Rinehart and Winston.

Dunn, J.F. (1977). Patterns of early interaction: Continuities and consequences. In H.R. Schaffer (Ed.), Studies in Parent-Infant Interaction, Academic Press: London.

Dunn, J. (1979). The first year of life: Continuities in individual differences. In D. Shaffer & J. Dunn (Eds.), The First Year of Life: Psychological and Medical Implications of Early Experience, Toronto: Wiley.

Dunn, J.F., & Richards, M.P.M. (1977). Observations on the developing relationship between mother and baby in the newborn. In H.R. Schaffer (Ed.), Studies in Parent-Infant Interaction, Academic Press: London.

Eaton, W. O. (1983). Measuring activity level with actometers: Reliablility, validity, and arm length. Child Development, 54, 720-726.

Eaton, W.O., & Enns L.R. (1986). Sex differences in human motor activity level. Psychological Bulletin, 100, 19-28.

Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. Journal of Personality and Social Psychology, 37, 1097-1126.

Epstein, S. (1980). The stability of behavior: II. Implications for psychological research. American Psychologist, 35, 790-806.

Escalona, S.K. (1968). The Roots of Individuality. Chicago: Adline.

Escalona, S.K. (1973). The differential impact of environmental conditions as a funtion of different reaction patterns in infancy. In J.C. Westman (Ed.), Individual Differences in Children, Toronto: John Wiley & Sons.

Escalona, S.K., & Heider, G.M. (1959). Predictions and outcome: A study of child development. New York: Basic Books.

Feiring, C., & Lewis, M. (1980). Temperament-Sex differences and stability in vigor, acitivity, and persistence in the first 3 years of life. The Journal of Genetic Psychology, 1980, 36, 65-75.

Fullard, W., McDevitt, S.C., & Carey, W.B. (1978). Toddler temperament scale. Unpublished manuscript. Temple U.,

Philadelphia, PA.

Garcia-Coll, C., Kagan, J., & Reznick, J.S. (1984). Behavioral inhibition in young children. Child Development, 55, 1005-1019.

Garside, R.F., Birch, H., Scott, D. McI., Chambers, S., Kolvin, I., Tweddler, E.G., & Barber, L.M. (1975). Dimensions of temperament in infant school children. Journal of Child Psychology and Psychiatry, 16, 219-231.

Goldsmith, H.H., & East, P.L. (1981, April). Parental perception of infant temperament: Validity and genetics. Paper presented at society for Research in Child Development, Boston.

Goldsmith, H.H., & Gottesman, I.I. (1981). Origins of variation in behavioral style: A longitudinal study of temperament in young twins. Child Development, 52, 91-103.

Graham, F. K. (1973). Habituation and dishabituation of responses innervated by the autonomic nervous system. In H. V. S. Peeke & M. J. Herz (Eds.), Habituation. Volume 1. New York: Academic Press.

Graham, F. K., & Clifton, R. K. (1966). Heart rate change as a component of the orienting response. Psychological Bulletin, 65, 305-320.

Graham, P., Rutter, M., & George, S. (1973). Temperamental characteristics as predictors of behavior disorders of children. American Journal of Orthopsychiatry, 43, 328-339.

Halverson, C.F., & Martin, C.L. (1981, April). Parent-child stabilitiy over time. Paper presented at the meeting of Society for Research in Child Development, Boston.

Hegvik, R.L., McDevitt, S.C., & Carey, W.B. (1981, April). The assessment of temperamental characteristics in 8-12 year old children. Paper presented at the meetings of Society for Research in Child Development, Boston.

Horn, J., Plomin, R., & Rosenman, R. (1976). Heritability of personality traits in adult male twins. Behavior Genetics, 6, 17-30.

Kagan, J. (1971). Change and continuity in infancy. New York: Wiley, 1971.

Klein, D. B. A history of scientific psychology: Its origins and philosophical backgrounds. New York: Basic Books.

Korner, A.F. (1971). Individual differences at birth: Implications for early experience and later development. American Journal of Orthopsychiatry, 41, 608-619.

Krakow, J.B., Kopp, C.B., & Vaughn, B.E. (1981, April). Sustained attention during the second year: Age trends, individual differences, and implications for development. Paper presented at the meeting of Society for Research in Child Development, Boston.

Kretschmer, E. (1925). Physique and character. New York: Harcourt Brace.

Kronstadt, D., Oberklaid, F., Ferb, T.E., & Swartz, J.P. (1979).

Infant behavior and maternal adaptations in the first six months of life. American Journal of Orthopsychiatry, 49, 454-464.

Lacey, J. I. (1956). The evaluation of autonomic responses: Toward a general solution. Annals of the New York Academy of Sciences, 67, 123-163.

Lacey, J. I. (1959). Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome. In E. A. Rubenstein & M. B. Parloff (Eds.), Research in psychotherapy. Washington, D.C.: APA.

Lacey, J. I., Bateman, D. E., & Van Lehn, R. (1953). Autonomic response specificity: An experimental study. Psychosomatic Medicine, 15, 8-21.

Lacey, J. I., & Lacey, B. C. (1958). The relationship of resting autonomic activity to motor impulsivity. Research Publications of the Association for Nervous and Mental Disease, 36, 144-209.(a)

Lacey, J. I., & Lacey, B. C. (1958). Verification and extension of the principles of autonomic response-stereotypy. American Journal of Psychology, 71, 50-73.(b)

Lee, C.L. (1981, April). Perceived difficult temperament and mother-toddler interaction sequences. Paper presented at Society for Research in Child Development, Boston.

Lewis, M. (1970). Individual differences in the measurement of early cognitive growth. In J. Hellmuth (Ed.), The

Exceptional Infant, Vol. 2. Washington: Brumner-Mazel.

Lewis, M., & Baldini, N. (1979). Attentional processes and individual differences. In G. A. Hale & M. Lewis (Eds.), Attention and cognitive development. New York: Plenum.

Lipton, E. L., Steinschneider, A., & Richmond, J. B. (1960). Autonomic function in the neonate: II. Physiologic effects of motor restraint. Psychosomatic Medicine, 22, 57-65.

Lipton, E. L., Steinschneider, A., & Richmond, J. B. (1961). Autonomic function in the neonate: IV. Individual differences in cardiac reactivity. Psychosomatic Medicine, 23, 472-484.

Lipton, E. L., Steinschneider, A., & Richmond, J. B. (1966). Autonomic function in the neonate: VII. Maturational changes in cardiac control. Child Development, 37, 1-16.

Loehlin, J.C., & Nichols, R.C. (1976). Heredity, Environment, and Personality: A Study of 850 Twins. Austin: University of Texas Press.

Lynn, R. (1966). Attention, arousal and the orienting reaction. London: Pergamon Press.

Malatesta, C.Z. & Haviland, J.M. (1981, April). Emotional socialization in the infant: Age and sex differences and the influence of maternal emotional traits. Paper presented at society for Research in Child Development, Boston.

Martin, P.P., & Wachs, T.D. (1981, April). A longitudinal study of temperament and its correlates in the first year of life.

Paper presented at society for Research in Child Development, Boston.

Matheny, A.P., & Dolan, A.B. (1980). A twin study of personality and temperament during middle childhood. Journal of Research in Personality, 14, 224-234.

Maurer, R., Cadoret, R.J., & Cain, C. (1980). Cluster analysis of childhood temperament data on adoptees. American Journal of Orthopsychiatry, 50, 522-534.

McDevitt, S. C., & Carey, W. B. (1978). The measurement of temperament in 3-7 year old children. Journal of Child Psychology & Psychiatry, 19, 245-253.

McDevitt, S. C., & Carey, W. B. (1981). Stability of ratings vs. perceptions of temperament from early infancy to 1 to 3 years. Paper presented at society for Research in Child Development, Boston.

McGehee, L.J., Eckerman, C.O., & Gross, S.J. (1981, April). The predicitive value of postnatal behavior observations of the high-risk premature infant. Paper presented at the meeting of Society for Research in Child Development, Boston.

McInery, T., & Chamberlin, R. (1978). Is it feasible to identify infants who are at risk for later behavioral problems? Clinical Pediatrics, 17, 233-238.

Milner, P. (1970). Physiological psychology. New York: Holt, Rinehart & Winston.

Morgan, L.J., & Camp, B.W. (1981, April). Stability of infant

vocalizations and maternal verbal responses. Paper presented at the meeting of Society for Research in Child Development, Boston.

Olweus, D. (1980). Familial and temperamental determinants of aggressive behavior in adolescent boys: A causal analysis. Developmental Psychology, 16, 644-660.

Pain, K. (1979). The effects of temperament on mother-infant interactions. Paper presented at the annual conference of the Canadian Psychological Association, Ottawa, Canada, 1979.

Pain, K.S. (1980). Infant temperament and assessment of cognitive development. Paper presented at the annual conference of the Canadian Psychological Association, Ottawa, Canada, 1980.

Pederson, F.A., Zaslow, M., Cain, R.L., Anderson, B.J., & Thomas, M. (1980). A methodology for assessing parent perception of baby temperament. J.S.A.S. Catalogue of Selected Documents in Psychology, 10, (10), Ms. No. 1978.

Persson-Blennow, I., & McNeil, T.F. (1979). A questionnaire for measurement of temperament in six-month-old infants: Development and standardization. Journal of Child Psychology and Psychiatry, 20, 1-13.

Plomin, R. (1976). A twin and family study of personality in young children. Journal of Psychology, 94, 233-235.

Plomin, R., & Rowe, D.C. (1979). Genetic and Environmental etiology of social behavior in infancy. Developmental

Psychology, 16, 62-72.

Powell, L.F. (1974). The effect of extra stimulation and maternal involvement on the development of low-birth-weight infants and on maternal behavior. Child Development, 45, 106-113.

Roe, K.V. (1978). Infants´ mother-stranger discrimination at 3 months as a predictor of cognitive development at 3 and 5 years. Developmental Psychology, 14, 191-192.

Roe, K.V., McClure, A., & Roe, A. (1980, September). Infant´s mother-stranger discrimination and cognitive functioning twelve years later. Paper presented at the annual meeting of APA, Montreal.

Rothbart, M.K. (1981). Measurement of temperament in infancy. Child Development, 52, 569-578.

Rothbart, M.K. (1986). Longitudinal observation of infant temperament. Developmental Psychology, 22, 356-365.

Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. E. Lamb & A. L. Brown (Eds.). Advances in developmental psychology (Vol. 1). Hillsdale, New Jersey: Erlbaum, Inc.

Rothbart, M. K., & Derryberry, D. (1982). Theoretical issues in temperament. In M. Lewis & L. Taft (Ed.), Developmental disabilities: Theory, assessment and intervention. New York: S. P. Medical and Scientific Books.

Rowe, D. C., & Plomin, R. (1977). Temperament in early

childhood. Journal of Personality Assessment, 41, 150-156.

Rutter, M, Korn, S., & Birch, H.G. (1963). Genetic and environmental factors in the development of "primary reactive patterns". British Journal of the Society for Clinical Psychology, 2, 161.

Sameroff, A.J. (1975). Early influences on development: Fact or fancy? Merrill-Palmer Quarterly, 21, 267-294.

Sameroff, A.J. (1980). Issues in early reproductive and caretaking risk: Review and current status. In D.B. Sawin, R.B. Hawkins, L.O. Walker, & J.H. Penticuff (Eds.), Exceptional Infant IV: Psychosocial Risks in Infant-Environment Transactions. New York: Brunner/Mazel.

Sameroff, A.J., & Sandler, M.J. (1975). Reproductive risk and the continuum of caretaking casualty. In F.D. Horowitz (Ed.), Review of Child Development Research. Chicago: U. of Chicago Press.

