Development and Psychopathology, **14** (2002), 463–475
Copyright © 2002 Cambridge University Press
Printed in the United States of America
DOI: 10.1017.S0954579402003048

One measure, one meaning: Multiple measures, clearer meaning

JEROME KAGAN, NANCY SNIDMAN, MARK McMANIS, SUE WOODWARD, AND CHRISTINA HARDWAY
Harvard University

Abstract
This paper tries to make three points. First, current constructs in personality and psychopathology are based on the restrictive evidence contained in self-reports. As a result, heterogeneous categories of individuals are assigned to the same category. Second, it is suggested that when different sources of evidence are included, theoretically distinct groups will be detected within the prior heterogeneous category. Third, the authors argue that physiological information has the potential to parse individuals with similar phenotypes on self-report data into distinct groups that reveal the temperamental origins of their phenotype.

Scientists engage nature in a contest in which she presents a sign and investigators try to determine its meaning, especially those features in the meaning network that reveal the origin of the natural event. When a sufficiently large number of meanings have been decoded, scientists can predict some signs that nature has not yet uncovered. This theoretically more mature stage is restricted, at present, to physics and parts of chemistry and molecular biology. Those who study human behavior are constrained to a dependent posture in which nature holds all the answers and scientists have to guess at the meaning of every phenomenon.

This state of affairs is illustrated with an example from development. Some school-age children are quiet, subdued, and withdrawn when they first meet children or adults they do not know (Kagan, 1994). Scientists wish to learn three things about this phenomenon: What combination of biological and experiential factors produced it? Can this profile be produced by more than one set of conditions? What other properties, correlated with the behavioral pattern, are part of a coherent configuration? When these three questions have been answered, scientists can decide what name, or names, to apply to a child who is shy and subdued in the presence of unfamiliar people.

Unfortunately, some scientists interested in this profile assume two things: first, that only one set of conditions generated it and, second, that the shyness, which is obvious to observers, is the most important feature of the set of processes that is the foundation of the behavior. Therefore, the name of the construct should refer explicitly to that property.

The problem associated with the habit of attributing primacy to measurements that are easily observed or imagined is illustrated in a series of investigations of infant rat pups. When scientists took newborn rat pups away from their nursing mother for a few minutes each day over the first 2 weeks of life, the adult animals showed less emotionality in a

This research was supported by the W. T. Grant Foundation and the Bial Foundation. We thank Robin Schacht, Tracey Shissler, Eric Peterson, and Melissa Lewis for their contributions to the research.

Address correspondence and reprint requests to: Jerome Kagan, Daniel and Amy Starch Professor of Psychology, Department of Psychology, Harvard University, William James Hall, 33 Kirkland Street, Cambridge, MA 02138-2044; E-mail: jk@wjh.harvard.edu.

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

novel situation than the adults who had been reared normally as infants. Other scientists discovered later that the rat pups separated from the mother underwent a permanent change in brain physiology; specifically, the hippocampus developed a larger number of receptors for glucocorticoids compared to normally reared animals. This finding was regarded as proof that stressful events, in this case separation from the mother, had a permanent effect on the developing rat. However, to the surprise of many, later work revealed that the primary cause of the change in the hippocampus was not the stress of being removed from the mother but rather the fact that the mother licked and groomed the pups more vigorously when they were returned to her because the cooler temperature of the separated rat pups alerted her and led her to vigorously lick her infants. The licking increased brain serotonin that, in turn, led to increased numbers of glucocorticoid receptors in the hippocampus, which then led to increased activity in the hypothalamic nuclei that produce glucocorticoids. As a result, the adult rats who had been separated as infants were less emotional when placed in an unfamiliar environment. However, it was the vigorous, tactile stimulation of the infant pups, not the stress of being removed from the mother, that produced the altered brain anatomy and later adult behavior (Denenberg, 1964). This story illustrates how easy it is to attribute causal influence to the most obvious events and how difficult it is to imagine the subtle, or invisible, events that are of greater priority. The habit of being seduced by what appears obvious is an example of Francis Bacon's "Idols of the Tribe," one of his four categories of mistaken beliefs. This error is being made in classifications of anxiety and depressive disorder as investigators award priority to the affective features because they are obvious and easy to imagine.

Fatigue supplies an analogy. Physicians have learned that a feeling of fatigue is a feature of many different natural phenomena; each is a component of a unique configuration with a special set of prior conditions. For example, a viral infection, leukemia, and insomnia due to worry over a spouse's illness are three different causes of fatigue. That is why physicians gather additional information, including biological data, in order to decide on the best diagnostic category for the patient. The diagnostic category will not refer to fatigue but instead to its presumed cause and correlated properties.

Many conceptual terms in current writing on personality and psychopathology resemble the concept of fatigue. That is, most personality constructs and diagnostic categories have more than one history and physiological configuration. Each construct in these systems refers to one or two features that are a small part of the larger configuration that nature intended. Progress in these domains of inquiry will require gathering additional information (behavioral, attitudinal, cognitive, and biological) in order to arrive at a richer understanding and more accurate conceptual name for each configuration.

The Validity of Questionnaires

Psychologists and psychiatrists resist the implications of this state of affairs more than biologists. Most scholars who study personality rely only on self-reports to questionnaires; those interested in the epidemiology of psychopathology typically rely only on interviews. At least six factors compromise the accuracy of self-report. We use parental reports of their children's behaviors to illustrate the nature of the compromise.

