Critique of the Act Frequency Approach to Personality

Jack Block
University of California, Berkeley

Evaluates the act frequency approach (AFA), noting that retrospective self-reports rather than behavioral acts are studied; act context and meaning are not considered; the AFA self-report inventories are incompletely developed and are psychometrically unsound; the AFA claim of absolute measurement of dispositions is insupportable; many of the self-report act statements used are technically unacceptable or conceptually unwarranted; the research agenda of the AFA primarily involves only "internal analyses" of self-report "acts" inventories and indices and proposes the further creation of "act" inventories to index thousands of conceptually unordered dispositions. Were the problems involved in the current implementation of the AFA to be resolved, various large conceptual problems presently besetting the approach then would have to be confronted.

The act frequency approach (AFA; e.g., Buss & Craik, 1980, 1981, 1983a, 1983b, 1983c) has been presented as a new, needed, widely applicable, and generative way of studying personality. For example, it has been characterized by Buss and Craik (1984) as providing "a fresh look" (p. 283), "a distinctive perspective" (p. 283), "a novel level of analyses" (p. 257), "an exciting research agenda" (p. 291), and as an instance of the 1980s "resurgence in the field of personality psychology (p. 242). Vigorously promulgated by its initiators as having multidirectional relevance (Buss, 1984, 1985a; Buss & Craik, 1986, 1987), the AFA is now represented by a sizable body of literature that continues to grow rapidly. Au courant personality psychologists have become aware of the AFA, and it has begun to be cited with some frequency and approval (e.g., Pervin, 1985).

However, some psychologists, I among them, do not view the AFA as being especially helpful to the understanding of personality processes, personality development, and personality differences, or as offering technical or conceptual improvements in the necessary field of personality assessment. Because only articles propounding the AFA have appeared thus far, now may be a time for presentation of an alternative view of the approach. A dialogue on the AFA can only benefit the field of personality psychology because important themes and issues are involved.

In this article I first discuss the essential features of the AFA. I then call attention to a long-used and effective principle of personality assessment in order to provide a larger perspective on a central emphasis of the AFA. Finally, I present some of my concerns regarding the AFA.

Essential Features—Conceptual, Methodological, and Empirical—Underlying the AFA

I begin by describing denotatively, and I believe fairly, the first article presenting the AFA (Buss & Craik, 1980). This article, on the disposition of dominance, announced a conceptual orientation and established a research paradigm that the AFA repeatedly has used since. Although I consider other AFA publications, I suggest that it is not necessary to closely consider all of the subsequent AFA work in scrupulous detail because, as will be seen, the later articles repeat or build in fundamental ways on the initial conceptualization and the initial research formula, although some changes have been introduced. Consequently, sufficient criticism or questioning of the AFA, as initially presented, is also a sufficient argument against the later usages of the approach. I attempt to keep separate the strict description of the AFA and my evaluations of the approach.

The conceptual introduction to the AFA in the initial 1980 article is terse, only a few paragraphs long. Citing Hampshire's (1953) discussion of dispositions, itself brief, the "summary" view of dispositions is essentially endorsed. This view asserts that a personality attribution or a dispositional statement about an individual depends on the relative incidence of occurrence of specified acts over a designated period of observation. From the standpoint of this "frequency concept of disposition," the statement "Mary is dominant" is said to mean that, over a specified interval (e.g., 3 months), Mary has exhibited many dominant acts and therefore can be expected to exhibit many dominant acts in the future.

Although Buss and Craik characterized acts as "the basic constituent elements of the behavioral world" (1984, p. 245), they do not discuss the nature of acts and how, conceptually, they are to be delineated. Rather, the act is viewed here and explicitly characterized in subsequent AFA articles as "a primitive term requiring further explication as the research agenda for this approach to personality unfolds" (Buss & Craik, 1984, p. 247). Consequently, the nature of an "act" within the AFA

This study was supported by National Institute of Mental Health Grant MH 16080 to Jack Block. Over the last decade, David Buss, Kenneth Craik, and I have discussed various aspects of their approach. I have found them to be unfailingly cooperative with my efforts to understand their research program.

Earlier versions of this article benefited from the comments of Alois Angleitner, Oliver John, Randy Larsen, Daniel Ozer, Avril Thorne, and the journal reviewers.

Correspondence concerning this article should be addressed to Jack Block, Department of Psychology, Tolman Hall, University of California, Berkeley, California 94720.
can only be understood ostensively, by the way the approach in fact proceeds to operationalize "acts." I note, however, that there is explicit turning away by the AFA in this and later articles (Buss & Craik, 1984, p. 243) from propositional, "if-then" conceptions of dispositions, such as Ryle's (1949), that invoke (and require) recognition of the situation or context in which an act occurs before a dispositional assertion can be made.

Having adopted the act frequency concept of disposition, Buss and Craik (1980) then were interested in identifying acts that are considered to be manifestations of common dispositions, such as "dominance." Some behaviors are construable as instances of dominance and some are not. Having established a set of "acts," each of which may be said to fall within the specified dispositional category (i.e., dominance), Buss and Craik (1980) then sought to order or scale these "dominant acts" with respect to their degree of relevance or appositeness as a manifestation of the dispositional category of "dominance" behaviors. For this scaling task, Buss and Craik (1980) opted for the prototypicality rating procedure developed by Rosch (1975; Rosch & Mervis, 1975). Within the Roschian terminological framework, semantic categories are said to have an "internal structure" when

subjects consider it a meaningful task to rate members of [a category] according to how well they fit the subjects' idea or image of the meaning of the category name and . . . there is high [italics added] agreement between subjects concerning these rankings. (Rosch, 1975, p. 198)

Thus, when the AFA talks in Roschian terms of the internal structure of the dispositional category of dominance, what is meant is that there are better and poorer act examples of what people consensually mean by the concept of dominance.