Sameroff, A., Seifer, R., & Elias, P.K. (1982). Socio-cultural variablility in infant temperament rating. Child Development, 53, 164-173.

Sarett, P.T. (1976). The relationship between infant temperament and selected maternal behaviors. Unpublished doctoral dissertation, State University of New Jersey at Rutgers.

Scarr, S. (1969). Social introversion-extraversion as a heritable response. Child Development, 40, 823-832.

Scholom, A., & Koller, T.J. (1980). Relating infant temperament to mother-infant and father-infant interaction. Paper presented at the Annual Convention of the American Psychological Association, Montreal.

Scholom, A., Zucker, R.A., Stollak, G.E. (1979). Relating early child adjustment to infant and parent temperament. Journal of Abnormal Child Psychology, 7, 297-308.

Shane, J., Brodsky, M., & Brodsky, P. (1982). The relationship between temperament, arousal and discrimination performance. Paper presented at the annual conference of the American Psychological Association, Washington, D.C.

Sheldon, W. H., & Stevens, S. S. (1946). The varieties of temperament. New York: Harper & Row.

Shirley, M.M. (1931). The First Two Years: Volume I: Postural and Locomotor Development. Minneapolis: U. of Minnesota Press.

Shirley, M.M. (1933). The First Two Years: Volume II: Intellectual Development. Minneapolis: U. of Minnesota Press. (a)

Shirley, M.M. (1933). The First Two Years: Volume III: Personality Manifestations. Minneapolis: U. of Minnesota Press. (b)

Sobesky, W.E., List, K.R., Holden, D.L., & Braucht, G.N. (1981, April). Dimensions of child temperament in school settings. Paper presented at society for Research in Child Development,

Boston.

Sokolov, E. N. (1963). Perception and the conditioned reflex. New York: Macmillan.

Sokolov, E. N. (1969). The modeling properties of the nervous system. In M. Cole & Maltzman (Eds.), A handbook of contemporary solvent psychology. New York: Basic Books.

Sostek, A., & Anders, T.F. (1977). Relationship among the Brazelton Neonatal Scale, Bayley Infant Scales, and early temperament. Child Development, 48, 320- 323.

Stevenson, M.B., & Lamb, M.E. (1979). Effects of infant sociability and the caretaking environment on infant cognitive development. Child Development, 50, 340-349.

Tarchanoff, J. (1976). Galvanic phenomena in the human skin during stimulation of the sensory organs and during various forms of mental activity. In S.W. Porgea and M.G.H. Coles (Eds.), Psychophysiology, Stroudsburg, PA: Dowden Hutchingon & Ross, Inc.

Thomas, A., & Chess, S. (1977). Temperament and development. New York: Bruner/Mazel.

Thomas, A., Chess, S., & Birch, H. G. (1968). Temperament and behavior disorders in children. New York: New York University Press.

Thomas, A., Chess, S., Birch, H. G., Hertzig, M., & Korn, S. (1963). Behavioral individuality in early childhood. New York: New York University Press.

Torgersen, A.M., & Kringlen, E. (1978). Genetic aspects of temperamental differences in infants. Journal of the American Academy of Child Psychiatry, 17, 433-444.

Vandenberg, S. G., Clark, P. J., & Samuels, I. (1965). Psychophysiological reactions of twins: Hereditary factors in galvanic skin response, heartbeat, and breathing rates. Eugenico Quarterly, 12, 7-10.

Von Bargen, D. M. (1983). Infant heart rate: A review of research and methodology. Merrill-Palmer Quarterly, 29, 115-149.

Wenger, M.A. (1941). The measurement of individual differences in autonomic balance. Psychosomatic Medicine, 3, 427-434.

Westman, J. C., & Baldwin, H. H. (1973). Foundations for the study of individual children. In J. C. Westman (Ed.), Individual differences in children. Toronto: Wiley.

Wilhoit, P.D. (1976). Assessment of temperament during the first months of life. Unpublished doctoral dissertations, Florida State University.

Woodworth, R., & Schlosberg, H. (1954). Experimental Psychology. New York: Holt.

Table 1. Correlation Matrix

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | GENDER | | | | | | | | | |
| 2 | AGE (MO) | .19 | | | | | | | | |
| 3 | ORDER | -.02 | -.10 | | | | | | | |
| 4 | CLASS | -.05 | .26 | -.32 | | | | | | |
| 5 | CCTI SOC | -.12 | .04 | .01 | -.22 | | | | | |
| 6 | CCTI EMOT | -.26 | -.20 | .01 | .26 | -.08 | | | | |
| 7 | CCTI ACT | -.13 | -.28 | .12 | .01 | -.01 | -.03 | | | |
| 8 | CCTI ATT | -.07 | .15 | .24 | -.08 | .11 | -.32 | -.05 | | |
| 9 | CCTI FOOD | -.26 | -.03 | .05 | .29 | -.02 | .34 | -.15 | -.09 | |
| 10 | CCTI SOOTH | -.20 | .30 | .18 | -.30 | .21 | -.53 | -.17 | .34 | -.18 |
| 11 | EAS SHY | .18 | -.01 | -.03 | .23 | -.88 | .13 | -.04 | .04 | .07 |
| 12 | EAS EMOT | -.20 | -.29 | -.11 | .26 | -.15 | .84 | .10 | -.33 | .27 |
| 13 | EAS SOC | -.23 | -.06 | -.09 | -.12 | .55 | -.09 | .09 | -.06 | .24 |
| 14 | EAS ACT | -.09 | -.27 | -.01 | .05 | .04 | -.05 | .75 | -.14 | -.11 |
| 15 | BSQ ACT | -.08 | -.19 | -.15 | .10 | .14 | .27 | .50 | -.47 | .08 |
| 16 | BSQ RHYTH | -.06 | -.01 | -.31 | .44 | -.18 | .29 | .17 | -.18 | .13 |
| 17 | BSQ APPR | -.08 | -.13 | -.14 | .33 | -.65 | .31 | -.03 | -.15 | .52 |
| 18 | BSQ ADAPT | -.03 | -.12 | -.31 | .39 | -.15 | .46 | .11 | -.39 | .23 |
| 19 | BSQ INT | -.12 | -.16 | .01 | .12 | .06 | .60 | .08 | -.09 | .20 |
| 20 | BSQ MOOD | -.20 | -.05 | -.32 | .28 | -.15 | .52 | .13 | -.37 | .08 |
| 21 | BSQ PERSIS | .05 | -.23 | -.19 | .02 | .05 | .03 | .14 | -.67 | .06 |
| 22 | BSQ DISTR | -.06 | -.20 | .08 | .01 | .01 | .06 | .22 | .08 | -.08 |
| 23 | BSQ THRESH | -.06 | .-04 | .08 | -.03 | .10 | .29 | .04 | .01 | .05 |
| 24 | DIAG CLUS | .02 | -.14 | -.25 | .39 | -.26 | .53 | .08 | -.35 | .20 |
| 25 | EMOT FACT | -.26 | -.20 | -.08 | .28 | -.05 | .87 | .07 | -.20 | .27 |
| 26 | SHY FACT | .14 | -.05 | -.03 | .24 | -.93 | .09 | .07 | .01 | .00 |
| 27 | DISTR FACT | .01 | -.10 | -.33 | .21 | -.03 | .08 | .08 | -.75 | .33 |
| 28 | ACT FACT | .01 | -.25 | .02 | -.04 | -.08 | -.14 | .85 | -.08 | -.43 |
| 29 | THRSH FACT | -.07 | -.09 | .01 | .03 | -.07 | .18 | .13 | .15 | .12 |
| 30 | GFIRST AV | .16 | -.04 | -.02 | .31 | .11 | .24 | -.11 | -.08 | -.02 |
| 31 | HFIRST AV | .16 | .13 | -.06 | -.17 | -.11 | -.33 | -.05 | -.09 | -.16 |
| 32 | GFIRST DIF | -.08 | -.16 | .05 | -.23 | -.09 | .09 | -.27 | -.22 | .23 |
| 33 | HFIRST DIF | -.04 | -.26 | -.04 | -.12 | .00 | .00 | .21 | .14 | .12 |
| 34 | GFIRST MAX | -.18 | .01 | -.07 | .36 | .12 | .19 | -.07 | .07 | .03 |
| 35 | HFIRST MAX | .05 | .07 | .08 | -.08 | -.23 | -.24 | -.10 | .10 | -.08 |
| 36 | GSCND AV | .09 | .03 | .01 | .18 | .04 | -.26 | -.03 | .30 | .09 |
| 37 | HSCND AV | -.04 | .41 | .16 | -.05 | -.13 | -.14 | -.14 | .17 | -.03 |
| 38 | GSCND DIF | .17 | .11 | -.12 | -.10 | -.06 | -.09 | -.21 | -.17 | -.19 |
| 39 | HSCND DIF | .07 | .39 | -.29 | .25 | -.12 | -.06 | -.20 | .16 | .00 |
| 40 | GSCND MAX | -.13 | -.09 | .12 | .24 | .11 | .22 | .00 | .19 | .27 |
| 41 | HSCND MAX | -.08 | .31 | .21 | .02 | -.16 | .03 | -.21 | .01 | .04 |
| 42 | GTHRD AV | .22 | .07 | -.22 | .30 | -.03 | -.01 | -.10 | .15 | .05 |
| 43 | HTHRD AV | .35 | .11 | -.12 | -.02 | -.01 | -.12 | .20 | -.17 | -.12 |
| 44 | GTHRD DIF | .19 | -.28 | -.14 | -.31 | -.14 | -.17 | .09 | -.27 | -.25 |
| 45 | HTHRD DIF | -.23 | -.26 | -.09 | .13 | -.27 | .06 | .03 | .02 | .09 |
| 46 | GTHRD MAX | .03 | .16 | -.25 | .40 | .02 | .21 | -.04 | -.04 | .17 |
| 47 | HTHRD MAX | .16 | .07 | -.12 | -.10 | .04 | -.05 | -.02 | -.23 | -.11 |
| 48 | GLAST AV | .35 | .12 | -.29 | .32 | -.07 | -.13 | -.06 | .00 | -.07 |
| 49 | HLAST AV | .22 | .14 | -.12 | .11 | -.09 | -.23 | -.20 | .13 | -.11 |
| 50 | GLAST DIF | -.03 | -.17 | -.07 | .09 | -.21 | .20 | .14 | -.28 | -.06 |
| 51 | HLAST DIF | -.08 | .01 | .14 | -.06 | -.02 | -.17 | -.20 | -.02 | -.02 |
| 52 | GLAST MAX | .25 | .10 | .32 | .42 | .01 | .07 | -.10 | -.03 | .14 |
| 53 | HLAST MAX | .16 | .10 | -.03 | .14 | -.06 | -.17 | -.08 | .20 | -.05 |