First, parent's verbal descriptions of the child are influenced by their notions of the ideal for their society. Parents who want outgoing children are threatened by very quiet ones and may deny their children's timidity. Second, parents are sensitive to the logical consistency of their answers to a series of questions. A mother who reports that her daughter is happy and relaxed will resist acknowledging on other questions that her child is occasionally sad. Third, it is difficult to describe blends of emotions and motives because the words in questionnaires refer to discrete behavioral categories. There is no English word, for example, that describes the feeling generated when a child hopes for parental forgiveness for a misdemeanor but fears

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

punishment. Fourth, every question assumes a comparison referent. When a parent reads on a questionnaire, "Does your child like to go to parties?" he or she unconsciously compares that preference with others. If one mother compares going to parties with an activity that is disliked whereas a second parent compares it with one that is also preferred, the former is more likely to endorse the item than the latter, even though both children may like parties to an equivalent degree. Fifth, parents who have not had extensive experience with children have a less accurate basis for judging their first child than experienced parents judging their third or fourth child. Sixth, parents differ in their understanding of the meanings of the words on questionnaires. In an unpublished study from our laboratory with Maureen Razendes (1998), adolescent subjects who had watched a specially prepared film of two adolescents competing in a contest were asked "Which one in the film was more anxious?" Some subjects interpreted the word anxious as meaning "eager to perform well," while others interpreted the same question as meaning that the actors in the film were fearful. Some mothers regard the word shy as implying that their child is sensitive to others; other parents interpret the same word as meaning that their child is cautious about trusting others; still others interpret the term as if it meant a fear of strangers. These are some of the reasons why parental descriptions of children on questionnaires are not always valid indexes of the child's profile and why there is poor to modest correspondence between these descriptions and direct behavioral observations (Kagan, 1998; Klein, 1991; Perrin & Last, 1992; Rosicky, 1993; Seifer, Sameroff, Barrett, & Krafchuk, 1994; Spiker, Kramer, Constantine, & Bryant, 1992).

The Emphasis on Feelings

Because feelings of fear, anxiety, and sadness are salient to consciousness, the use of self-reports tempts investigators to award this feature priority and leads them to ignore less salient properties that might provide more relevant information. The diagnostic category of "depressive disorder" illustrates this problem.

Adolescents who report feelings of depression can belong to at least three different groups. One group consists of minority youth from chaotic families in poverty who believe they have little chance of escaping from their frustrating conditions to gain a more satisfying life. A second group consists of youths in economically secure, suburban homes with predictable parents who set such extremely high standards for the child's academic success that the adolescents feel they are unable to meet these standards. The third group consists of adolescents with a first degree relative with bipolar disorder who have inherited a biological vulnerability to a depressed mood.

If clinicians and investigators gathered information on the patient's family and their values, neighborhood conditions, school achievement, friendships, and physiology, they would discover that these patients belonged to three different configurations. Youths in the first group are likely to be ashamed of their family's disadvantaged status, angry at society for its prejudice and indifference to their state of poverty, academically compromised, and discouraged over the probability of changing these unhappy conditions. Youths in the second group are more likely to experience guilt over their failure to meet the family's high academic standards. Youths in the third group will feel both less helpless and less guilty than those in the first two groups. Finally, each of the three groups will display a distinct physiological profile.

Clinical psychologists and psychiatrists are not the only investigators who assign heterogeneous phenomena to a single unitary category. Behavioral biologists and neuroscientists who study fear in rats often restrict their evidence to one dependent variable. For example, some scientists treat the duration of freezing to a conditioned stimulus that had been associated with electric shock as an index of fear (LeDoux, 1996). Others regard the magnitude of potentiated startle as an index of fear (Davis, Schlesinger, & Sorenson, 1989). When scientists discovered that rats with a lesioned amygdala did not acquire an association between a conditioned stimulus and either freezing or potentiated startle, they concluded that the amygdala mediates the ac-

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

quisition of a fear state. However, rats with a lesioned amygdala can acquire a conditioned defecation reaction to a distinctive chamber in which they received electric shock, even though they fail to acquire conditioned freezing behavior (Antoniadas & McDonald, 2000).

This last fact means that fear cannot be a unitary state, and it is necessary to assume at least two different fear constructs. One is the state of animals who have an amygdala; another refers to the state of animals who do not have an amygdala. Further, if the intensity of the electric shock used as the unconditioned stimulus in the acquisition of a startle response is very intense, then the startle reflex is not potentiated. The startle is only potentiated if the shock is of moderate intensity (Davis & Astrachan, 1978; Walker & Davis, 1997). It seems a mistake to conclude that rats who experienced intense shock during training were not in a state of fear when the conditioned stimulus appeared.

There is additional support for rejecting the notion of a single fear state. Rats placed in an elevated maze with brightly lit alleys avoided those alleys and moved to a darker place. Scientists claimed that the brightly lit alleys provoke a state of fear. Rats also learn to avoid an object that delivers an electric shock when they contact it; the avoidance is assumed to reflect a fear state. However, the avoidance of brightly lit alleys requires the septum but not the amygdala; avoidance of an object that delivers shock requires the amygdala but not the septum (Treit, Pesold, & Rotzinger, 1993). Even though both reactions are assumed to reflect a fear state, each is mediated by a different brain circuit. Thus, the gathering of additional data leads to a more valid understanding.

The remainder of this paper illustrates the advantages of gathering multiple sources of evidence when investigators explore a psychological phenomenon by using a corpus of information gathered on children participating in a longitudinal study of the temperamental categories we call high and low reactivity to unfamiliar events (Kagan, 1994).

High and Low Reactivity in Infants

The corpus of evidence comes from a large group of Caucasian 10- to 11-year-old children born at term to healthy middle-class mothers who experienced an untoward pregnancy. When the infants were 4 months old, they were filmed while presented with a standard battery that included the mother looking at the infant, brief sentences played through a speaker baffle, colorful mobiles moved in front of the infant's face, a cotton swab dipped in dilute butyl alcohol applied to the nostrils, and tape-recorded nonsense syllables played at different levels of loudness (Kagan, 1994).