With the expectation that it would prove feasible to order a set of "act" instances with regard to their relevance for the dispositional concept of dominance or, alternatively put, their "centrality of membership" within the dominance category, Buss and Craik (1980) then proposed to evaluate two primary hypotheses: (a) In replication of a study by Jaccard (1974), single "act" instances of dominance were expected, on the average, to correlate less well than would the sum of many single "act" instances of dominance with established, conventional personality inventory measures of dominance; and (b) responses to a set of "act" instances judged to be relatively pertinent or central to the dispositional category of dominance were expected to correlate higher with established, conventional personality inventory measures of dominance than would the responses to a set of "act" instances judged as relatively peripheral to the disposition of dominance.

To test for the superiority of a "multiple-act criterion" over a "single-act criterion" and to demonstrate the existence of the "internal structure" of a dispositional category, two preparatory efforts were required. First, following the procedure initially used by Jaccard (1974), "act nominations" were sought from undergraduates. Each of 75 students was asked to write down 10 acts or behaviors that dominant people they knew "might [sic] perform that would reflect or exemplify their dominance" (Buss & Craik, 1980, p. 381). These 750 written statements describing hypothetical "acts" were then, to an unspecified extent, "supplemented by acts gleaned from a perusal of [conventional] dominance scales" (p. 382). From the 750 "act" statements plus the supplementations from extant dominance scales, the researchers selected or formulated 100 "act" statements for subsequent use. Two of these "act" statements proved to be slightly modified items from the California Psychological Inventory (CPI) Dominance (Do) scale of Gough (1964). The 652 (87%) nominated "act" statements excluded by the AFA researchers were considered by the investigators to be either redundant, statements about "nonacts," or statements regarding behavioral tendencies rather than statements about specific "acts." Each "act" statement was then phrased in two versions, with a man as the actor and again with a woman as the actress. Thus, the dominant "act" statement "He demanded a backrub" was paralleled by the statement "She demanded a backrub."

The second preliminary step involved presentation of the selected "act" statements to another set of undergraduates and to a set of personality psychologists, asking the participants to rate each "act" statement on the extent to which it represented a prototypically good example of the concept of dominance. A modification of the instructions by Rosch and Mervis (1975) was used. For each panel of judges, the average prototypical rating was calculated for each "act" statement. The 100 prototypicality ratings, when averaged over the set of judges, proved to be reliable, with the alpha coefficients being uniformly high (from .82 to .89) for each panel of judges. From these figures, Buss and Craik (1980) concluded, incorrectly, that "considerable agreement exists among [italics added] members of each panel regarding which acts are prototypically dominant." (p. 383). Their conclusion was wrong because they did not recognize that the reliability or reproducibility of a set of average scores (the 100 mean prototypical ratings) is always higher, and generally much higher, than the coefficient found when the prototypicality ratings of each judge are correlated with the prototypicality ratings of every other judge, with the average of these correlations then being taken. In the present instance, although the average intercorrelation between prototyping judges was not presented, it was readily estimated by reversing the Spearman-Brown formula. The average correlation among judges could be estimated to be about .20, a figure promptly verified by direct calculation (D. M. Buss, personal communication, July 17, 1980). That is, there was little agreement among the judges as to which "act" statements were prototypically dominant. The high agreement between judges implied by the Rosch (1975) definition of the internal structure of a category did not in fact exist.¹ Nevertheless, because so many raters were involved, the averaging process across raters built up cultural-

¹ Indeed, high and stable agreement among judges in their prototypicality ratings may not exist even among concepts of the kind studied by Rosch (Barsalou, 1987; Chaplin, John, & Goldberg, 1988). Subsequent articles on the act frequency approach do not make this early interpretable error of confusing the average correlation among judges with the correlation between two sets of averages. It should also be noted that the degree of agreement among judges is a manipulable rather than an intrinsic function of the set of "act" statements presented for prototypicality judgments. When these range widely in degree of prototypicality, interjudge agreement will be relatively high; when the prototypicality range is limited, interjudge agreement will be relatively low.
norm indices that broadly characterize the relevance or centrality of dominance-related "act" statements even though individual judges correlated only slightly with each other.

A new set of undergraduates was then administered the CPI Dominance scale, the Dominance scale from the Personality Research Form (PRF; Jackson, 1967), and a form for the self-rating of dominance. A week later these subjects were administered the Act Report, a first-person version of the 100 "act" statements previously chosen by the investigators and rated for prototypicality. Thus, the "act" statement "He demanded a backrub" became "I demanded a backrub." Subjects indicated whether they had, during their entire lives, ever performed the "act" described in the statement.

One emphasized result was that the average correlation of individual "act" statements with scores on CPI and PRF Dominance scales with the self-ratings was low, with the averages ranging from .10 to .20. Not emphasized was the finding that individual "act" statements correlated negatively with the dominance scales and self-ratings (for the CPI, as negatively as -.26 for men and -.38 for women; for the PRF, as negatively as -.19 for men and -.26 for women; for the self-ratings, as negatively as -.26 for men and -.38 for women). These negative correlations, and others that can be inferred to be present, imply that a number of the "act" statements supposedly manifesting the disposition of dominance should have been reversed in order to relate to dominance as indexed by extant inventory scales and self-ratings. The Jaccard (1974) study, which Buss and Craik (1980) were approximately replicating, reported no negative correlations of "act" statements with the CPI or PRF Do scales.