Table 1 continued

| | | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|
| 10 | CCTI SOOTH | | | | | | | | | |
| 11 | EAS SHY | -.17 | | | | | | | | |
| 12 | EAS EMOT | -.70 | .21 | | | | | | | |
| 13 | EAS SOC | -.09 | -.61 | .03 | | | | | | |
| 14 | EAS ACT | -.14 | -.04 | .08 | .12 | | | | | |
| 15 | BSQ ACT | -.27 | -.10 | .26 | .10 | .65 | | | | |
| 16 | BSQ RHYTH | -.40 | .14 | .30 | -.20 | .11 | .34 | | | |
| 17 | BSQ APPR | -.45 | .72 | .45 | -.22 | -.07 | -.06 | .24 | | |
| 18 | BSQ ADAPT | -.58 | .31 | .55 | -.15 | .21 | .46 | .48 | .50 | |
| 19 | BSQ INTEN | -.40 | .03 | .57 | .07 | .22 | .34 | .14 | .15 | .34 |
| 20 | BSQ MOOD | -.57 | .14 | .56 | -.01 | .09 | .34 | .44 | .36 | .68 |
| 21 | BSQ PERS | -.28 | -.08 | .12 | .09 | .31 | .58 | .31 | .08 | .41 |
| 22 | BSQ DISTR | .06 | .01 | -.01 | .08 | .12 | .20 | .14 | -.11 | -.16 |
| 23 | BSQ THRSH | -.04 | -.08 | .21 | .23 | -.01 | .11 | .03 | -.06 | -.16 |
| 24 | DIAG CLUS | -.64 | .29 | .64 | -.04 | .13 | .33 | .56 | .52 | .73 |
| 25 | EMOT FACT | -.72 | .13 | .89 | .00 | .07 | .27 | .29 | .38 | .60 |
| 26 | SHY FACT | -.19 | .96 | .18 | -.66 | .00 | -.10 | .23 | .72 | .27 |
| 27 | DIST FACT | -.28 | -.04 | .12 | .15 | .21 | .58 | .50 | .25 | .52 |
| 28 | ACT FACT | -.11 | .07 | .00 | -.10 | .87 | .61 | .14 | -.15 | .17 |
| 29 | THRSH FACT | .08 | .09 | .09 | .17 | .05 | .11 | .11 | .04 | -.22 |
| 30 | GFRST AV | -.11 | -.03 | .14 | -.10 | -.08 | -.10 | .00 | .09 | .18 |
| 31 | HFRST AV | .06 | .08 | -.34 | -.05 | -.08 | -.03 | -.08 | .00 | -.13 |
| 32 | GFRST DIF | -.10 | .10 | .12 | .16 | -.28 | -.01 | .01 | .20 | .06 |
| 33 | HFRST DIF | .04 | .04 | .07 | .02 | .06 | -.11 | -.10 | .07 | .01 |
| 34 | GFRST MAX | -.09 | -.11 | .07 | .03 | .04 | -.13 | -.01 | .05 | .09 |
| 35 | HFRST MAX | -.03 | .19 | -.24 | -.12 | -.09 | -.19 | -.03 | .13 | -.10 |
| 36 | GSCND AV | .11 | -.08 | -.30 | .08 | -.05 | -.32 | -.05 | .01 | -.28 |
| 37 | HSCND AV | .14 | .11 | -.18 | .02 | -.19 | -.31 | -.05 | -.04 | -.15 |
| 38 | GSCND DIF | -.05 | .03 | .05 | -.09 | -.17 | -.11 | -.03 | .06 | -.04 |
| 39 | HSCND DIF | .25 | .28 | -.09 | -.13 | -.15 | -.14 | .02 | .13 | .14 |
| 40 | GSCND MAX | -.01 | -.11 | -.09 | .23 | -.01 | -.19 | -.13 | .06 | .16 |
| 41 | HSCND MAX | .05 | .13 | -.15 | -.06 | -.26 | -.22 | -.06 | .01 | -.07 |
| 42 | GTHRD AV | -.05 | .05 | .03 | -.16 | -.04 | -.12 | .06 | .15 | .07 |
| 43 | HTHRD AV | .08 | .06 | -.06 | -.05 | .11 | .10 | .19 | .05 | .09 |
| 44 | GTHRD DIF | -.12 | .12 | -.03 | -.10 | .11 | .25 | .16 | -.03 | .00 |
| 45 | HTHRD DIF | -.25 | .26 | .10 | -.02 | -.09 | -.16 | .22 | .35 | .18 |
| 46 | GTHRD MAX | -.15 | -.05 | .13 | .01 | -.03 | .00 | .13 | .13 | .20 |
| 47 | HTHRD MAX | .04 | .02 | -.14 | -.14 | .00 | .04 | .27 | -.01 | .19 |
| 48 | GLAST AV | -.01 | .07 | .01 | .00 | .04 | -.03 | .03 | .18 | .17 |
| 49 | HLAST AV | .04 | .10 | -.22 | -.21 | -.23 | -.24 | .11 | .12 | .10 |
| 50 | GLAST DIF | -.18 | .27 | .20 | -.33 | .28 | .38 | .23 | .04 | .36 |
| 51 | HLAST DIF | -.02 | .06 | -.21 | -.11 | -.05 | .05 | -.22 | .00 | .02 |
| 52 | GLAST MAX | -.08 | .02 | .14 | .05 | .03 | .07 | .17 | .21 | .27 |
| 53 | HLAST MAX | -.05 | -.04 | -.11 | -.03 | -.17 | -.31 | .25 | .09 | .03 |

Table 1 continued

| | | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 |
|---|---|---|---|---|---|---|---|---|---|---|
| 19 | BSQ INTEN | | | | | | | | | |
| 20 | BSQ MOOD | .43 | | | | | | | | |
| 21 | BSQ PERSIS | .09 | .19 | | | | | | | |
| 22 | BSQ DISTR | .00 | -.06 | .00 | | | | | | |
| 23 | BSQ THRESH | .43 | .15 | -.13 | .48 | | | | | |
| 24 | DIAG CLUS | .52 | .76 | .27 | -.07 | .07 | | | | |
| 25 | EMOT FACT | .74 | .69 | -.01 | -.11 | .25 | .69 | | | |
| 26 | SHY FACT | -.04 | .18 | -.07 | .08 | -.07 | .29 | .09 | | |
| 27 | DIST FACT | -.04 | .31 | .86 | .01 | -.14 | .39 | .00 | -.01 | |
| 28 | ACT FACT | .09 | .11 | .27 | .22 | -.05 | .08 | -.02 | .15 | .10 |
| 29 | THRSH FACT | .27 | .02 | -.17 | .80 | .85 | .02 | .06 | .12 | -.12 |
| 30 | GFIRST AV | .25 | .07 | .02 | .02 | .16 | .15 | .23 | -.03 | -.06 |
| 31 | HFIRST AV | -.33 | -.33 | .20 | .07 | -.07 | -.26 | -.40 | .10 | .16 |
| 32 | GFIRST DIF | -.22 | .18 | .18 | -.14 | -.04 | .15 | .02 | .04 | .27 |
| 33 | HFIRST DIF | -.14 | -.15 | -.09 | .12 | -.09 | -.18 | -.03 | .05 | -.10 |
| 34 | GFIRST MAX | .25 | .00 | -.13 | .08 | .14 | .06 | .19 | -.08 | -.14 |
| 35 | HFIRST MAX | -.20 | -.26 | -.07 | .04 | -.07 | -.16 | -.24 | .22 | -.05 |
| 36 | GSCND AV | -.07 | -.24 | -.15 | .00 | -.11 | -.12 | -.23 | -.05 | -.15 |
| 37 | HSCND AV | -.16 | -.10 | -.20 | -.03 | .01 | -.13 | -.16 | .08 | -.15 |
| 38 | GSCND DIF | -.11 | -.07 | .05 | -.05 | .07 | -.02 | -.06 | .06 | .03 |
| 39 | HSCND DIF | -.11 | .04 | -.16 | .08 | .10 | .02 | -.09 | .22 | -.06 |
| 40 | GSCND MAX | .17 | -.11 | -.09 | -.04 | -.02 | -.06 | .02 | -.14 | -.11 |
| 41 | HSCND MAX | -.10 | -.02 | -.11 | -.12 | -.09 | -.09 | -.07 | .09 | -.06 |
| 42 | GTHRD AV | .15 | .07 | -.21 | -.06 | .06 | .16 | .11 | .09 | -.15 |
| 43 | HTHRD AV | .09 | .05 | .07 | -.14 | .19 | .13 | -.05 | .07 | .14 |
| 44 | GTHRD DIF | -.04 | .08 | .36 | .00 | -.12 | .07 | -.12 | .12 | .28 |
| 45 | HTHRD DIF | -.06 | .17 | -.04 | -.11 | -.10 | .19 | .11 | .28 | .06 |
| 46 | GTHRD MAX | .33 | .21 | -.16 | -.10 | .09 | .28 | .29 | -.03 | -.02 |
| 47 | HTHRD MAX | .11 | .08 | .18 | -.11 | .03 | .17 | -.05 | .02 | .24 |
| 48 | GLAST AV | .17 | .11 | -.03 | -.11 | .04 | .32 | .04 | .10 | .02 |
| 49 | HLAST AV | -.12 | -.03 | -.07 | -.21 | -.19 | .10 | -.14 | .13 | -.02 |
| 50 | GLAST DIF | .18 | .24 | .28 | .00 | -.06 | .26 | .20 | .24 | .24 |
| 51 | HLAST DIF | .07 | -.14 | .24 | .02 | -.23 | -.19 | -.14 | .02 | .08 |
| 52 | GLAST MAX | .30 | .13 | .00 | -.13 | .06 | .37 | .19 | .01 | .10 |
| 53 | HLAST MAX | -.05 | .04 | -.19 | -.26 | -.05 | .14 | -.03 | .03 | -.09 |

Table 1 continued

| | | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 |
|---|---|---|---|---|---|---|---|---|---|---|
| 28 | ACT FACT | | | | | | | | | |
| 29 | THRSH FACT | .04 | | | | | | | | |
| 30 | GFIRST AV | -.12 | .06 | | | | | | | |
| 31 | HFIRST AV | .02 | -.03 | -.03 | | | | | | |
| 32 | GFIRST DIF | -.30 | -.09 | -.30 | -.04 | | | | | |
| 33 | HFIRST DIF | .07 | .02 | -.17 | -.14 | -.07 | | | | |
| 34 | GFIRST MAX | -.08 | .11 | .85 | -.07 | -.53 | -.15 | | | |
| 35 | HFIRST MAX | -.05 | .00 | .06 | .70 | -.21 | -.11 | .12 | | |
| 36 | GSCND AV | -.12 | .00 | .41 | -.05 | -.09 | -.03 | .46 | .00 | |
| 37 | HSCND AV | -.17 | .04 | -.21 | .41 | .02 | .03 | -.16 | .50 | -.11 |
| 38 | GSCND DIF | -.13 | -.03 | .10 | .14 | -.09 | -.09 | .02 | .09 | -.25 |
| 39 | HSCND DIF | -.13 | .15 | .00 | .07 | .02 | .06 | .03 | .10 | .00 |
| 40 | GSCND MAX | -.14 | .03 | .40 | -.21 | -.12 | -.01 | .52 | -.04 | .82 |
| 41 | HSCND MAX | -.22 | -.09 | -.13 | .34 | .07 | .02 | -.12 | .45 | -.08 |
| 42 | GTHRD AV | -.09 | .03 | .57 | -.04 | -.26 | -.28 | .52 | -.08 | .50 |
| 43 | HTHRD AV | .15 | .06 | .28 | .13 | -.07 | -.14 | .13 | -.07 | .19 |
| 44 | GTHRD DIF | .24 | -.09 | -.33 | .17 | .28 | .04 | -.51 | .13 | -.18 |
| 45 | HTHRD DIF | -.07 | -.07 | -.02 | -.23 | .19 | .02 | .00 | -.16 | .01 |
| 46 | GTHRD MAX | -.12 | .03 | .50 | -.14 | -.40 | -.33 | .61 | -.18 | .30 |
| 47 | HTHRD MAX | .01 | -.05 | .31 | .23 | -.07 | -.22 | .19 | .12 | .23 |
| 48 | GLAST AV | .01 | .00 | .40 | -.14 | -.17 | -.26 | .44 | -.14 | .39 |
| 49 | HLAST AV | -.16 | -.21 | .31 | .36 | -.12 | -.07 | .29 | .33 | .25 |
| 50 | GLAST DIF | .31 | -.07 | -.01 | -.02 | -.02 | .02 | -.10 | -.11 | -.28 |
| 51 | HLAST DIF | -.02 | -.16 | -.03 | .09 | .02 | -.11 | .02 | .00 | .11 |
| 52 | GLAST MAX | -.08 | .02 | .40 | -.22 | -.29 | -.29 | .51 | .21 | .21 |
| 53 | HLAST MAX | -.16 | -.13 | .33 | .22 | -.15 | -.01 | .34 | .25 | .39 |