Quantification of the motor activity and crying to the entire battery revealed that 22% of the 450 infants showed high levels of motor activity combined with frequent distress; these infants are called high reactive. About 40% of the sample showed minimal motor activity and minimal distress; these infants are called low reactive. The remaining 38% who showed other patterns, belong to other temperamental groups that will not be considered in this article.

Most of these children (n = 369) returned to the laboratory at 14 and 21 months, when they were exposed to a variety of unfamiliar events, some social and some nonsocial. The events included rotation of small plastic objects in a wheel; an examiner who frowned and spoke in a serious voice, the placement of heart rate electrodes and a blood pressure cuff; and inviting the child to display novel actions, to accept liquid on the tongue, and to approach a stranger, a clown, and an unfamiliar metal robot. Each episode was scored for the presence of a fear reaction, where fear was defined as fretting or crying to any one of the unfamiliar incentives or failing to approach the stranger, clown, or robot, despite a friendly invitation to do so. About 30% of the sample showed four or more fears at both 14 and 21 months; this score reflected a high level of fear. The children who had been high reactive infants at 14 and 21 months were significantly more fearful than those who had been low reactive (see Table 1). Fifty percent of the children who had been high reactive infants had four or more fears at both ages, compared to only 10% of those who had been low reactive. By contrast, 75% of those who had been low reactive had either no fears or one fear at both ages (Kagan, 1994).

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

Table 1. *Differences between high reactive and low reactive infants from 14 months to 7.5 years*

Variable	High Reactive	Low Reactive	Significance Test
Fear at 14 months	4.2	2.5	t (286) = 11.1, $p < .001$
Fear at 21 months	3.4	2.0	t (243) = 4.67, $p < .001$
Spontaneous comments	40.7	53.7	t (191) = 1.96, $p < .05$
Smiles at 4.5 years	17.8	27.7	t (191) = 3.69, $p < .01$
Inhibited with peers at 4.5 years (%)	46	10	
Uninhibited with peer at 4.5 years (%)	27	67	$\chi^2 = 35.5$, $p < .001$
Spontaneous comments at 7.5 years	19.6	32.8	t (109) = 2.37, $p < .01$
Smiles at 7.5 years	14.4	22.0	t (109) = 1.97, $p < .05$
Anxious at 7.5 years (%)	45	15	$\chi^2 = 12.8$, $p < .0l$

A proportion of these children was evaluated again when they were 4.5 and 7.5 years old. If we assume that the inhibition of speech in an unfamiliar social setting reflects uncertainty to unfamiliar events, then the high reactives should be less talkative and less likely to smile while interacting with an unfamiliar adult than low reactives. Each child was interviewed by an unfamiliar female examiner over a 60-min battery that involved cognitive tests and measurements of autonomic reactivity. The videotapes of these sessions were coded for the number of spontaneous comments and smiles, where a comment was defined as any remark that was not a direct answer to the examiner's question. High reactives displayed significantly fewer comments and smiles than the low reactives (Kagan, Snidman, & Arcus, 1998a).

Each child and their parents returned between 3 and 6 weeks later for a play session with two other unfamiliar peers of the same age and sex. Coders behind a one-way vision screen recorded each child's behavior for time sitting or standing close to the mother, time talking, time close to another peer, and total time staring at another child while not proximal to that child. After 30 min of play, the mothers left the room for 5 min and the examiner placed a single attractive toy in the room for a 3-min period to see which child would seize it. Finally, after the parents had returned, an adult dressed in a gorilla costume carrying toys entered the room and, after a minute, invited each child to approach and play with the toys.

This entire corpus of data was coded reliably for whether the child was sociable or uninhibited, on the one hand, or shy or inhibited, on the other. Fifty-one percent were classified as uninhibited, 26% were classified as inhibited, and the remaining 23% fit neither category. The children who had been high reactive infants were more likely than the low reactives to be classified as inhibited (46 vs. 10%; see Table 1).

A group of 164 of the children from this sample was evaluated when they were 7.5 years old. Interviews with the mothers and the teachers permitted us to classify each child as meeting or failing to meet a criterion for anxious symptoms. Forty-three of the 164 children were classified as possessing anxious symptoms, and 45% of the high reactives, but only 15% of the low reactives, had anxious symptoms (Kagan, Snidman, Zentner, & Peterson, 1999; see Table 1). In addition, the high reactives made fewer comments and smiles than the low reactives (see Table 1).

The 23 high reactive children who had developed anxious symptoms were compared with 27 high reactives who did not have such symptoms. The former group had a distinct profile on other variables. More of the anxious high reactives had narrow facial skeletons, higher sitting diastolic blood pressure, and greater magnitude of cooling of the temperature of the fingertips while listening to a series of digits that they were instructed to remember. Moreover, 7 of the 23 high reactive, high anxious children, but not one of the high reactive, low anxious children, had been extremely fearful in the second year and at 4 months displayed extremely high motor arousal and distress scores, high heart rates, and large heart rate decelerations to the tape

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

recording of a female voice speaking short sentences. By combining the behavioral and biological data, we were able to detect a small group that was behaviorally and physiologically different from the other high reactive infants at as early as 16 weeks of age. This group may represent a distinct category.

The Assessment at Age 10 Years

A total of 49 children who had been high reactive and 65 who had been low reactive were seen in our laboratory when they were between 10 and 11 years of age for a 3-hr battery that included evaluations of electroencephalogram (EEG) patterns, brain stem auditory evoked potential, startle to acoustic probes, and social behavior with the examiner. In addition, the mothers and children were interviewed at a later time in their home by Christina Hardaway for her unpublished dissertation research. Because the number of spontaneous comments and smiles displayed with a female examiner had been sensitive correlates of high or low reactivity in the past, we coded these two variables during the first 18 min of the interaction between the child and the examiner. After this time the children were required to be very quiet while we measured the biological variables.