The second emphasized result was that the number of yes responses to a subset of act statements that previously had been rated as relatively relevant or prototypical with respect to dominance tended to correlate higher (albeit not highly) with CPI and PRF Dominance scale scores than did the number of yes responses to a subset of "act" statements rated earlier as being relatively peripheral to the concept of dominance. Contrasting the most prototypical item set with the least prototypical item set, for the male and female samples, respectively, the correlations with the CPI were .48 versus .14 and .31 versus .05; with the PRF the comparable correlation pairs were .67 versus .33 and .47 versus .26. This pattern did not hold for the self-ratings of dominance. I note that, typically, the correlation between the CPI and PRF Dominance scales is in the mid-70s; for example, Gough (1987) reported correlations of .75 and .78 and Jackson (1967) reported .78.

The pattern holding for the CPI and PRF Do scales was viewed by Buss and Craik (1980) as supporting two conjoined hypotheses: (a) A summation of responses to their "act" statements (which they call "a multiple-act criterion" or an "act trend") would, on the average, generate higher correlations with related measures than would, on the average, single responses (termed "specific acts"); and (b) a subset of "act" items judged to be relatively relevant or prototypical in assessing a disposition would generate higher correlations with related measures than would a subset of "act" statements judged to be relatively distant from or unrepresentative of the disposition being assessed.

The discussion of these results viewed the administration of "act" statements to subjects as providing a superior alternative to traditional questionnaire or inventory approaches. In particular, Buss and Craik (1980) declared that "multiple-act criteria [the number of yes responses to the act statements] constitute the appropriate [italics in the original] criteria in personality research from the perspective of a summary or frequency interpretations of the concept of personal dispositions" (p. 390).

Because they considered the scores on their "act" statement scales to be "criterion" scores, Buss and Craik (1980) contended that other approaches to personality should seek to predict these "multiple-act criteria." Thus, for example, they did not talk simply of the correlations between their dominance "multiple-act criteria" and the CPI and PRF Do scales. Rather, they phrased their findings in terms of how well the CPI and PRF Do scales "predicted" scores on the "act" statement scales. Throughout their many AFA articles, Buss and Craik (1980) have been consistent and emphatic in phrasing relations in terms of how well other kinds of measures "predict" their "multiple-act criteria" or "act trend indices.

Subsequent to this first AFA article on dominance, the AFA has gone on to the construction of several dozen more "act reports," to analyses of the relations between "act reports" and various self-report personality measures, to the study of various "act" indices, to a proposal that "act trend indices" achieve absolute measurement of dispositions, and to elaboration of the implications of the AFA for personality psychology and psychiatry. I discuss various aspects of these AFA articles later in this article but, as will be seen, the conceptualizing and empirical procedures underlying the first AFA article have continued to be fundamental to the approach.

A Basic Principle of Personality Assessment and Inventory Scale Construction

Before embarking on specific evaluation of the AFA, it is useful, as a reminder, to call attention to a basic, conventional, even archaic but also effective principle of personality assessment and the psychometrics of scale construction. There is a rich literature on these matters (cf. Jackson & Paunonen, 1985; Lovinger, 1957; Lord, 1955; Nunnally, 1978; Wiggins, 1973).

The assessment of personality generally depends on the availability or generation of a host of characterological indicators, each of which by itself may be equivocal. A crucial assumption, which generally is subsequently supported, is that although each indicator may be ambiguous in implication, the sum or some kind of scaling of these individually uncertain indicators has an implication that is relatively clear. How does one tell that something is a duck? One indicator is webbed feet, another is flying north in the spring and south in the fall, a third is that water rolls off its back, a fourth is that it has a platypus-shaped bill, a fifth attribute is that it quacks, a sixth is that it waddles when it walks, and so forth. Each of these indicators is equivocal; other living entities also manifest each of these characteristics. Frogs also have webbed feet, geese also are seasonal flyers, wearers of slickers also find that water rolls off their backs, the platypus also has a similarly shaped bill, my grandson has a quacking bathtub toy, I have been known to waddle when I walk, and so forth. However, when one sums up these characteristics for each of a variety of entities, which gets the highest
score with respect to the set of "duckiness" attributes or indicators? The duck does, and by far.

This strategy has worked reasonably well in constructing measures of personality, measures of intelligence, life data indices, and in a host of behavioral and experimental applications. Literally thousands of instances of the usefulness of the summing or compositing principle could be cited. There is a long psychometric basis for this approach, going back to Spearman (1910) but also involving, at a more conceptual level, the domain sampling approaches of Thomson (1935) and Tryon (1935), the recent important extensions of this principle by Epstein (1979), and the work of generations of psychometricians. The sense of this recognition goes back centuries, at least to the times of the pharaoh's land surveyors. In personality assessment, one does not depend on the validity of the response to any single item; rather, the sum or average or cumulative sense over a number of conceptually linked items is what researchers generally must (and generally can) depend on to attain a higher reliability and the consequent possibility, often realized, of a higher validity (cf. Horowitz, Inouye, & Siegelman, 1979). This general principle underlies the construction of personality inventory scales, sets of phenotypically diverse but psychologically related items that index various personality dimensions, vocational interests, or psychiatric conditions.

Concerns and Questions Regarding the AFA

1. **The label, the act frequency approach, is a misnomer: The AFA does not study the "acts" of people.** Although the AFA frequently indicates explicitly as well as implicitly that it is based on direct behavioral observation, on "what people do all day" (Buss & Craik, 1983c, p. 119), as of this writing, the AFA has not yet directly observed the actual "acts" of individuals. This point is basic, incontestable, and I believe casts a long shadow on all claims made for the AFA.

Instead of on-line behavioral observation, "act" data are said to be generated when subjects retrospect about the applicability to them of each of 100 positively phrased statements contained in a booklet. The sum of the affirming responses offered by a subject to the statements within an act report is then explicitly posited by the AFA as the "multiple-act criterion" or "act trend index" and is vigorously represented by the AFA as a conceptually superior and "heretofore neglected validational criterion[on] for personality scales" (Buss & Craik, 1983b, p. 406) "grounded in everyday conduct [that] acquires prior status." (Buss & Craik, 1984, p. 274). However, the indisputable fact remains that nowhere have acts been directly observed in the AFA.