Table 1 continued

| | | | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 37 | HSCND | AV | | | | | | | | | |
| 38 | GSCND | DIF | .02 | | | | | | | | |
| 39 | HSCND | DIF | .33 | -.01 | | | | | | | |
| 40 | GSCND | MAX | -.17 | -.44 | -.09 | | | | | | |
| 41 | HSCND | MAX | .86 | -.10 | .11 | .00 | | | | | |
| 42 | GTHRD | AV | -.37 | .07 | .06 | .29 | -.42 | | | | |
| 43 | HTHRD | AV | .04 | .20 | .11 | .01 | .02 | .28 | | | |
| 44 | GTHRD | DIF | -.04 | .13 | -.24 | -.20 | .01 | -.29 | .12 | | |
| 45 | HTHRD | DIF | -.09 | .16 | -.10 | .07 | .04 | -.16 | .15 | .04 | |
| 46 | GTHRD | MAX | -.34 | -.02 | -.06 | .32 | -.28 | .74 | .21 | -.46 | .07 |
| 47 | HTHRD | MAX | .19 | .05 | .13 | .13 | .24 | .21 | .71 | .08 | .14 |
| 48 | GLAST | AV | -.31 | .08 | .17 | .21 | -.37 | .62 | .26 | -.23 | -.02 |
| 49 | HLAST | AV | .22 | .10 | .25 | -.01 | .19 | .31 | .31 | -.12 | -.05 |
| 50 | GLAST | DIF | -.07 | -.02 | -.13 | -.20 | .04 | .06 | . 15 | .24 | .03 |
| 51 | HLAST | DIF | -.13 | -.03 | -.04 | .17 | .01 | -.04 | -.02 | .09 | -.02 |
| 52 | GLAST | MAX | -.39 | .16 | .05 | .16 | -.37 | .56 | .25 | -.32 | .16 |
| 53 | HLAST | MAX | .17 | -.05 | .02 | .18 | .15 | .39 | .40 | -.17 | .08 |

| | | | 46 | 47 | 48 | 49 | 50 | 51 | 52 |
|---|---|---|---|---|---|---|---|---|---|
| 46 | GTHRD | MAX | | | | | | | |
| 47 | HTHRD | MAX | .23 | | | | | | |
| 48 | GLAST | AV | .58 | .13 | | | | | |
| 49 | HLAST | AV | .17 | .31 | .28 | | | | |
| 50 | GLAST | DIF | .11 | .16 | -.19 | -.06 | | | |
| 51 | HLAST | DIF | -.09 | .04 | -.01 | .09 | .11 | | |
| 52 | GLAST | MAX | .73 | .12 | .86 | .17 | -.11 | -.08 | |
| 53 | HLAST | MAX | .29 | .33 | .30 | .78 | -.11 | -.17 | .23 |

Table 2. Means and Standard Deviations

| | | MEAN | SD |
|---|---|---|---|
| 1 | GENDER | 1.53 | 0.50 |
| 2 | AGE (MO) | 60.31 | 9.19 |
| 3 | ORDER | 4.47 | 2.32 |
| 4 | CLASS | 5.18 | 2.47 |
| 5 | CCTI SOC | 3.53 | 0.85 |
| 6 | CCTI EMOT | 2.52 | 0.90 |
| 7 | CCTI ACT | 3.90 | 0.61 |
| 8 | CCTI ATT | 3.59 | 0.78 |
| 9 | CCTI FOOD | 2.51 | 0.90 |
| 10 | CCTI SOOTH | 3.10 | 0.74 |
| 11 | EAS SHY | 2.33 | 0.88 |
| 12 | EAS EMOT | 2.66 | 0.98 |
| 13 | EAS SOC | 3.71 | 0.74 |
| 14 | EAS ACT | 4.08 | 0.69 |
| 15 | BSQ ACT | 3.73 | 0.77 |
| 16 | BSQ RHYTH | 2.76 | 0.63 |
| 17 | BSQ APPR | 2.87 | 0.79 |
| 18 | BSQ ADAPT | 2.69 | 0.78 |
| 19 | BSQ INT | 4.30 | 0.66 |
| 20 | BSQ MOOD | 3.15 | 0.72 |
| 21 | BSQ PERSIS | 2.95 | 0.73 |
| 22 | BSQ DISTR | 3.84 | 0.60 |
| 23 | BSQ THRESH | 3.86 | 0.54 |
| 24 | DIAG CLUS | 2.06 | 1.29 |
| 25 | EMOT FACT | 0.00 | 1.00 |
| 26 | SHY FACT | 0.00 | 1.00 |
| 27 | DISTR FACT | 0.00 | 1.00 |
| 28 | ACT FACT | 0.00 | 1.00 |
| 29 | THRSH FACT | 0.00 | 1.00 |
| 30 | GFIRST AV | 53.43 | 38.00 |
| 31 | HFIRST AV | -8.68 | 3.86 |
| 32 | GFIRST DIF | -19.04 | 35.14 |
| 33 | HFIRST DIF | 1.04 | 5.76 |
| 34 | GFIRST MAX | 94.54 | 56.23 |
| 35 | HFIRST MAX | -18.24 | 6.61 |
| 36 | GSCND AV | 48.54 | 30.63 |
| 37 | HSCND AV | -9.45 | 4.60 |
| 38 | GSCND DIF | -20.26 | 38.92 |
| 39 | HSCND DIF | 0.72 | 5.70 |
| 40 | GSCND MAX | 92.21 | 56.54 |
| 41 | HSCND MAX | -19.48 | 7.75 |
| 42 | GTHRD AV | 40.89 | 26.45 |
| 43 | HTHRD AV | -8.47 | 4.16 |
| 44 | GTHRD DIF | -16.12 | 28.13 |
| 45 | HTHRD DIF | 0.60 | 6.00 |
| 46 | GTHRD MAX | 79.20 | 50.87 |
| 47 | HTHRD MAX | -20.05 | 8.39 |
| 48 | GLAST AV | 44.48 | 31.98 |
| 49 | HLAST AV | -9.06 | 4.39 |
| 50 | GLAST DIF | -6.29 | 31.25 |
| 51 | HLAST DIF | 0.34 | 6.91 |
| 52 | GLAST MAX | 84.21 | 58.17 |
| 53 | HLAST MAX | -20.44 | 9.12 |

Table 3

Rotated Factor Loadings from Principal Components Analysis of 18 Temperament Scores

| | | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Factor 5 |
|---|---|---|---|---|---|---|
| CCTI Sociability | 61 | -0.021 | -0.920 | -0.057 | 0.060 | 0.058 |
| CCTI Emotionality | 62 | 0.853 | 0.048 | 0.117 | -0.120 | 0.047 |
| CCTI Activity | 63 | 0.079 | -0.088 | 0.071 | 0.867 | 0.072 |
| CCTI Attention Span/ Persistence | 64 | -0.161 | -0.018 | -0.735 | 0.061 | 0.081 |
| CCTI Reaction-to-Food | 65 | 0.280 | -0.038 | 0.376 | -0.466 | 0.097 |
| CCTI Soothability | 66 | -0.641 | -0.106 | -0.150 | -0.046 | 0.195 |
| BSQ Activity Level | 67 | 0.317 | -0.177 | 0.506 | 0.618 | 0.098 |
| BSQ Rhythmicity | 68 | 0.229 | 0.070 | 0.623 | -0.016 | 0.164 |
| BSQ Approach/Withdrawal | 69 | 0.197 | 0.758 | 0.267 | -0.258 | 0.014 |
| BSQ Adaptability | 70 | 0.605 | 0.214 | 0.518 | 0.089 | -0.152 |
| BSQ Intensity | 71 | 0.714 | -0.180 | 0.047 | 0.085 | 0.310 |
| BSQ Quality of Mood | 72 | 0.705 | 0.167 | 0.359 | 0.089 | 0.020 |
| BSQ Attention Span/ Persistence | 73 | 0.077 | -0.026 | 0.799 | 0.280 | -0.110 |
| BSQ Distractibility | 74 | -0.097 | -0.083 | 0.014 | 0.158 | 0.802 |
| BSQ Threshold | 75 | 0.182 | -0.104 | -0.062 | -0.056 | 0.836 |
| EASI Shyness | 76 | 0.033 | 0.956 | -0.003 | -0.074 | -0.018 |
| EASI Emotionality | 77 | 0.862 | 0.087 | 0.107 | 0.019 | 0.068 |
| EASI Sociability | 78 | -0.070 | -0.614 | 0.200 | 0.050 | 0.321 |
| EASI Activity | 79 | 0.019 | -0.196 | 0.078 | 0.872 | 0.064 |
| | VP | 3.612 | 2.941 | 2.542 | 2.340 | 1.687 |

The VP for each factor is the sum of the squares of the elements of the column of the factor pattern matrix corresponding to that factor. When the rotation is orthogonal, the VP is the variance explained by the factor.

Table 4

Rearranged, Simplified Presentation of Factor Loadings[1]

| | | Emot. Factor 1 | Shyness Factor 2 | Att'n. Factor 3 | Activ. Factor 4 | Thresh. Factor 5 |
|---|---|---|---|---|---|---|
| EASI Emotionality | 77 | 0.862 | 0.0 | 0.0 | 0.0 | 0.0 |
| CCTI Emotionality | 62 | 0.853 | 0.0 | 0.0 | 0.0 | 0.0 |
| BSQ Intensity | 71 | 0.714 | 0.0 | 0.0 | 0.0 | 0.310 |
| BSQ Quality of Mood | 72 | 0.705 | 0.0 | 0.359 | 0.0 | 0.0 |
| CCTI Soothability | 66 | -0.641 | 0.0 | 0.0 | 0.0 | 0.0 |
| BSQ Adaptability | 70 | 0.605 | 0.0 | 0.518 | 0.0 | 0.0 |
| EASI Shyness | 76 | 0.0 | 0.956 | 0.0 | 0.0 | 0.0 |
| CCTI Sociability | 61 | 0.0 | -0.920 | 0.0 | 0.0 | 0.0 |
| BSQ Approach/Withdrawal | 69 | 0.0 | 0.758 | 0.267 | -0.258 | 0.0 |
| EASI Sociability | 78 | 0.0 | -0.614 | 0.0 | 0.0 | 0.321 |
| BSQ Attention Span/ Persistence | 73 | 0.0 | 0.0 | 0.799 | 0.280 | 0.0 |
| CCTI Attention Span/ Persistence | 64 | 0.0 | 0.0 | -0.735 | 0.0 | 0.0 |
| BSQ Activity Level | 67 | 0.317 | 0.0 | 0.506 | 0.618 | 0.0 |
| BSQ Rhythmicity | 68 | 0.0 | 0.0 | 0.623 | 0.0 | 0.0 |
| EASI Activity | 79 | 0.0 | 0.0 | 0.0 | 0.872 | 0.0 |
| CCTI Activity | 63 | 0.0 | 0.0 | 0.0 | 0.867 | 0.0 |
| BSQ Threshold | 75 | 0.0 | 0.0 | 0.0 | 0.0 | 0.836 |
| BSQ Distractibility | 74 | 0.0 | 0.0 | 0.0 | 0.0 | 0.802 |
| CCTI Reaction-to-Food | 65 | 0.280 | 0.0 | 0.376 | -0.466 | 0.0 |
| | | 26% | 18% | 11% | 7% | 7% |
| | VP | 3.612 | 2.941 | 2.542 | 2.340 | 1.687 |

[1]The above factor loading matrix has been rearranged so that the columns appear in decreasing order of variance explained by factors. The rows have been rearranged so that for each successive factor, loadings greater than 0.5000 appear first. Loadings less than 0.2500 have been replaced by zero.