As expected, the 10-year-old children who had been high reactive were less talkative and smiled less often than the low reactives. Thirty-four percent of the high reactives, but only 13% of the low reactives, made fewer than eight comments and displayed fewer than eight smiles during the first 18 min (this profile is regarded as subdued; $\chi^2 = 7.1$, $p < .01$). Further, these subdued 10-year-old high reactives were also subdued at 7 years (see Figure 1). Forty-two percent of the high reactives, but only 10% of the low reactives, were subdued with the examiner at both 7 and 10 years. By contrast, 43% of the low reactives, but only 7% of the high reactives, displayed frequent spontaneous comments at both ages and 30% of the low reactives, but only 3% of the high reactives, smiled frequently at both ages. The odds ratio for high versus low reactive with respect to comments was 24; the odds ratio for smiles was 25, $p < .01$. To our surprise, there was even a predictive relation between frequency of smiling during the original battery at 4 months and frequency of smiling at 10 years. Significantly more low than high reactives were frequent smilers at both age while more high reactives were infrequent smilers, $\chi^2 = 10.3$, $p < .01$.

Although more high than low reactives were subdued with the examiner, the low reactives who were subdued appear to observers to be equally shy. Thus, we asked whether, among the subdued children, the high and low reactives displayed different physiological profiles. Such a result would support the plea for combining sources of evidence. The physiological variables of interest were asymmetry of frontal activation in the EEG, amplitude of the brain stem auditory evoked potential from the inferior colliculus (wave 5), and amount of beta power in the EEG at rest.

Davidson (1995) and Fox, Calkins, and Bell (1994) showed that greater desynchronization of EEG alpha frequencies (usually 6–13 Hz) in the right compared with the left frontal area is most characteristic of anxious adults and inhibited children, although the origins of this asymmetry are not yet understood.

The magnitude of the evoked potential produced by the inferior colliculus to a series of clicks was quantified because the colliculus is the first structure in the auditory pathway that is primed by the amygdala and is part of a fear circuit in animals. Electrical stimulation of the inferior colliculus in rats leads to freezing and other reactions usually regarded as signs of a fear state (Brandao, Cardoso, Melo, Motta, & Coimbra, 1994). The basilar membrane, cochlear nucleus, superior olive, and medial lemniscus are not primed by the amygdala. Thus, Woodward, McManis, Kagan, Deldin, Snidman, Lewis, and Kahn (2001) argued that a larger evoked potential from the inferior colliculus, called wave 5, might reflect an excitable amygdala. If high reactives had a larger wave 5, the hypothesis of greater amygdalar activity for this group would be supported. The variable of interest was the difference (in microvolts, μV) between the trough of wave 3, from the superior olive, and the peak of wave 5, from the inferior colliculus. Finally, research has demonstrated that

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

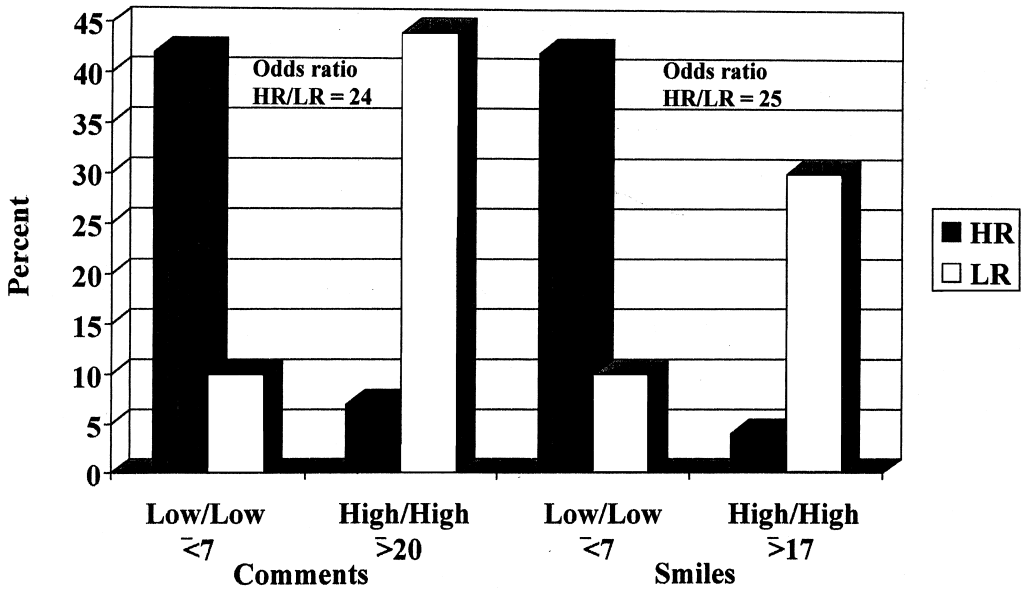


Figure 1. The percentage of high (HR) and low (LR) reactive children showing low versus high comments and smiles at 7 and 10 years.

when humans are cognitively active, or in a state of alerted arousal, higher EEG frequencies replace lower frequencies in the alpha range; that is, there is more power in the beta band (14–30 Hz) when subjects are cognitively active.

As expected, more high than low reactives, especially those who had been highly fearful in the second year, showed more right frontal asymmetry (see Figure 2), a larger wave 5 from the inferior colliculus (see Figure 3), and larger beta power values while the child sat quietly with eyes open. In addition, the children who had been highly subdued with the examiner (values for comments and smiles were in the lowest quartile) were more likely than others to show right frontal asymmetry (see Figure 4).