2. **Acquiescence to an "act" statement is accepted without question within the AFA as an instance of a single "act"; it is assumed that the written endorsement of a statement stands in one-to-one correspondence with the past event portrayed by the booklet statement. However, there are many memorial problems associated with the use of questions or items inquiring of subjects about events during designated but long periods of their past, as the AFA requires (Bernard, Killworth, Kronenfeld, & Sailer, 1984; Sikkel & Jelierse, 1987; Som, 1973; Sudman & Bradburn, 1974). Did I demand a back rub 12 weeks (less than 3 months) ago or was it 14 weeks (more than 3 months) ago? Furthermore, when recall fails or is inexact, individuals bring into play, in diverse ways, various inferential strategies in order to generate a reasonable if not quite accurate answer (Bradburn, Rips, & Shevell, 1987). Although informant inaccuracy is widely recognized as presenting major and vexing difficulties when retrospectives about factual events are used, the AFA has essentially ignored this issue.

In the first AFA article on dominance (Buss & Craik, 1980), it was recognized that the approach relied on written responses to "verbal presentation of acts rather than directly observed acts. One evident next step in this line of research calls for field studies . . . of acts observed in situ . . . . Such naturalistic field studies . . . . are urgently needed in personality research" (p. 390). However, 10 years later, the AFA has remained wedded to the "retrospective self-recording of act performance [RSRAP] within a designated time interval" (Buss & Craik, 1984, p. 249), as revealed through the responses of subjects to booklets of statements about "acts." The naturalistic observation of acts in situ, noted as "urgently needed" in the first AFA article, still remains a task, certainly arduous and apparently unscheduled, for the future (Buss & Craik, 1983b, p. 407). I submit that many of the AFA statements and interpretations take on a different and modest meaning once it is recognized that written retrospective responses of uncertain validity to an inventory of "act" statements rather than observable behaviors are involved.

3. **The AFA has adopted a conceptual position that reiterates (and in effect celebrates) an ancient, well-used and not ineffective orientation to personality in which "oil personality psychologists nevertheless wish to advance beyond.** Personality psychologists have long recognized the necessity of conceptualizing behavior as a serious function of both the qualities of the person and the nature of the situation or context or environmental surround. There have been various efforts in personality psychology to move, conceptually and empirically, in this personality-in-situ direction (e.g., Arsenian & Arsenian, 1948; Block & Block, 1981; Chein, 1954; Murray, 1938), none of them as yet truly satisfying. However, the need to go toward this goal (classifying or scaling both individuals and situations) is, I suggest, widely recognized and accepted even if it remains far from being implemented.

In preferring the "summary" view of dispositional, the AFA is reasserting a long-used conceptual position that is formally unconcerned with the context that surrounds an "act." "This assessment approach sums displays of prototypical acts without regard to situational analysis or to attributions of causality to person, role, situation, or other factors. It remains strictly descriptive rather than explanatory" (Buss & Craik, 1983b, p. 249).
The following is another instance of the Buss and Craik (1983b) position regarding context:

The act frequency approach avoids the complexities of . . . inferred opportunity structures [previously defined as the alternative possibilities for action that occur to the person, or to an observer of the person, in each situation] by restricting itself to frequency tallies, without considering presumed action possibilities or counting acts that might be ascribed to social role. (p. 290)

As still another instance, consider the AFA rejection (Buss & Craik, 1980, pp. 389–390; 1984, pp. 242–244) of conceptions of disposition that involve motives, needs, beliefs, and desires that energize, direct, and select behaviors as a function of the individual's values, capacities, and situational opportunities.

Thus, if I affirm the context-unspecifying statement "I demanded a backrub" (a dominance "act" statement), the AFA is not interested in and cannot know the meaning of my yes response. Did I pay for a backrub at a massage parlor that I then did not receive? Did I in a jocular but loving vein "demand" a backrub from my lover? Did I, at long last, having given hundreds of backrubs to others, finally "demand" one for myself?

Currently, in many fields of psychology and the behavioral sciences, there are intensive efforts underway (e.g., Rommetveit, 1987) to incorporate context and meaning into conceptual formulations and empirical research. The conceptual position of the AFA, by explicitly adopting the "summary" or frequency concept of disposition view, excludes both the internal motivational context influencing the actor and the external situational context as it registers on the actor.

As the preceding, but not unique, quotations indicate, the AFA crisply rejects analysis of the situation in which an act occurs and chooses to not recognize the motivations underlying "acts." It is therefore intellectually distressing when elsewhere, and as categorically, it is declared, "acts, within the act frequency approach, are interpreted acts in that they involve elements of context, style, and inferences about intentionality" (Buss & Craik, 1985, p. 399). I submit that an approach that both rejects and embraces situational analysis and inferences about the intentions of the actor may be all-encompassing but is difficult to live with.

3. To date, the methodology used by the AFA in developing RSRAps scored to represent "the multiple-act criterion" or an "act trend" is nothing more than, and provides appreciably less than, traditional procedures in personality assessment for developing inventory scales. Setting aside for the moment the issue of the content of the "act" statements, I suggest that the methodology underlying the AFA is no more than a conventional approach because its procedures are fully equivalent to the initial procedures for developing "rational" inventory scales. The methodology underlying the AFA provides appreciably less in the way of a measuring procedure because, unlike scale development in personality assessment, it does not go on to use a variety of logics and procedures and a sequential revision process through which an initial scale becomes a dependable measure that has both convergent and discriminant validity. Instead, the AFA, after formulating a set of 100 "act" statements, posits that set (or subset) of "act" items as a criterion and does not seriously analyze RSRAps so as to eliminate the "act" statements that are ineffective or invalid discriminators.