Table 5

Canonical Correlation Results Between Physiological Reactivity Scores and Three Sets of Temperament Scores

| | Reactivity Scores | All 18 Temperament Scores | 5 Temperament Factors |
|---|---|---|---|
| GSR & HR MAX's (1,2,3,4) | ev* = .29 $X^2_{(64)} = 49.85$ $p = .90$ | ev = .47 $X^2_{(152)} = 149.48$ $p = .54$ | ev = .17 $X^2_{(40)} = 20.63$ $p = 1.00$ |
| GSR MAX's (1,2,3,4) | ev = .13 $X^2_{(32)} = 25.42$ $p = .79$ | ev = .28 $X^2_{(76)} = 62.58$ $p = .87$ | ev = .10 $X^2_{(20)} = 10.95$ $p = .95$ |
| HR MAX's (1,2,3,4) | ev = .17 $X^2_{(32)} = 22.37$ $p = .90$ | ev = .38 $X^2_{(76)} = 73.17$ $p = .57$ | ev = .12 $X^2_{(20)} = 13.59$ $p = .85$ |
| GSR & HR Firsts (1,2,3,4) | ev = .29 $X^2_{(64)} = 46.79$ $p = .95$ | ev = .56 $X^2_{(152)} = 140.04$ $p = .75$ | ev = .22 $X^2_{(40)} = 32.95$ $p = .78$ |
| GSR Firsts (1,2,3,4) | ev = .11 $X^2_{(32)} = 22.36$ $p = .90$ | ev = .28 $X^2_{(76)} = 53.60$ $p = .98$ | ev = .15 $X^2_{(20)} = 19.33$ $p = .50$ |
| HR Firsts (1,2,3,4) | ev = .17 $X^2_{(32)} = 23.08$ $p = .88$ | ev = .39 $X^2_{(76)} = 72.94$ $p = .58$ | ev = .13 $X^2_{(20)} = 15.24$ $p = .76$ |
| GSR & HR First MAX | ev = .15 $X^2_{(16)} = 21.31$ $p = .17$ | ev = .27 $X^2_{(38)} = 40.20$ $p < .37$ | ev = .12 $X^2_{(10)} = 13.71$ $p = .19$ |

*ev = eigenvalue (variance accounted for by the canonical variable)

Table 6

Canonical Correlation Results Between Physiological Difference Scores and Three Sets of Temperament Scores

| | Habituation Scores | All 18 Scores | 5 Factor Scores |
|---|---|---|---|
| GSR & HR Difference Scores (1,2,3,4) | ev* = .22<br>$X^2_{(40)} = 46.81$<br>$p = .21$ | ev = .43<br>$X^2_{(152)} = 156.49$<br>$p = .38$ | ev = .23<br>$X^2_{(40)} = 39.16$<br>$p = .51$ |
| GSR Difference Scores (1,2,3,4) | ev = .16<br>$X^2_{(20)} = 26.29$<br>$p = .16$ | ev = .33<br>$X^2_{(76)} = 74.70$<br>$p = .52$ | ev = .20<br>$X^2_{(20)} = 24.41$<br>$p = .22$ |
| HR Difference Scores (1,2,3,4) | ev = .10<br>$X^2_{(20)} = 19.51$<br>$p = .49$ | ev = .30<br>$X^2_{(76)} = 77.00$<br>$p = .45$ | ev = .08<br>$X^2_{(20)} = 14.03$<br>$p = .83$ |
| GSR & HR Difference Scores (Last only) | ev = .15<br>$X^2_{(10)} = 16.54$<br>$p = .09$ | ev = .34<br>$X^2_{(38)} = 52.35$<br>$p = .06$ | ev = .16<br>$X^2_{(10)} = 17.86$<br>$p = .06$ |

*ev = eigenvalue (variance accounted for by the canonical variate)

Table 7

Simultaneous Regressions of Three Sets of Temperament Scores (All 18 Subscale Scores, Habituation Scores, and 5 Factor Scores) on GSR and HR Difference Scores for the Last Stimulus

| | $R^2$ | F | df | p |
|---|---|---|---|---|
| GSR last Difference score | | | | |
| - with 19 scores | .31 | 1.56 | 19,65 | <.10 |
| - with habituation scores | .15 | 2.71 | 5,79 | <.03 |
| - with 5 factors | .16 | 2.94 | 5,79 | <.02 |
| HR last Differencescores | | | | |
| - with 19 scores | .29 | 1.43 | 19,65 | =.15 |
| - with habituation scores | .05 | .84 | 5,79 | =.52 |
| - with 5 factor scores | .06 | .93 | 5,79 | =.47 |

Table 8

Stepwise Regression of Habituation Scores on GSR Difference Scores for the Last Stimulus

| Variables In Order Entered | Coefficient Last Step | Standard Coefficient | r |
|---|---|---|---|
| BSQ Adaptability | 11.01 | .26 | .31 |
| CCTI Attention Span | -7.01 | -.18 | -.27 |
| BSQ Distractability | 6.09 | .12 | .06 |

$R = .37$

$R^2 = .14$

$F_{(3,81)} = 4.28,\ p < .01$

Table 9

Stepwise Multiple Regression of All 18 Temperament Scores With GSR Difference Scores for the Last Stimulus

| Variables In Order Entered | Coefficient Last Step | Standard Coefficient | r |
|---|---|---|---|
| BSQ Adaptability | ---- | ---- | .31 |
| BSQ Activity | 9.38 | .24 | .30 |
| EAS Shyness | 20.10 | .58 | .21 |
| BSQ Adaptability - Removed | ---- | ---- | |
| BSQ Rhythmicity | 9.25 | .18 | .26 |
| BSQ Approach | -17.89 | -.46 | .07 |
| CCTI Attention Span | -6.32 | -.16 | -.27 |
| CCTI Reaction-To-Food | 4.77 | .14 | .04 |

$R = .51$

$R^2 = .26$

$F_{(6,78)} = 4.62,\ p < .01$

Table 10

Stepwise Multiple Regression of the Five Temperament Factors With the GSR Difference Scores for the Last Stimulus

| Variables In Order Entered | Coefficient Last Step | Standard Coefficient | r |
|---|---|---|---|
| Distractability Factor | 8.30 | .28 | .27 |
| Shyness Factor | 5.80 | .19 | .19 |
| Activity Level Factor | 5.39 | .18 | .18 |
| Emotionality Factor | 3.45 | .11 | .11 |

Multiple R = .40

Multiple $R^2$ = .16

$F_{(4,80)} = 3.71, p < .01$

Table 11

Repeated Measures MANOVA for Diagnostic Cluster, Stimulus Presentation, and Trials Effects on GSR Scores

| Effect | Statistic | F | df | p |
|---|---|---|---|---|
| Diagnostic Cluster | SS = 85683.1 | 1.97 | 3, 78 | .125 |
| Error | SS = 1128235.2 | | | |
| Stimulus | $T^2$ = 25.71 | 8.35 | 3, 76 | .000 |
| S x Diagnostic Cluster | Lambda = .74 | 2.72 | 9, 185.11 | .005 |
| Error | SS = 592354.05 | | | |
| Trials | $T^2$ = 91.17 | 17.30 | 5, 74 | .000 |
| T x Diagnostic Cluster | Lambda = .69 | 1.95 | 15, 204.68 | .021 |
| Error | SS = 429570.3 | | | |
| Stimulus x Trials | $T^2$ = 30.72 | 1.68 | 15, 64 | .078 |
| ST x Diagnostic Cluster | Lambda = .52 | 1.06 | 45, 190.91 | .383 |
| Error | SS = 1208705.9 | | | |

Table 12

Correlations Between HR and GSR Scores

| | r | p | partial r[1] | p |
|---|---|---|---|---|
| **Maximum Scores** | | | | |
| GSR First Max with HR First Max[2] | -.065 | ns | -.064 | ns |
| GSR Second Max with HR Second Max[2] | -.312 | <.01 | -.318 | <.01 |
| GSR Third Max with HR Third Max[2] | .025 | ns | .025 | ns |
| GSR Last Max with HR Last Max[2] | .111 | ns | .075 | ns |
| **Difference Scores** | | | | |
| GSR First Dif with HR First Dif[3] | -.029 | ns | -.036 | ns |
| GSR Second Dif with HR Second Dif[3] | -.080 | ns | -.118 | ns |
| GSR Third Dif with HR Third Dif[3] | -.076 | ns | -.086 | ns |
| GSR Last Dif with Hr Last Dif[3] | .074 | ns | .079 | ns |

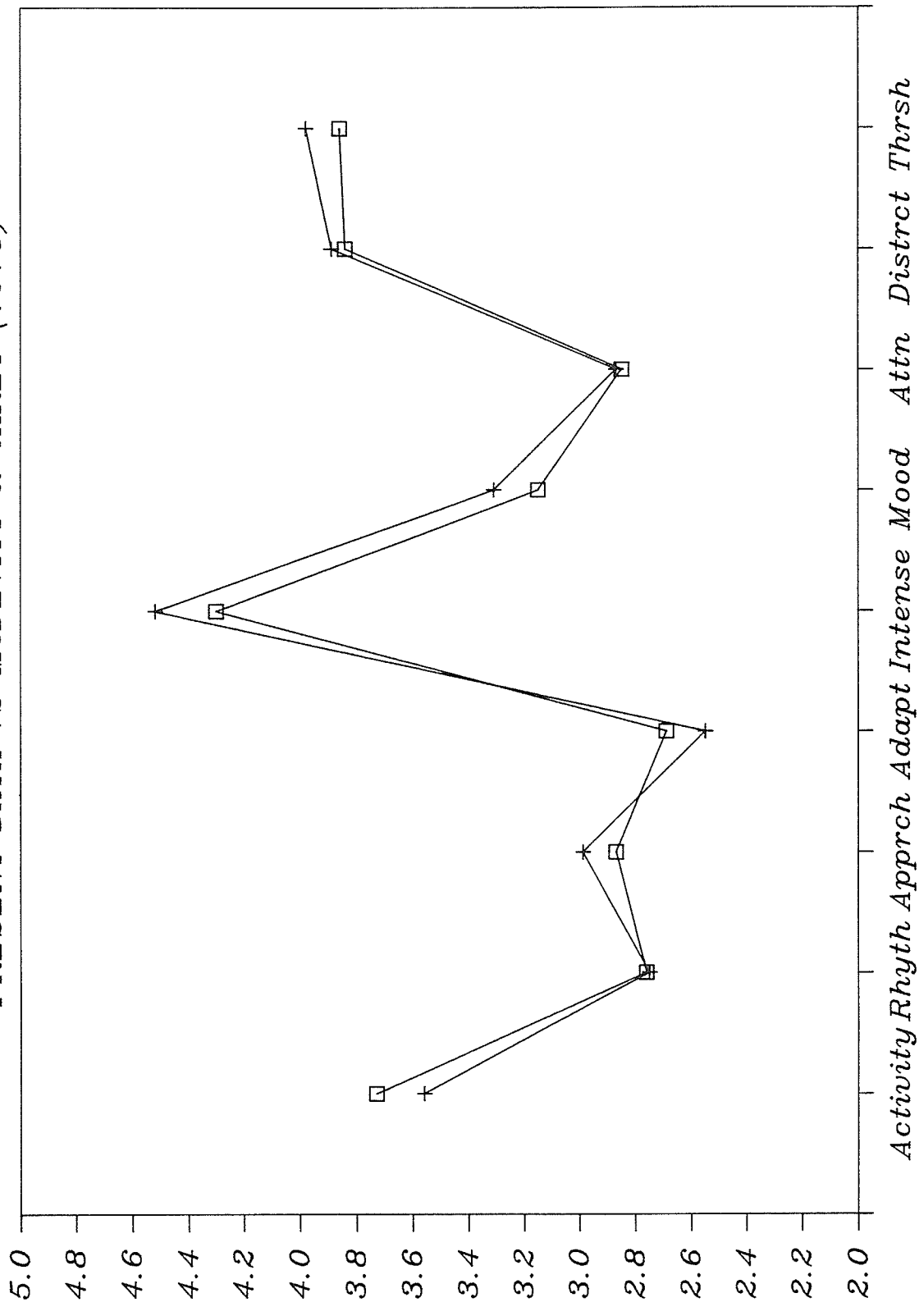
[1]partialling out the linear effects of age, sex and order