We then asked whether, among the subdued children, the two temperamental groups differed on these measures. Fifty percent of the subdued high reactives showed right frontal asymmetry, compared with only 23% of the low reactives; 38% of the subdued high reactives had a wave 5 larger than .90 μV compared with only 15% of the low reactives; and 67% of the subdued high reactives had beta power values greater than 3.0, compared with 46% of the low reactives ($p < .05$ by the exact test for each comparison; see Table 2).

Thus, among children whose external behavior was equally subdued, the addition of physiological data revealed that those who had been high reactive infants were biologically different from those who had been low reactive. Therefore, the two groups should be given different classifications. The subdued profile of the high reactives is a product of temperament; the subdued profile of the low reactives was probably acquired.

Eye Blink Startle

Another example of the utility of combining biological and behavioral information involved the magnitude of the child's eye blink startle to an auditory probe delivered when the child was watching a series of 27 pictures presented through goggles; nine pictures were pleasant in connotation, nine were neutral, and nine were unpleasant or aversive. Acoustic probes (95-dB white noise) during the presentation of each picture produced an eye blink reflex. We quantified the magnitude of the eye blink reflex to each acoustic probe as

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

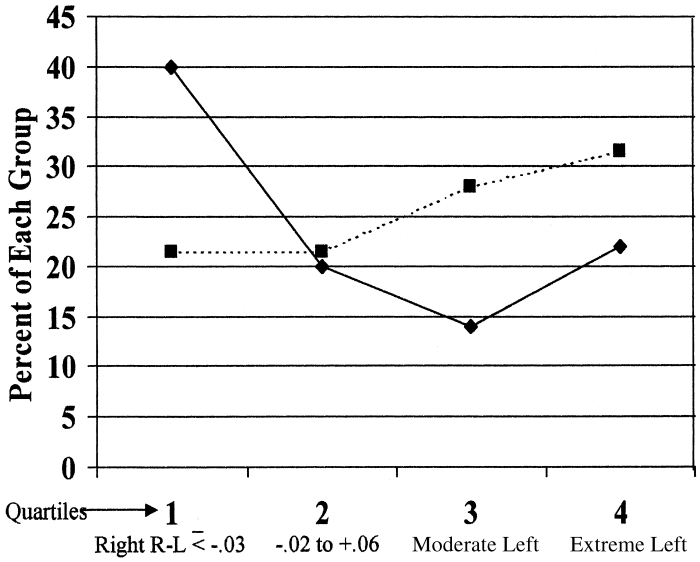


Figure 2. The quartiles of right frontal asymmetry distribution under eyes open for (◆) High and (■) low reactive children.

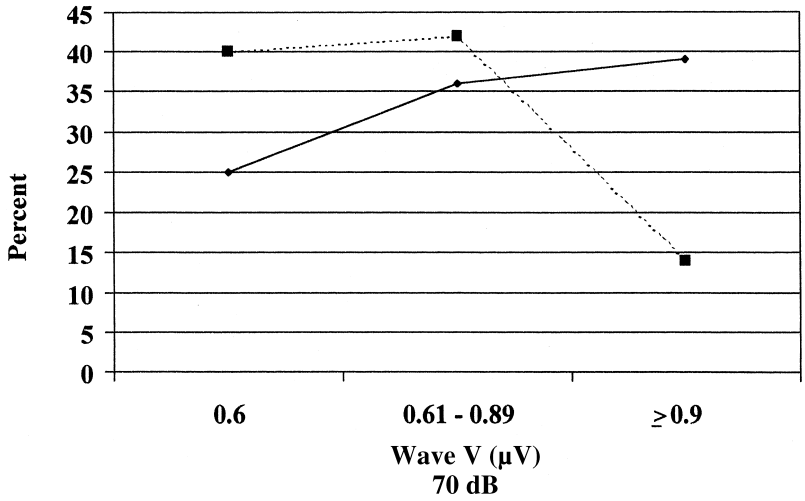


Figure 3. The percentage of (◆) high and (■) low reactive children showing low, moderate, or high values for wave 5 at 70 dB.

well as corrugator activity and heart rate to the pictures.

There was no significant difference between high and low reactive children in mean startle, corrugator activity, or heart rate when each variable was examined separately. However, only high reactive boys showed a significant positive correlation between magnitude of the startle and corrugator activity to the unpleasant pictures (r = .56 for the high reactives vs. .04 for the low reactives).

The high reactive girls showed a significant correlation between corrugator activity and the lowest heart rate displayed to the unpleasant pictures, but the low reactive girls did not (r = .41 for high reactives vs. .10 for the low reactives). Thus, by combining variables, in this case startle and corrugator for boys and corrugator and heart rate for girls, we distinguished between the high and low reactives. These data imply that combining measures can lead to fruitful conclusions. If