To facilitate discernment of the correspondence between traditional procedures of personality assessment and those presented by the AFA as "novel," "fresh," "distinctive," and "alternative," I recount the sequence of procedures involved in the construction of rational scales, calling attention to the transliterated equivalents of these procedures within the AFA. I rely heavily on the excellent account of Jackson (1971).

(a) "An item pool may be viewed as a sample of content; it should be evaluated in terms of how adequately and how representatively it reflects the content domain or universe of content implied by the definition. . . . At the initial stages of personality scale development, one is not likely to have items in sufficient quantity and quality with which to work. It is necessary, therefore, to define what one is seeking to measure and to develop an item pool" (Jackson, 1971, p. 237). How does one generate an initial item pool? One personally writes items deemed pertinent, asks colleagues to write items they believe to be relevant, or even asks students in an introductory psychology course to write items they believe to be related to a defined dimension (e.g., Ashton & Goldberg, 1973; Jaccard, 1974; Jackson, 1975). I submit that the garnering of items from available undergraduates or laypeople is the full equivalent of what AFA labels "the act nomination procedure."

(b) The collection of possibly relevant items is only a beginning. According to Jackson (1971), "the . . . question of the adequacy of item sampling remains. Short of actually sampling the behaviors of very large numbers of individuals in their natural state . . . there is probably no way to judge the adequacy of an item pool and of a definition of a trait except by the use of rational processes, either of an individual psychologist or of a group of judges" (p. 237).

Rational judgments of the adequacy of items involve two steps. First, there needs to be some simple editorial work by a judge or two to ensure that items are in the verbal format desired and to weed out unquestionable or likely redundancies, and so on. Ordinarily, this is a quick, mostly clerical task and relatively few items are excluded as inappropriate at this juncture.

The AFA also first applies an editing function but, as applied, an unusually large number of the items generated by the "act nomination" procedure typically are excluded from subsequent consideration because they are judged by the AFA investigators to be redundant or to be statements about "nonacts" or statements about behavioral tendencies. Thus, for example, 87% (652 of 750) of the "acts" nominated by undergraduates as possibly exemplifying the disposition of dominance were excluded by the editing process. After the AFA editing to achieve 100 "act" statements deemed appropriate, these immediately become the RSRAP used without further change in subsequent AFA research.

Recall that the "act" nomination procedure is presented by the AFA as a means of gathering a wide and representative array of "acts" that are manifestations of a given disposition, as an approach to provide "rich and varied instantiation" (Buss & Craik, 1985, p. 944) of the nature of dispositional concepts, and as a means of identifying the behavioral phenomena that are subsumed by or generate the need for the dispositional category. The AFA rejection of so many of the "acts" proposed by their "act" nominators operates against this AFA intention and raises
questions regarding the bases for exclusion of nominated "acts." I believe that redundancies alone would not cause the exclusion of so many of the proposed items. Rather, my observation has been that subject-generated "act" statements frequently involve direct trait descriptors or are so laden with interpretation or involve so prolonged a time unit or behavior as to be deemed inappropriate within the AFA for consideration as an "act."

In any event, it is clear that various unspecified, unacknowledged, domain-restricting, perhaps unwitting and certainly crucial considerations are involved when Buss and Craik fix on the 100 "act" statements then used in their research on a particular disposition. For example, how did two items from the CPI find their way into the RSRAP for dominance? Although the AFA has emphasized the openness and broad base provided by the "act nomination process," I suggest that this emphasis has been misleading; the closed and unspecified "act selection process" has been at least as important in determining the crucial content of RSRAPS. The AFA has been laconic and uninformative regarding the process of "act selection," so fundamental when the ratio of the number of "acts" nominated to the number of "acts" chosen is so high and when all subsequent AFA data and relations are highly dependent on the particular set of "act" statements that just happen to have been used within the particular RSRAP. In the construction of traditional inventory scales, it is certainly the case that, as with the AFA, the items finally constituting a scale are selected by the scale constructor in ways that are not fully specified or specifiable. It is generally impossible, understandably so, to fully trace the many considerations, often conceptual and contextual rather than statistical, that influence final choice of the items finally used in an inventory scale. Within the AFA, therefore, it would be fairer if the inevitable, even acceptable subjectivity in the selection or construction of the "act" statements used were to be acknowledged and the emphasis on the "act nomination process" as implying extensive domain sampling were diminished.

A second step in traditional scale construction, after the editing process, requires psychological judgments, item by item, of the match between the content of the proposed item and the personality concept one is seeking to index. Generally, the pooled judgments of a number of persons are used to gain consensus and generalizability on this matter, to rise above idiosyncratic viewpoints (e.g., Ashton & Goldberg, 1973; Jackson, 1970). Thus, the proposed items are rationally placed on a gradient according to their judged relevance as indicators of the concept of interest. Those items judged to be irrelevant or largely so may be discarded at this juncture in the interest of economy; those items deemed to have potential are tried out empirically and are subsequently retained or discarded depending on their empirically attained reliability and validity. As Jackson (1971, 1975), Ashton and Goldberg (1973), and others have shown, rational judgments of the anticipated relevance of inventory items have proven to have appreciable validity; Pooled judgments can indeed identify reasonably well the items or subsets of items that will later prove to be valid.