[2]N = 92

[3]N = 91

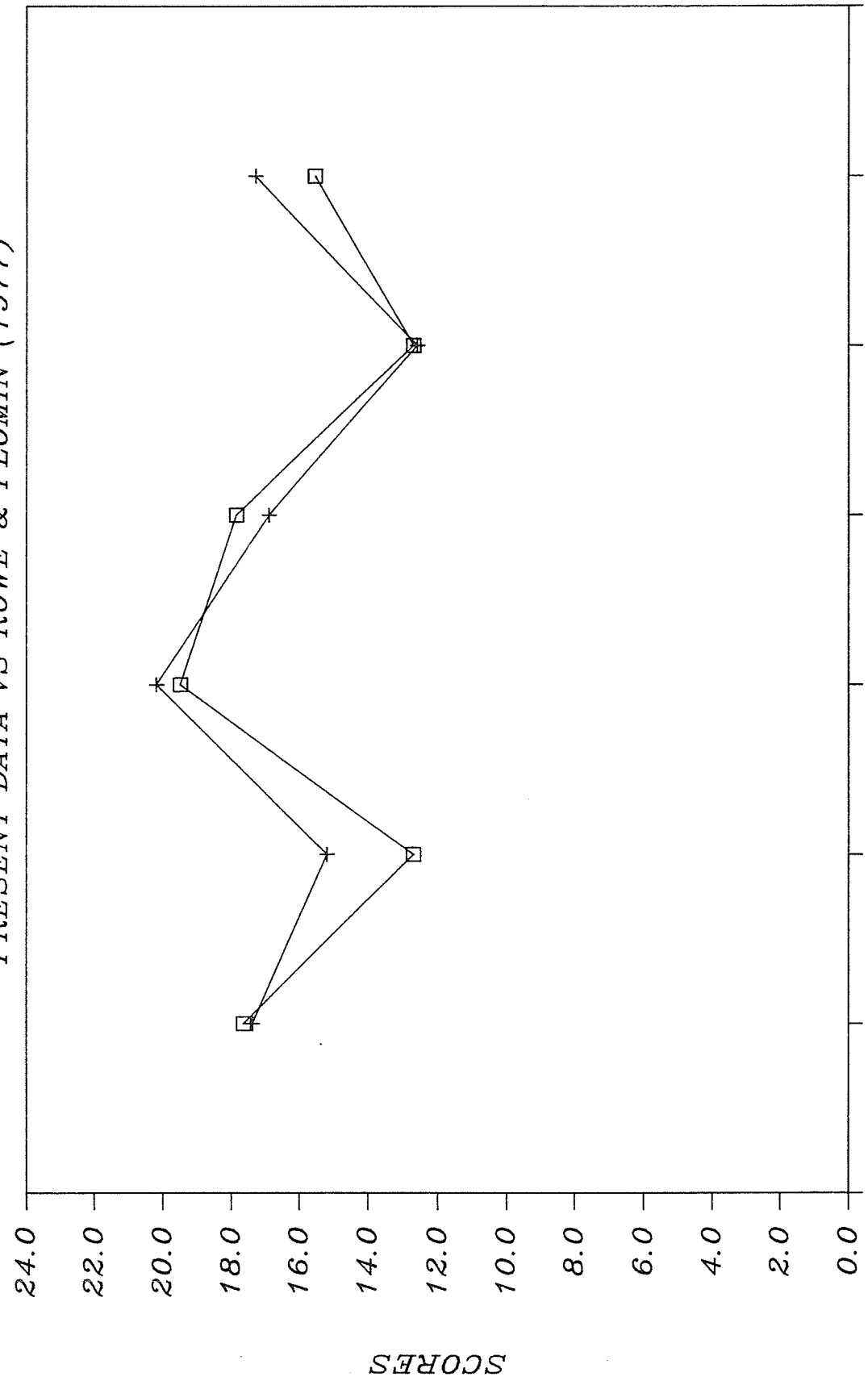
FIGURE 1

*PRESENT DATA VS McDEVITT & CAREY (1978)*



# FIGURE 2

*PRESENT DATA VS ROWE & PLOMIN (1977)*



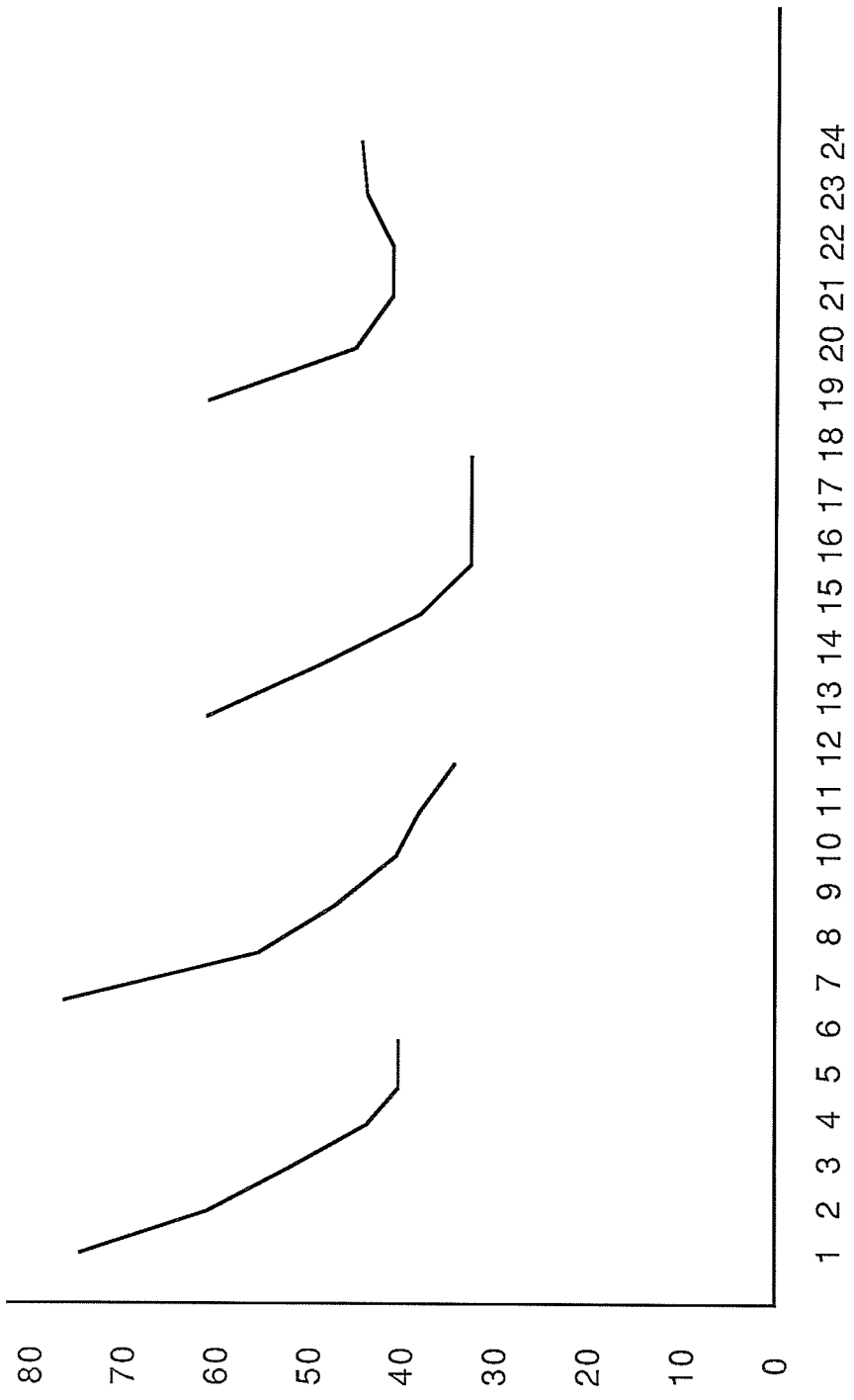



FIGURE 3: GSR MAGNITUDE ACROSS TRIALS

FIGURE 4

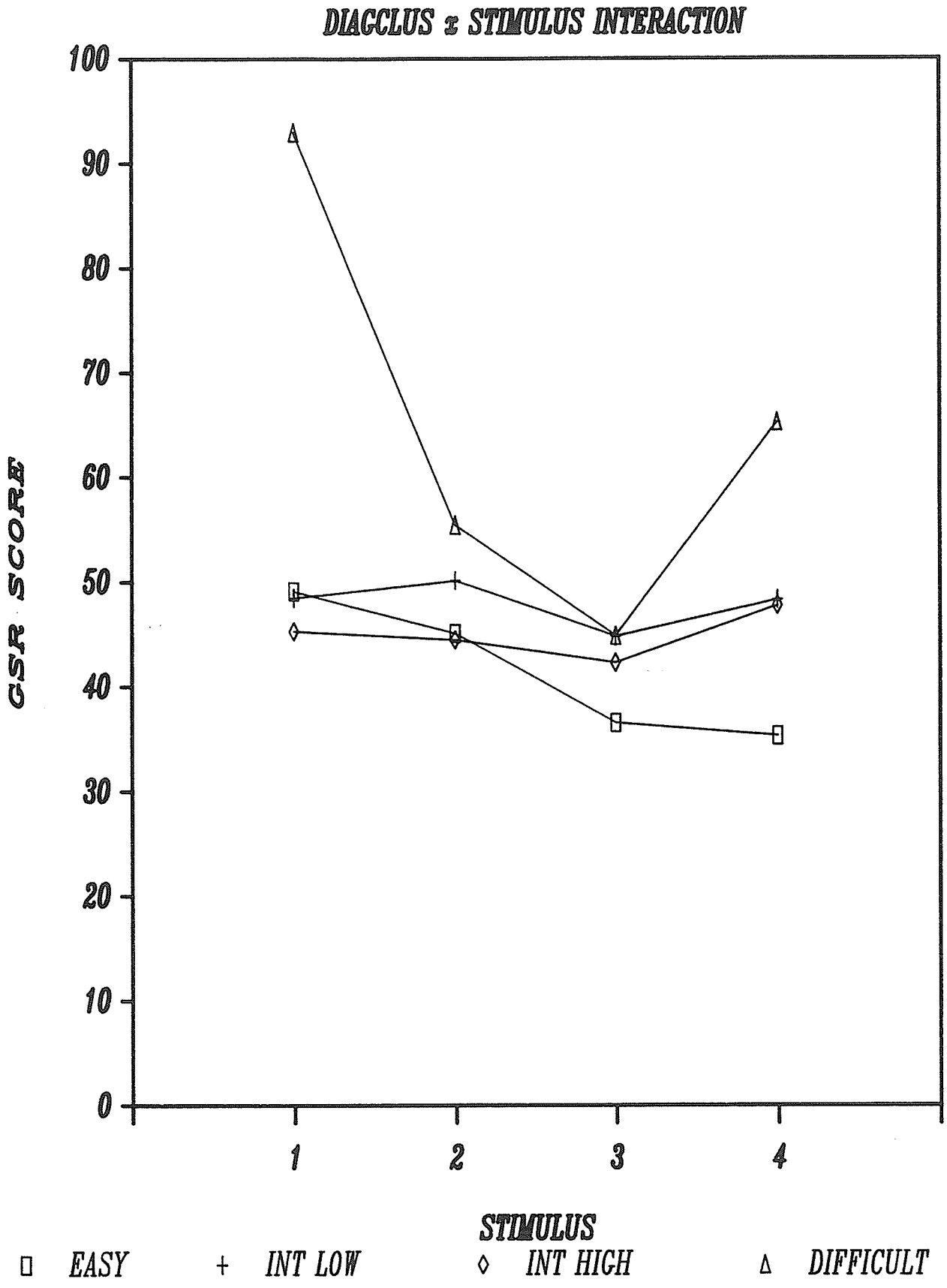
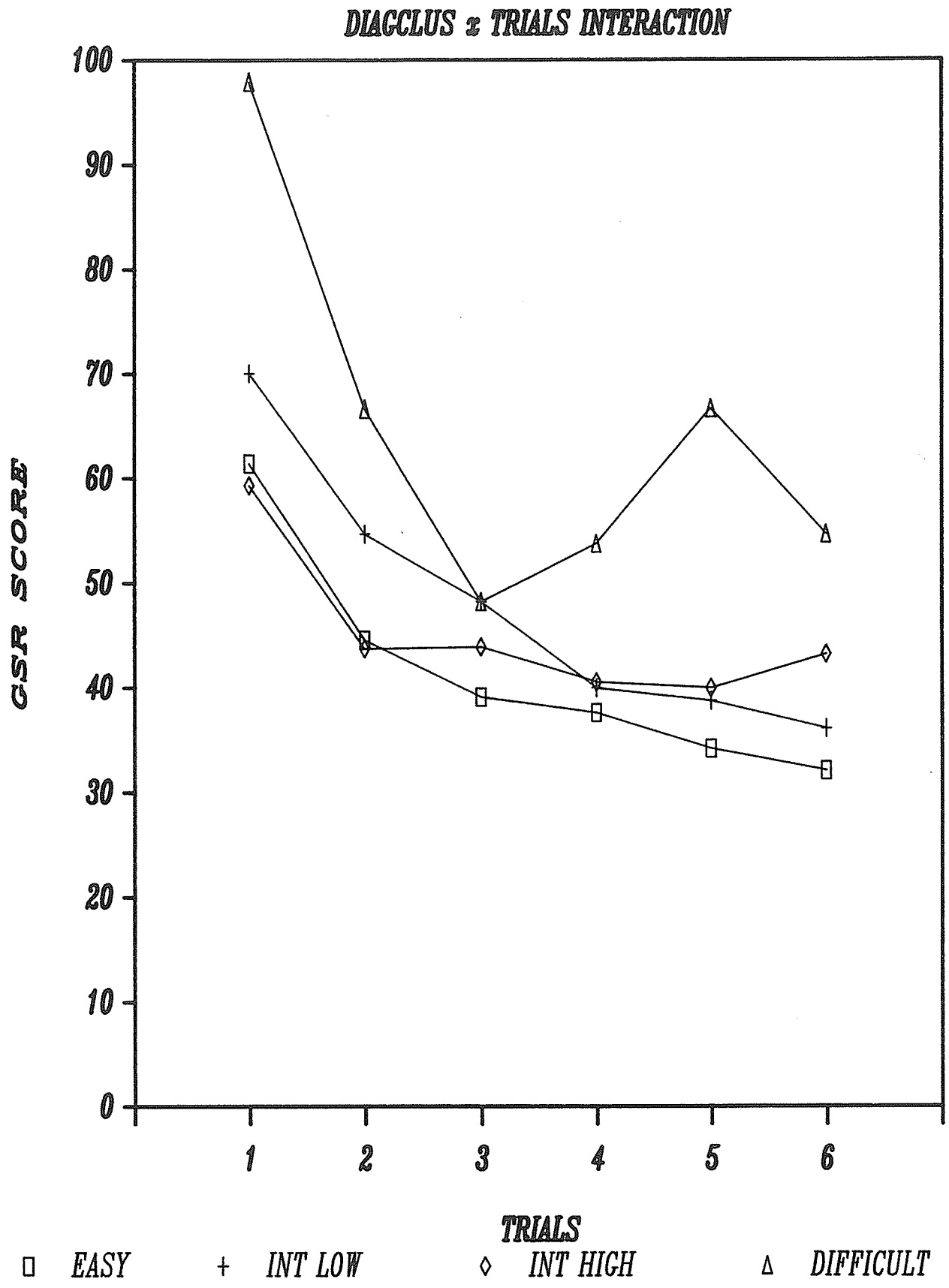

DIAGCLUS x STIMULUS INTERACTION
GSR SCORE
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
STIMULUS
EASY
INT LOW
INT HIGH
DIFFICULT

FIGURE 5


DIACCLUS x TRIALS INTERACTION
GSR SCORE
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
TRIALS
EASY
INT LOW
INT HIGH
DIFFICULT

# Appendix A

## Letter and Permission Form Sent to Parents

November, 1984

Dear parent:

I am conducting my doctoral research at the Friend's University Preschool, which your child attends, and would like to include you and your child in this project. My research is an investigation of the relationship between children's external behavior and their internal biological processes. Anyone who has contact with children knows that they differ in their typical style of interacting with the world almost from the moment they are born. The general area of these stylistic differences is called temperament, and researchers have identified many different sub-categories of temperament such as activity level, emotionality, soothability, attention span/persistence, intensity of response to stimulation, and adaptability to new stimuli. Researchers in this area generally agree that these differences in behavior are related to internal biological processes, but no one has studied exactly how they are related. In my doctoral research I want to see how temperament measures are related to heart rate and galvanic skin response.

This area of research is important because of the implications for parenting and teaching. Before these individual differences were generally recognized, people expected all children to respond in the same way to a given situation. For example, all school-aged children were expected to be able to sit still in their chair through class at school. If a highly active child were unable to do this, the adults involved

might have thought he or she was a disobedient child, and he or she might have been punished. Now that we have more understanding, we recognize that sitting still for that length of time might be easy for one child and next to impossible for another, and we can adjust our expectations according to the abiility of each particular child.

Let me tell you a bit about myself. I came to Wichita in the summer of 1983 to complete my internship training at the Wichita Guidance Center, an agency which specializes in helping children and their families. The internship year was a requirement of my Ph.D. program in Clinical Psychology (this program is approved by the American Psychological Association) at the University of Manitoba in Canada. I completed the intern year in August, 1984, and have completed all the other requirements for the Ph.D. degree, except for this research project. I am married (my husband does human factors work at Boeing) and we have a 14 year old son who is a sophomore at East High School.