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

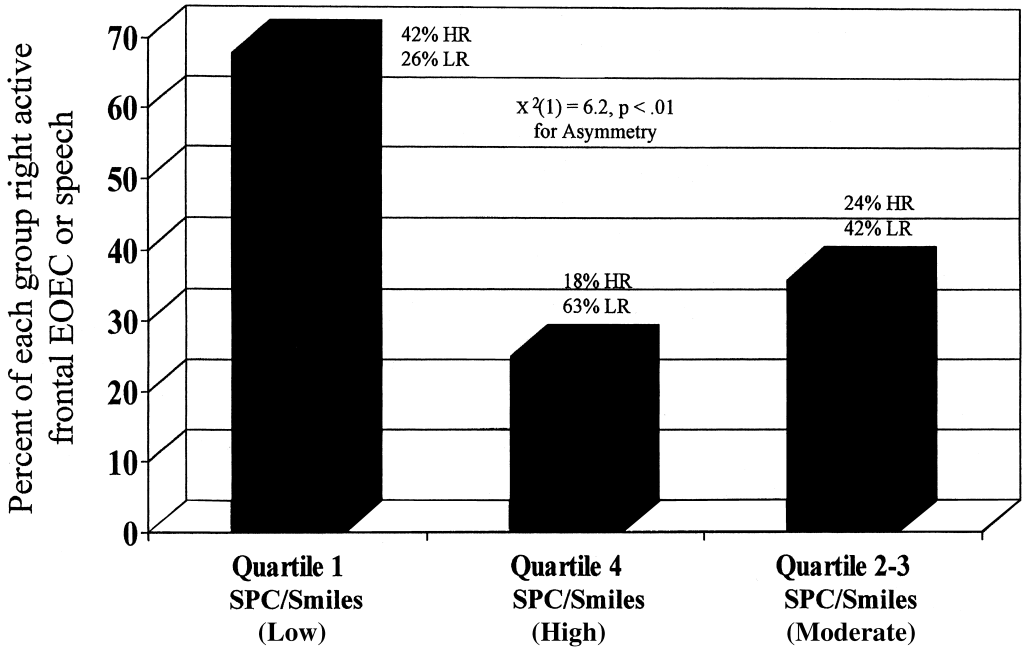


Figure 4. The percentage of spontaneous and subdued children showing right frontal activation under eyes open or eyes closed (EOEC) or speech preparation (SPC).

Table 2. *Percentage of subdued high and low reactives meeting criteria for three biological measures*

	High Reactive (N = 49)	Low Reactive (N = 65)
Percent subdued	37	20
Among subdued		
Right active	50	23
β power > 3.0	67	46
Wave 5 > 0.9 μV	38	15
Among remaining subjects		
Right active	21	19
β > 3.0	36	45
Wave 5 > 0.9	31	13

we had only coded startle magnitude, we would have concluded that high and low reactives are similar in their psychological reaction to the unpleasant pictures.

Mother and Child Q-Sorts

A third example of the utility of combining evidence comes from the Q-sorts of the 10- to 11-year-old children. During an interview at home, each child ranked 20 printed statements about his or her behavior from most to least characteristic of self. Six of the 20 items reflected an inhibited profile: "I get a little scared in new places; I worry too much about making mistakes; I am a little nervous when I sleep over at a friend's house; I am a shy person; I don't like to read in front of the class; If I have a question in the store, I like to have a parent ask the sales person." A total of 16 high reactives and 12 low reactives had a mean score across these six items that was in the lowest percentile of the distribution, reflecting a self-reported inhibited profile. The maternal Q-sorts of their child involved 28 traits and the mothers' descriptions were in

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

closer agreement with the self-reports of their high reactive children than those of the low reactives because a large number of low-reactive children described themselves as shy and nervous. However, their mothers denied this fact and the children's behavior in the laboratory did not suggest an inhibited style. Further, among the children who said they were inhibited, four times as many high as low reactives (38 vs. 8%) showed right frontal asymmetry. Thus, among children who rated themselves as equally shy, timid, and inhibited, additional data from the mother's Q-sort and EEG separated the high reactives from the low reactives. Once again, combining measurements reveals differences between groups of children who were similar on one variable.

Extreme Scores

Most relations between behavior and physiology are nonlinear; hence, it is important to examine subjects who have extreme scores. The relation between any two phenomena is rarely the same for all ranges of the variables in question. Consider as an illustration the relation between body weight, on the one hand, and health, popularity, or speed of running, on the other. If a scientist based her conclusions on the first 30 adult women who volunteered for a study and all the volunteers happened to weigh between 110 and 140 lbs, the investigator would find no relation between weight and any of the three characteristics. However, if 300 subjects were assessed, the sample would probably contain some very heavy and some very light women who should differ in their health, speed of running, and popularity.

A small group of boys in our longitudinal sample combined low reactivity and a low heart rate at 4 months, low fear and frequent smiling to the examiner at 2 years, and a cooler forehead temperature on the left compared with the right at 4.5 years. These boys were different from all other children because they were extremely exuberant and full of vitality. Most low reactive boys did not show this intensely ebullient profile, and no high reactive boy came close. The fact that less than 10% of this sample displayed this special combination of five characteristics has implications for current research practices. The vast majority of laboratory studies of children contain fewer than 50 unselected volunteers; laboratory studies of patients usually have smaller samples. If a theoretically important category is appropriate for less than 10% of the population, samples as small as 50 will contain too few subjects with the profile and the investigators will not detect them (Kagan, Snidman, & Arcus, 1998b).

A second example of the utility of examining extreme scores comes from a collaboration with Jerrold Rosenbaum and Joseph Biederman, two psychiatrists at Massachusetts General Hospital (Rosenbaum, Biederman, Hirshfeld–Becker, Kagan, Snidman, & Faraone, 2000). We evaluated 148 children whose parents were diagnosed as having panic disorder, panic disorder with depression, or no psychiatric condition. A group of 65, 4.5-year-olds and 83, 6.5-year-olds had a parent with panic disorder or panic with depression; 42, 4.5-year-olds and 36, 6.5-year-old children came from families, matched on ethnicity and social class, in which neither parent had anxiety or depressive disorder. All families were Caucasian and of middle-class status. Nineteen percent of children with a panic parent, but only 5% of controls, had scores in the lowest quartile of the distributions for both comments and smiles with the examiner, $\chi^2 = 6.6$, $p < .01$. However, the mean scores for smiles and comments did not differentiate the two groups.

Ten girls with a panic parent displayed fewer than 10 spontaneous comments and fewer than 10 smiles (these 10 girls represent 15% of all girls with a panic parent). These girls were qualitatively different from the remaining girls with a panic parent because they had a higher heart rate at rest, were more likely to have a very cool right versus left index finger, and were more likely to have very light blue eyes. These three features are characteristic of extremely inhibited children in our other longitudinal samples.