I submit that the prototypical rating procedure to get at Roschian centrality of category membership, so emphasized by the AFA as a link to the popular field of cognitive psychology, is simply another way of evaluating the relevance of items for the indexing of a concept when applied to inventory items. The AFA prototypical rating procedure, besides providing the catch of a connection with the field of cognition, delivers no information or recognitions or insights beyond what has long been available through the rational item evaluation procedures used in personality assessment. The Buss and Craik finding, heavily emphasized, of "the internal structure of a dispositional category" (1983c, p. 111) celebrates no more than the long-established recognition that items can be ordered along a continuum of conceptual relevance by rational judgments and that these rational judgments have a significant measure of validity. I have, informally and in an abbreviated way, compared the item scalings provided by the Roschian with the older item evaluation procedures and found them to be fully equivalent. I invite the reader, if only as a gedanken experiment, to do the same. Interestingly, and perhaps as further support for my assertion that prototypicality per se is not an integral aspect of the AFA, it may be noted that in a recent report (Buss, 1987), the prototypicality rating procedure was abandoned for direct judgments of the "effectiveness" of "act" statements.

(c) Traditionally, personality scales are scored by adding up the number of keyed responses to relevant items. This approach presumes a linear relation between individual items and the underlying latent continuum (Jackson, 1971). However, single items, even if deemed relevant, are recognized to insufficiently sample the relevant domain and therefore are viewed as equivalent in their implication. It is by summing over many relevant items, thus sampling the conceptual domain representatively, that a measure takes on reliability and an improved validity. This approach, as already noted, goes back many decades to the time when the conceptual foundations of psychometrics were formulated.

Often, however, by its phrasings the AFA seems to take solitary and special credit for discounting the importance of single items, instead proclaiming the importance of "multiple-act criteria" or "act trend indices." Thus, in criticism of psychological research wherein judges were compelled to make dispositional inferences from a single "act" of a person, Buss and Craik (1983c) declared: "The act frequency approach holds explicitly that prototypicality per se is not an integral aspect of the AFA, it may be noted that in a recent report (Buss, 1987), the prototypicality rating procedure was abandoned for direct judgments of the "effectiveness" of "act" statements.

So, I suggest that the RSRAPS constructed by the AFA are tantamount to a beginning effort at scale construction as long practiced within the field of personality assessment: Potential items are generated, they are judged a priori regarding their likely relevance, and they are used as a scale scored by summing keyed items.

Whereas good scale construction in personality assessment goes on to a sequential, spiraling revision process involving convergent and discriminant analyses of items and scales, the AFA does not. Instead, the sum of affirmed responses to the particular set of "act" statements that has been constituted becomes
defined as the “multiple-act index.”^4 Then, by its central tenet, the AFA declares that this “multiple-act index, grounded in everyday conduct, is prior and the search is for empirically effective [personality] scales that can serve as efficient predictors of it” (Buss & Craik, 1983a, p. 1082). The arbitrarily assembled “act” statements in the RSRAPs are held to be inviolable, as not requiring further evaluation because, by assertion, the “multiple-act index” is the “theoretically sanctioned” criterion against which all other measures are to be assessed.

The “priority” claimed for RSRAPs represents a strong (even breathtaking) position, given that only retrospective responses to booklets of statements are involved. There are many consequences if this position is accepted. Consequently, the reader should ponder the empirical basis of this claim, on which so much is based by the AFA, and evaluate its entitlement.

4. **The AFA has derived various RSRAP indices and then interpreted them in absolute terms without realizing these absolutely interpreted indices are, wittingly or not, readily manipulable.** Consider, for example, the AFA “act density index” and the “act extensity index.”^5 “Act density” is the number of significant correlates an established personality scale has with the individual “act” statements included in a nominally appropriate RSRAP. For example, the CPI Dominance scale correlates significantly with 39 of the 100 “act” statements in the dominance RSRAP. “Act extensity” in the AFA indexes the extent to which a traditional personality scale correlates with “act” statements in RSRAPs other than its semantically equivalent RSRAP. For example, Buss and Craik reported that the CPI Dominance scale correlated significantly with 13 “aloofness act statements,” 39 “gregarious act statements,” 6 “quarrelsome act statements,” and 15 “agreeableness act statements” (1983a, p. 1091).

Interestingly, and alarmingly, Buss and Craik (1983a) reported a positive correlation of .60 between “act density” and “act extensity,” indicating that “scales that predict act frequencies well in the domains for which they were developed also predict act frequencies in other, sometimes quite disparate and possibly inappropriate categories” (p. 1093). The extent of this problem—that many “act” statements relate as well or better to dispositions other than those they have been assigned to represent—has been well documented (Angleitner & Demtroeder, 1988; Borkenau, 1986). The multiple category membership of “act” statements presents large difficulties for the AFA; it means that many “act” statements appear to have been misclassified or have ambiguous implications for dispositions.

In traditional approaches to the construction of personality scales, the general problem of the confounded, multiple meanings of items is well recognized and there are various procedures for minimizing the invalidity or “confuseability” of items, thus improving scale validity and discriminant validity.

In the AFA, however, the “act density index” is viewed as a fixed quantity that represents either the “inherent predictability” (Buss & Craik, 1983a, p. 1094) of the act category represented by the particular RSRAP or the quality of the particular personality scales being used as “predictors of act frequencies.” The “act extensity index” is interpreted as “revealing the precise, underlying psychological dimension of each scale” (p. 1092). It is not recognized or acknowledged within the AFA that the “act density” or “act extensity” of a personality scale may also indicate the presence, within the RSRAP, of inadequate or conceptually imprecise “act” items and, furthermore, that these indices will change, should the “act” statements in the particular RSRAP be changed. Thus, the absolutely interpreted “act density index” is fortuitously based, is readily manipulable, and therefore has no necessary implication for either the “inherent predictability” of an act category or the “empirical effectiveness” of personality scales in predicting “act trend indices,” as the AFA claims.

In conventional personality assessment, on the other hand, the “act density index” would be thought of as a transitional quantity, to be used to identify those “act” items in the RSRAP that relate and do not relate to conceptually appropriate external measures, as a basis for creating an improved RSRAP by deleting the undiscriminating “act” items.