If you choose to be included in this research, I will ask you to complete three temperament measures about your child. This should take about 15 minutes of your time, and I will be glad to discuss the results with you when I have finished collecting the measures on your child (I won't be scoring these scales until then to avoid any bias).

The procedures the children will be involved in will take place as follows. I plan to spend some time in each classroom before testing to let the children get to know me. As part of this time the children and their teacher will visit the testing room as a learning experience. While we're there, I'll tell them what the equipment does, and demonstrate it for them, and tell them that they may be coming back to

this room with me another time. When I test the children, I'll escort them to the testing room, and put a heart rate sensor on one finger, and skin response sensor on another finger or on their palm. Then I will measure their heart rate and skin response while they listen to two sounds, and watch a red and a blue light which will be projected onto a screen in front of them. I don't anticipate any of this to be upsetting to the child, but since individual differences do exist, some children may become upset anyway. If this happens, that child will be returned immediately to his or her classroom. When the testing procedure is finished, and I expect it to take 15-20 minutes, children will be allowed to choose a sticker or a balloon for themselves from a box as a "thank you."

I sincerely hope that this letter has given you sufficient information, both about me and my research to make a decision about participating. If it has not, please do not hesitate to call me with any questions or concerns you have. I can be reached at home (681-1393) on Monday, Tuesday, or Wednesday, or at the Wichita Guidance Center (686-6671) on Thursday or Friday (if I am in a session please leave a message and I will call you back). I will also be available for you to meet me and ask questions at the parent's meeting on November 13. If you want to be included in this project, please complete the permission form on the following page and return it to your child's teacher. Thank you.

Sincerely,

Donna Von Bargen, M.A.

I have read the preceding letter, and give my permission for my child, to be included in Donna Von Bargen's doctoral research project at the Friend's University Preschool. My child's name is _____, and his/her birth date is _____.

DATE_____

_____

parent's signature

Appendix B

BEHAVIORAL STYLE QUESTIONNAIRE

(for 3-7 year old children)

by

Sean C. McDevitt, Ph.D. and William B. Carey, M.D.

DATA SHEET

Child's Name _____ Sex _____

Date of Child's Birth _____ Present Age _____
month day year

Rater's Name _____

Relationship to Child _____

Date of Rating _____
month day year

RATING INFORMATION

1. Please base your rating on the child's recent and current behavior (the last four to six weeks).

2. Consider only your own impressions and observations of the child.

3. Rate each question independently. Do not purposely attempt to present a consistent picture of the child.

4. Use extreme ratings where appropriate. Avoid rating only near the middle of the scale.

5. Rate each item quickly. If you cannot decide, skip the item and come back to it later.

6. Rate every item. Circle the number of any item that you are unable to answer due to lack of information or any item that does not apply to your child.

1.