A third example of the value of extreme scores is the fact that a large temperature asymmetry between the pair of index fingers while the children were watching a series of

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

film clips differentiated the children with a panic parent from the controls, but the mean asymmetry values did not (Kagan, Snidman, & Peterson, 2000). That is, the relation between temperature asymmetry and diagnostic status of the parents only held for extreme asymmetry values. More than twice as many children with a panic parent, compared with controls, had temperature asymmetry values of ±0.5°C or greater between the pair of index fingers (±0.5 represented the top percentile of the distribution of asymmetry values).

In addition, each child in this study was rated reliably on a 4-point scale for degree of inhibited behavior during a 90-min interaction with the examiner. Although more children with panic parents were given ratings of 3 or 4, there was a nonlinear relation between this rating of inhibition and the probability of having a large temperature asymmetry on the pair of index fingers. Thirty-five percent of the children with extreme asymmetries were given a rating of 4, compared with 18, 24, and 20% for the children given ratings of 1, 2, and 3, respectively.

About 1 in 5 children in this study born to a parent who had panic disorder was likely to be inhibited in the laboratory, compared with 1 in 20 for the matched controls. If we add a large temperature asymmetry on the index fingers as a second diagnostic feature, then 1 in 10 children with a panic parent, versus 1 in 100 control children, combined an inhibited style of interaction with this index of sympathetic lability. If having a parent with panic disorder was regarded as a risk factor for anxiety, the difference between 1 in 10 and 1 in 100 would have clinical significance.

Discussion

The principal argument in this paper is that many currently popular constructs in personality and psychopathology are based on only one source of information; usually, semantic constructs derived from self-report questionnaires or interviews. Investigators who use these methods assume that constructs, like extroversion or anxiety disorder, are unitary and all members of a category are fundamentally similar. However, when scientists combine two different sources of data presumed to reflect the same or a similar construct, they often discover that individuals with different biologies have been assigned to the same category. The children who had been high reactive infants and subdued with an examiner at age 10 showed right frontal asymmetry in the EEG and a large wave 5 in the brain stem evoked potential. Low reactive children who were equally subdued with the examiner typically showed left frontal asymmetry and a small wave 5. The first group is temperamentally inhibited; the second group is merely quiet with an unfamiliar adult in a laboratory setting.

The features that should define a personality or psychopathology category are stable profiles that combine behaviors, beliefs, and moods in order to differentiate one group of individuals from another. The biological concept of species, which parses animals into groups with relatively stable profiles, might provide a useful guide to a more fruitful conception of personality and pathology. Biologists usually take into account three sources of information, not one, when they define a species. They accommodate to the animal's phylogenetic history, a distinctive set of inherited anatomical and physiological features, and reproductive isolation from related groups. This strategy may help theory in personality and psychopathology.

Although childhood history must be a component in some adult pathologies, some psychologists have ignored this information because they are not certain about the childhood events that are most relevant to an adult profile and they do not possess valid ways to measure the early experiences they believe are relevant in adults.

The child's temperamental biases, based on inherited biological processes, are analogous to the anatomical and physiological features that define a species or a strain within a species. The research on temperaments emphasizes variation in energy level, capacity for sustained attention, and the ease with which anger, sadness, fear, guilt, hope, and joy are provoked. High and low reactives represent two of the many temperamental types that will be discovered. These discoveries will

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

require gathering behavioral and psychophysiological evidence.

The third contribution to personality, which is analogous to the criterion of reproductive isolation, is the context of social relationships and associated experiences that maintain or change the profile developed during childhood and adolescence. For example, an anxious man marries a woman whose fiercely loyal love helps him mute his chronic anxiety or a confident man marries a woman whose continued infidelity forces him to question his adequacy. Similarly, the civil protests against sexism, racism, and the Vietnam War during the 1960s led many academically conscientious high school and college students living in cities like Boston, Chicago, New York, and San Francisco to reject the professional careers their parents had encouraged because such a choice symbolized conformity to establishment values.

The fact that profiles of asocial behavior and social phobia are overrepresented in individuals who grew up with less well-educated families implies that a personality category must acknowledge the individual's social settings (Halfon & Newacheck, 1999). The frequency of depression among east European adults increased after these countries were released from Soviet domination because the high expectations for improvement in their lives were not realized. High reactive infants raised in homes where school achievement is encouraged should be more likely to become conscientious students. High reactives raised in homes that do not encourage academic achievement should be more likely to become social phobics. High reactives raised by families that are extremely permissive of aggression and disobedience should be at higher risk for the development of impulsive or asocial behavior when frustrated.

The invention of more valid personality or pathology categories will require the discovery of theoretically fruitful ways to conceptualize and measure the psychological products of childhood histories, early temperaments, and the contextual factors that maintain or alter the profiles acquired during earlier stages of development. These suggestions can be applied to the current diagnosis of social phobia. The bases for the classification are verbal self-reports of psychological distress when in a large gathering of strangers, reluctance to be seen in public places, and an unhappiness with these symptoms. If the last feature is absent, the patient would not be diagnosed as a social phobic. This configuration could be the result of temperament (a child born with a high reactive temperament) or a variety of environmental histories, including social rejection, abuse, moving to a city in which one is an obvious minority, or embarrassment over a cleft palate. Thus, it is necessary to gather additional information if one wishes to discover the correct category to which the individual should be assigned.

Physics supplies exquisite examples of how additional data change the meaning of important constructs. Before the discovery of microwave radiation in the radio spectrum from the cosmos and the subsequent invention of the concept of dark matter, astronomers thought that most of the mass in the universe came from stars and planets. They now believe that most of the matter in the universe is invisible and that the planets and stars comprise a very small proportion of all cosmic matter.