These diametrically different ways of looking at the “act density index” may perhaps be evaluated pragmatically by comparing RSRAPs and established personality scales with respect to their developmental and empirical bases. The RSRAP for the disposition of dominance was established as earlier described, then used without further improvement, and is said to “constitute the appropriate criteria” (Buss & Craik, 1980, p. 391) and to “provide[e] a basic contribution to the analysis of personality scale validity” (Buss & Craik, 1983a, p. 1093). In comparison, the CPI Dominance scale wasearly and carefully developed using a succession of item analysis procedures and, over many years, it has demonstrated its validity many times vis-à-vis a variety of relevant criteria (Gough, 1987). Both the dominance RSRAP and the CPI Dominance scale are based on self-report data. Should one measure serve as the criterion for the other? Should the CPI Dominance scale be evaluated with respect to how well it “predicts” the dominance “act trend”? Should the recent, empirically unevaulated dominance RSRAP be tested with respect to how well it correlates with the CPI Dominance scale? When the correlation is low, or at least not high, which measure should be faulted? Indeed, against a fallible real-life criterion of dominance behavior, which self-report procedure is likely to predict better, the dominance RSRAP postulated by

[^4]: It should also be noted that, by calling only for yes responses to index a disposition, retrospective self-recordings of act performances (RSRAPs) are susceptible to the possible response set of acquiescence or readiness to endorse “act” statements. This factor may account for the act frequency approach (AFA) finding that conceptually opposed or unrelated RSRAPs correlate with each other in theoretically dismaying ways; for example, the aloofness and agreeableness “act trend indices” correlated positively (Buss & Craik, 1983c, Table 4). Only recently has the necessity of controlling for this factor in the AFA been recognized (Botwin & Buss, 1987). For many years now, good practice in personality assessment has called for a balancing of the number of keyed yes responses with an equal number of keyed no responses in developing scales. It should also be noted, as a conceptual point, that a disposition may be revealed frequently by disavowals of certain acts in certain contexts. The AFA, by the terms of its approach to measurement, misses the possibilities of useful no responses, although in later articles, it acknowledges this problem (Buss & Craik, 1983a).

[^5]: The “act density” and “act extensity” indices were introduced with specific reference to the Wiggins model of the interpersonal circumplex (Wiggins, 1979). However, in its discussions of these indices, the act frequency approach appears to apply these notions in a broader context, not restricted to the Wiggins model.
the AFA as an ultimate criterion or the CPI Do Scale? The reader should formulate a personal judgment or anticipation regarding these questions.

Before going on, the AFA claim that RSRAPs achieve a new level of psychological measurement warrants some mention and comment. It is argued that "act frequencies over a period of observation entail, in principle, zero-based magnitudes" (Buss & Craik, 1983b, p. 402). That is, a true, absolute, zero-point is considered to exist and ratios of "act trend indices" are considered meaningful by the AFA. A person who affirms 5 dominance "act" statements at one testing and 10 dominance "act" statements some months later is said by the AFA to have become twice as dominant. The AFA also proposes that (a) individuals can be compared absolutely (i.e., Individual A has three times more aloof than Individual B, and an individual who does not affirm any of the provided "act" statements for gregariousness has a zero amount of this disposition); and (b) different dispositions can be absolutely compared (i.e., this individual is two times more "dominant" than he or she is "agreeable"). Model human tendencies, as revealed by the mean scores of groups on various RSRAPs, are said to provide insight into human nature.

Conventional personality assessment orientations are modest and do not claim the possibility of ratio, absolute measurement; they remain relativistic, presuming only ordinal or interval scales of measurement. The reason is that, for a claim of ratio scale properties to be made for a measure of a concept, it is necessary to establish that the entire universe of content for that concept is sampled by the measure and no one wishes to make this claim. The absolute measurement aspirations of the AFA, presented as achievable "in principle," are in practice never realizable because, again, the "act trend indices" gathered and compared will depend on the particularities of the set of "act" statements being used to represent the disposition and on the dependability and validity of the retrospective self-reports regarding these "act" statements.

I submit that the AFA does not seem to recognize that many of the empirical relations it describes in absolute terms within its conceptual framework can be readily changed or manipulated, wittingly or not, by using a different mix of the "act" statements in an RSRAP.

5. The "act" statements within an RSRAP, when unusual in the conditions they specify, are often technically unacceptable and, when acceptable, are often not especially different from the kinds of items used in conventional inventory scales. Further and dismaying, many "act" statements seem to have no conceptual or empirical connection with the disposition being measured. If the procedures for the initial development and scoring of an RSRAP are no different from the procedures long used in personality assessment, perhaps the signal contribution of the AFA is the introduction of new and importantly different kinds of item content. Perhaps "act" statements open up new vistas or domains of inquiry in the way the primitive term, "act," is verbally operationalized. By reading "act" statements and reflecting on them, one can gain a sense, finally, of just how the AFA means the concept or primitive term of "act." To understand the empirical contribution of the AFA, there is no substitute for the experience of actually reading the "act" statements that have been placed into the RSRAPs, which now number several dozen.

Consider the RSRAP for the disposition of "aloofness." One "aloof act" statement is "I displayed no emotion when meeting long lost friends at the airport." This item was judged to be highly prototypical of "aloofness." It is also an item to which, I suggest, perhaps no one validly would say yes. After all, what is the probability that one has conjointly (a) within a reasonably delineated period of time, say 3 months, (b) met long lost friends (c) at the airport and (d) displayed no emotion? All four of these conditions must be met before this statement can be validly affirmed. Although judged to be highly prototypical, this "act" statement is technically useless and unacceptable if it offers no or little believable discrimination.