USING THE SCALE SHOWN BELOW, PLEASE MARK AN "X" IN THE SPACE THAT TELLS HOW OFTEN THE CHILD'S RECENT AND CURRENT BEHAVIOR HAS BEEN LIKE THE BEHAVIOR DESCRIBED BY EACH ITEM.

| Almost never | Rarely | Usually does not | Usually does | Frequently | Almost always |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1. The child is moody for more than a few minutes when corrected or disciplined. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 2. The child seems not to hear when involved in a favorite activity. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 3. The child can be coaxed out of a forbidden activity. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 4. The child runs ahead when walking with the parent. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 5. The child laughs or smiles while playing. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 6. The child moves slowly when working on a project or activity. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 7. The child responds intensely to disapproval. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 8. The child needs a period of adjustment to get used to changes in school or at home. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 9. The child enjoys games that involve running or jumping. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 10. The child is slow to adjust to changes in household rules. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 11. The child has bowel movements at about the same time each day. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 12. The child is willing to try new things. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 13. The child sits calmly while watching TV or listening to music. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 14. The child leaves or wants to leave the table during meals. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 15. Changes in plans bother the child. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 16. The child notices minor changes in mother's dress or appearance (clothing, hairstyle, etc.). | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |

| Almost never | Rarely | Usually does not | Usually does | Frequently | Almost always |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 17. The child does not acknowledge a call to come in if involved in something. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 18. The child responds to mild disapproval by the parent ( a frown or shake of the head). | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 19. The child settles arguments with playmates within a few minutes. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 20. The child shows strong reaction to things, both positive and negative. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 21. The child had trouble leaving the mother the first three days when he/she entered school. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 22. The child picks up the nuances or subtleties of parental explanations (example: implied meanings). | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 23. The child falls asleep as soon as he/she is put to bed. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 24. The child moves about actively when he/she explores new places. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 25. The child likes to go to new places rather than familiar ones. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 26. The child sits quietly while waiting. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 27. The child spends over an hour reading a book or looking at the pictures. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 28. The child learns new things at his/her level quickly and easily. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 29. The child smiles or laughs when he/she meets new visitors at home. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 30. The child is easily excited by praise. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 31. The child is outgoing with strangers. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 32. The child fidgets when he/she has to stay still. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |
| 33. The child says that he/she is "bored" with his/her toys and games. | almost never | 1 | 2 | 3 | 4 | 5 | 6 | almost always |

3.

| Almost never 1 | Rarely 2 | Usually does not 3 | Usually does 4 | Frequently 5 | Almost always 6 |
|---|---|---|---|---|---|

| | | | |
|---|---|---|---|
| 34. The child is annoyed at interrupting play to comply with a parental request. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 35. The child practices an activity until he/she masters it. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 36. The child eats about the same amount at supper from day to day. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 37. Unusual noises (sirens, thunder, etc.) interrupt the child's behavior. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 38. The child complains when tired. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 39. The child loses interest in a new toy or game the same day. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 40. The child becomes engrossed in an interesting activity for one half hour or more. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 41. The child cries intensely when hurt. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 42. The child reacts strongly to kidding or light-hearted comments. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 43. The child approaches children his/her age that he/she doesn't know. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 44. The child plays quietly with his/her toys and games. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 45. The child is outwardly expressive of his/her emotions. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 46. The child is enthusiastic when he/she masters an activity and wants to show everyone. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 47. The child is sleepy at his/her bed-time. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 48. The child stops an activity because something else catches his/her attention. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 49. The child is hungry at dinner time. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 50. The child holds back until sure of himself/herself. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |

4.

| Almost never | Rarely | Usually does not | Usually does | Frequently | Almost always |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 51. The child looks up when someone walks past the door-way. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 52. The child becomes upset if he/she misses a regular television program. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 53. The child reacts strongly (cries or complains) to a disappointment or failure. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 54. The child accepts new foods within one or two tries. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 55. The child has difficulty getting used to new situations. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 56. The child will avoid misbehavior if punished firmly once or twice. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 57. The child is sensitive to noises (telephone, doorbell) and looks up right away. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 58. The child prefers active outdoor play to quiet play inside. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 59. The child dislikes milk or other drinks if not ice-cold. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 60. The child notices differences or changes in the consistency of food. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 61. The child adjusts easily to changes in his/her routine. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 62. The child eats about the same amount at breakfast from day to day. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 63. The child seems to take setbacks in stride. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 64. The child cries or whines when frustrated. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 65. The child repeats behavior for which he/she has previously been punished. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 66. The child looks up from playing when the telephone rings. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |
| 67. The child is willing to try new foods. | almost never | ___ | ___ | ___ | ___ | ___ | ___ | almost always |
| | | 1 | 2 | 3 | 4 | 5 | 6 | |

5.

| Almost never | Rarely | Usually does not | Usually does | Frequently | Almost always |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |

| | | | |
|---|---|---|---|
| 68. The child needs encouragement before he/she will try new things. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 69. The child cries or whines when ill with a cold or upset stomach. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 70. The child runs to get where he/she wants to go. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 71. The child's attention drifts away or lapses when listening to parental instructions. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 72. The child becomes angry with one of his/her playmates. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 73. The child is reluctant to give up when trying to do a difficult task. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 74. The child reacts to mild approval from the parent (a nod or smile). | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 75. The child requests "something to eat" between meals and regular snacks. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 76. The child rushes to greet the parent or greets loudly after absence during the day. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 77. The child looks up when he/she hears voices in the next room. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 78. The child protests when denied a request by by the parent. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 79. The child ignores loud noises when reading or looking at pictures in a book. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 80. The child dislikes a food that he/she had previously seemed to accept. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 81. The child stops what he/she is doing and looks up when the parent enters the room. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 82. The child cries for more than a few minutes when hurt. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 83. The child watches a long ( 1 hour or more) TV program without etting up to do something else. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 84. The child spontaneously wakes up at the usual time on weekends and holidays. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |

| Almost never 1 | Rarely 2 | Usually does not 3 | Usually does 4 | Frequently 5 | Almost always 6 |
|---|---|---|---|---|---|

| | | | |
|---|---|---|---|
| 85. The child responds to sounds or noises unrelated to his/her activity. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 86. The child avoids new guests or visitors. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 87. The child fidgets when a story is being read to him/her. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 88. The child becomes upset or cries over minor falls or bumps. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 89. The child interrupts an activity to listen to conversation around him/her. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 90. The child is unwilling to leave a play activity that he/she has not completed. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 91. The child is able to fall asleep when there is conversation in a nearby room. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 92. The child becomes highly excited when presented with a new toy or game. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 93. The child pays attention from start to finish when the parent tries to explain something to him/her. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 94. The child speaks so quickly that it is sometimes difficult to understand him/her. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 95. The child wants to leave the table during meals to answer the doorbell or phone. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 96. The child complains of events in school or with playmates that day. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 97. The child frowns when asked to do a chore by the parent. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 98. The child tends to hold back in new situations. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 99. The child laughs hard while watching television cartoons or comedy. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |
| 100. The child has "off" days when he/she is moody or cranky. | almost never | ___:___:___:___:___:___ 1 2 3 4 5 6 | almost always |

Appendix C

The EAS Temperament Survey for Children

Name_____ Date_____

Rate each of the items for your child on a scale of 1 (not characteristic or typical of your child) to 5 (very characteristic or typical of your child).

1. Child tends to be shy.

1----------2----------3----------4----------5
not typical very typical

2. Child cries easily.

1----------2----------3----------4----------5
not typical very typical

3. Child likes to be with people.

1----------2----------3----------4----------5
not typical very typical

4. Child is always on the go.

1----------2----------3----------4----------5
not typical very typical

5. Child prefers playing with others rather than alone.

1----------2----------3----------4----------5
not typical very typical

6. Child tends to be somewhat emotional.

1----------2----------3----------4----------5
not typical very typical

7. When child moves about, he usually moves slowly.

1----------2----------3----------4----------5
not typical very typical

8. Child makes friends easily.

1----------2----------3----------4----------5
not typical very typical

9. Child is off and running as soon as he wakes up in the morning.

1----------2----------3----------4----------5
not typical very typical

The EAS Temperament Survey for Children - Page 2

10. Child finds people more stimulating than anything else.

1----------2----------3----------4----------5
not typical very typical

11. Child often fusses and cries.

1----------2----------3----------4----------5
not typical very typical

12. Child is very sociable.

1----------2----------3----------4----------5
not typical very typical

13. Child is very energetic.

1----------2----------3----------4----------5
not typical very typical

14. Child takes a long time to warm up to strangers.

1----------2----------3----------4----------5
not typical very typical

15. Child gets upset easily.

1----------2----------3----------4----------5
not typical very typical

16. Child is something of a loner.

1----------2----------3----------4----------5
not typical very typical

17. Child prefers quiet, inactive games to more active ones.

1----------2----------3----------4----------5
not typical very typical

18. When alone, child feels isolated.

1----------2----------3----------4----------5
not typical very typical

19. Child reacts intensely when upset.

1----------2----------3----------4----------5
not typical very typical

20. Child is very friendly with strangers.

1----------2----------3----------4----------5
not typical very typical

Appendix D

Colorado Childhood Temperament Inventory

Instructions: Please circle the response that best represents your child's behavior. For instance,

| | Not at all like my child | A little like my child | On the average like my child | Most often like my child | A lot or always like my child |
|---|---|---|---|---|---|
| Child shows fear of stranger. | 1 | 2 | (3) | 4 | 5 |

You may change an answer by crossing out the incorrect response and circling another response. For instance,

| | | | | | |
|---|---|---|---|---|---|
| Child shows fear of strangers. | 1 | ~~2~~ | 3 | (4) | 5 |

For each of the following items, circle one response to describe your child's behavior.

| | Not at all like my child | A little like my child | On the average like my child | Most often like my child | A lot or always like my child |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 1. Child makes friends easily. | 1 | 2 | 3 | 4 | 5 |
| 2. Child gets upset easily. | 1 | 2 | 3 | 4 | 5 |
| 3. Child is very energetic. | 1 | 2 | 3 | 4 | 5 |
| 4. Plays with a single toy for long periods of time. | 1 | 2 | 3 | 4 | 5 |
| 5. Rarely took a new food without fussing. | 1 | 2 | 3 | 4 | 5 |
| 6. Whenever child starts crying, he can be easily distracted. | 1 | 2 | 3 | 4 | 5 |
| 7. Child is very friendly with strangers | 1 | 2 | 3 | 4 | 5 |
| 8. Child tends to be somewhat emotional. | 1 | 2 | 3 | 4 | 5 |
| 9. Child is always on the go. | 1 | 2 | 3 | 4 | 5 |

| | | Not at all like my child | A little like my child | On the average like my child | Most often like my child | A lot or always like my child |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| 10. | Child persists at a task until successful. | 1 | 2 | 3 | 4 | 5 |
| 11. | Child consistently dislikes many kinds of food. | 1 | 2 | 3 | 4 | 5 |
| 12. | When upset by an unexpected situation, child quickly calms down. | 1 | 2 | 3 | 4 | 5 |
| 13. | Child is very sociable. | 1 | 2 | 3 | 4 | 5 |
| 14. | Child reacts intensely when upset. | 1 | 2 | 3 | 4 | 5 |
| 15. | Child prefers quiet, inactive games to more active ones. | 1 | 2 | 3 | 4 | 5 |
| 16. | Child goes from toy to toy quickly. | 1 | 2 | 3 | 4 | 5 |
| 17. | Child makes faces at new foods. | 1 | 2 | 3 | 4 | 5 |
| 18. | Child stopped fussing whenever someone talked to him or picked him up. | 1 | 2 | 3 | 4 | 5 |
| 19. | Child takes a long time to warm up to strangers. | 1 | 2 | 3 | 4 | 5 |
| 20. | Child cries easily. | 1 | 2 | 3 | 4 | 5 |
| 21. | Child is off and running as soon as he wakes up in the morning. | 1 | 2 | 3 | 4 | 5 |
| 22. | Child gives up easily when difficulties are encountered. | 1 | 2 | 3 | 4 | 5 |

| | | Not at all like my child | A little like my child | On the average like my child | Most often like my child | A lot or always like my child |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| 23. | Once the child decides he doesn't like something, there is no getting him to like it. | 1 | 2 | 3 | 4 | 5 |
| 24. | If talked to, child stops crying. | 1 | 2 | 3 | 4 | 5 |
| 25. | Child tends to be shy. | 1 | 2 | 3 | 4 | 5 |
| 26. | Child often fusses and cries. | 1 | 2 | 3 | 4 | 5 |
| 27. | When child moves about, he usually moves slowly. | 1 | 2 | 3 | 4 | 5 |
| 28. | With a difficult toy, child gives up quite easily. | 1 | 2 | 3 | 4 | 5 |
| 29. | Child has strong likes and dislikes in food. | 1 | 2 | 3 | 4 | 5 |
| 30. | Child tolerates frustration well. | | | | | |