Chemistry made advances when it added information on the spectra that each heated element produced. Biologists attained deep insights when they added genetic information to descriptions of anatomy and physiology. Neuroscientists required electron microscopes to affirm the reality of the synapse. It is now time for psychology and psychiatry to enjoy their well-earned period of progress. In order to attain that goal, investigators must add behavioral, cognitive, and physiological data to the semantic self-reports that now dominate the methodologies used in most studies.

Alfred North Whitehead noted in *The making of the modern mind* (1925) that the physicists who invented quantum mechanics were not more intelligent than Boyle or Becquerel; they simply had the help of an apparatus that revealed clues that nature had withheld from 18th and 19th century scientists. A new source of information, Whitehead noted, was

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press

like visiting a new country for the first time. The fresh observations alter understanding in dramatic ways. The addition of biological data to verbal descriptions of behavior will reveal insights as profound as those that follow the study of the energy profiles produced by bombarding atoms with high energy particles.

References

Antoniadas, E. A., & McDonald, R. J. (2000). Amygdala, hippocampus, and discriminative fear conditioning to context. *Behavioural Brain Research, 108,* 1–9.

Bacon, F. (1878). *Novum organum* (T. Fowler, Trans.) Oxford: Oxford Univeristy Press. (Original work published 1605.)

Brandao, M. L., Cardoso, S. H., Melo, L. L., Motta, V., & Coimbra, N. C. (1994). Neural substrate of defensive behavior in the midbrain tectum. *Neuroscience and Biobehavior Reviews, 18,* 339–346.

Davidson, R. J. (1995). Cerebral asymmetry, emotion, and affective style. In R. J. Davidson & K. Hugdhal, (Eds.), *Brain asymmetry* (pp. 361–388). Cambridge, MA: MIT Press.

Davis, M., & Astrachan, B. (1978). Conditioned fear and startle magnitude. *Journal of Experimental Psychology, 4,* 95–103.

Denenberg, V. H. (1964). Critical periods, stimulus input, and emotional reactivity. *Psychological Review, 71,* 335–351.

Fox, N. A., Calkins, S. D., & Bell, M. A. (1994). Neuroplasticity and development in the first two years of life. *Development and Psychopathology, 6,* 677–696.

Halfon, N., & Newacheck, P. W. (1999). Prevalence and impact of parent-report describing mental health conditions among U.S. children. *Journal of American Academy of Child and Adolescent Psychiatry, 38,* 600–609.

Kagan, J. (1994). *Galen's prophecy*. New York: Basic Books.

Kagan, J. (1998). *Three seductive ideas*. Cambridge, MA: Harvard University Press.

Kagan, J., Snidman, N., & Arcus, D. (1998a). Childhood derivatives of high and low reactivity in infancy. *Child Development, 69,* 1483–1493.

Kagan, J., Snidman, N., & Arcus, D. (1998b). The value of extreme groups. In R. B. Cairns, L. R. Bergman, & J. Kagan (Eds.), *Methods and models for studying the individual* (pp. 65–80). Thousand Oaks, CA: Sage.

Kagan, J., Snidman, N., & Peterson, E. (2000). Temperature asymmetry and behavior. *Developmental Psychobiology, 36,* 1–8.

Kagan, J., Snidman, N., Zentner, M., & Peterson, E. (1999). Infant temperament and anxious symptoms in school age children. *Development and Psychopathology, 11,* 209–224.

Klein, R. G. (1991). Parent–child agreement in clinical assessment of anxiety and other psychopathology. *Journal of Anxiety Disorders, 5,* 182–198.

LeDoux, J. E. (1996). *The emotional brain*. New York: Simon & Shuster.

Perrin, S., & Last, C. G. (1992). Do childhood anxiety measures measure anxiety? *Journal of Abnormal Child Psychology, 20,* 567–578.

Ray, W. J., & Cole, H. C. (1985). EEG alpha activity reflects attentional demands and beta activity reflects emotional and cognitive processes. *Science, 228,* 750–752.

Rosicky, J. (1993, April). *The assessment of temperamental fearfulness in infancy.* Paper presented at the Meeting of the Society for Research in Child Development, New Orleans, LA.

Rosenbaum, J. F., Biederman, J., Hirshfeld–Becker, D. R., Kagan, J., Snidman, N., & Faraone, S. (2000). A controlled study of behavioral inhibition in children of parents with panic disorder and depression. *American Journal of Psychiatry, 157,* 2002–2010.

Seifer, R., Sameroff, A. J., Barrett, L. C., & Krafchuk, E. (1994). Infant temperament measured by multiple observations and mother reports. *Child Development, 65,* 1478–1490.

Spiker, D., Kramer, H. C., Constantine, N. A., & Bryant, D. (1992). Reliability and validity of behavior problem checklists as measures of stable traits in low birth weight premature preschoolers. *Child Development, 63,* 1481–1496.

Treit, D., Pesold, C., & Rotzinger, S. (1993). Dissociating the anti-fear effects of septal and amygdaloid lesions using two pharmacologically validated models of rat anxiety. *Behavioral Neuroscience, 107,* 770–785.

Walker, D. L., & Davis, M. (1997). Involvement of the dorsal periaqueductal gray in the loss of fear potentiated startle accompanied by high foot shock training. *Behavioral Neuroscience, 111,* 692–702.

Whitehead, A. N. (1925). *Science and the modern world*. New York: Macmillan.

Woodward, S. A., McManis, M. H., Kagan, J., Deldin, P., Snidman, N., Lewis, M., & Kahn, V. (2001). Infant temperament and the brain stem activity evoked response in later childhood. *Developmental Psychology, 37,* 533–538.

https://doi.org/10.1017/S0954579402003048 Published online by Cambridge University Press