Similar concerns apply to many of the "act" statements deemed highly prototypical. Some more aloofness examples: "In the sensitivity group, I changed the topic whenever someone asked me about my feelings." What is the likelihood of the intersection between (a) for a specified 3-month period, (b) one's being in a sensitivity group, (c) one's being asked about one's feelings, and (d) one's changing of the topic? Another aloofness "act" statement reads "While the other class members sat on the floor in a circle, I sat down behind them near the door." What is the likelihood of the intersection between (a) members of my class sitting on the floor, (b) in a circle, (c) I sitting down behind them, (d) near the door, (e) during the last 3 months? How literally should one contemplate these "act" statements? How should one answer if all but one of these conditions is met? What if the members of my class sat on the grass rather than on the floor, in an ellipse rather than in a circle? How do I respond to my logical concern that if I sit behind some of my class members arranged in a circle, I must sit in front of others? Seriously and closely responded to, such "act" statements do not or should not contribute to "the multiple-act criterion" or "act trend index" simply because almost no one who would affirm such statements could be believed literally.

If a paradoxical feature of the AFA is that many of the most prototypical "act" statements are also items to which almost no one can truly say yes, the other side of this coin involves "act" statements that everyone will affirm. For example, highly prototypical but, I wager, nondiscriminating "act" statements for the disposition of "agreeableness" include "I said thank you to someone who opened the door for me" and "I smiled and laughed at the jokes that were told." What citizen of the Western world, in retrospecting over the preceding 3 months, would not say yes to such "act" statements? "Act" statements such as these also contribute to the finding by A. Angleitner (personal communication, September 10, 1987) that the probability of endorsement of an AFA statement, across a set of diverse dispositions, correlates appreciably (about .5) with the rated social desirability of the AFA statement.

---

6 The "act trend indices" will also depend on the appropriateness or pertinence of the act statements in the retrospective self-recording of act performance (RSRAP) for the subjects responding to the RSRAP. Typically, thus far, college students have been the "act nominators" as well as the responders to RSRAPs. Were noncollege adults to respond to a RSRAP derived from undergraduate nominations, very different "mod-r" human tendencies would be found.
It is perhaps because of these several problems that, in recent AFA research, 25-item rather than 100-item RSRAPS are being used, the 25 items being those most prototypical after discarding “act” statements to which almost no one responds or to which almost everyone responds (Buss, 1984, p. 366). It is not indicated, in this later research, just how prototypical these 25 items prove to be, how they compare with earlier RSRAPS, and whether the relations devolving from the earlier used RSRAPS would also be found with these later, radically changed RSRAPS.

Besides the presence of prototypical but undifferentiating “act” statements, the RSRAPS also present subjects with many statements whose psychological meanings, I suggest, are difficult to assimilate to the dispositions nominally being indexed. Consider, for example, the disposition of “aloofness,” defined in my dictionary as meaning “reserved or reticent; indifferent; cool, detached, distant, standoffish; snobbish, haughty, disdainful” (Random House Dictionary of the English Language, 1987, p. 59). Is the “act” statement “I hid in my bedroom when the others came over to visit” deemed highly prototypical of aloofness by the AFA, really an expression of psychologically distancing oneself from others? Perhaps I was chatting with my guests while, with my back to them, I prepared their drinks.

There are more instances of the inclusion in RSRAPS, with no seeming conceptual or empirical justification, of “act” statements that, I suggest, reasonable people simply cannot comprehend as related to the dispositional category specified. These perplexing “act” statements are generally low in prototypicality, but it must be recalled that, by definition within the AFA, each “act” statement does fall within the dispositional category specified, although “act” statements may differ in their degree of prototypicality. It is also the case that these surprising “act” statements (examples to follow immediately) can have relatively high endorsement rates and so contribute appreciably and perhaps inordinately to the “act trend indices” ultimately computed from the RSRAPS.

Statements such as “At my party, I stood chatting with my back to most of the guests” deemed highly prototypical of aloofness by the AFA, really an expression of psychologically distancing oneself from others? Perhaps I was chatting with my guests while, with my back to them, I prepared their drinks.

The reader should reflect on the following question: How fundamental and far-reaching is the distinction between a generally stated extraversion inventory item such as “I tend to talk a lot” and the more specific or concrete extraversion “act” statement “I talked to almost everyone at the party”? Both statements represent self-reports by a subject regarding past behavior. Conceptually, how important is their difference? Empirically, which is the better item to index the construct of “extraversion?” Can researchers presume or declare that responses to the “act” statement provide a criterion or are superior indicators of extraversion? I submit that, conceptually, the difference is unimportant and, empirically, the inventory item, expressed in general terms, will do as well or perhaps even better than the “act” statement precisely because it is not constrained by the specific circumstances (“at the party,” “at the airport”), often making an “act” statement literally inapplicable or answerable only in (unacknowledged) general terms.

Interacting with the problem of perplexing AFA item content is the problem of the response format provided by the AFA to the respondent. The not unusual “act” statement “I made no plans for the weekend” is designated by the AFA as an “act” indicator of “introversion” when affirmed. One can, of course, debate whether not making plans can be construed as an “act” within the AFA and whether not making plans for the weekend was reasonably selected within the AFA as an indicator of “introversion.” However, consider especially the categories of response with which the subject is confronted: “Within the last 3 months, I made no plans for the weekend zero times, once time, twice times, once a month, twice a month, once a week, a few times a week, or daily.” A response of zero is a confusing way of saying, literally, that at no time in the last 3 months did I not make plans for the weekend (i.e., for 12 or 13 successive weekends, I had made weekend plans). A response of daily means, literally, that every single day during the last 3 months I have made plans for the weekend (i.e., weekend plans made one day were remade the next day and the day after, and so on). Will subjects understand the implications of these response categories and respond coherently?

Or consider the “ingenious act” statement “Within the last 3 months, I admired their idea of a Utopian society zero times, once, twice, once a month, twice a month, once a week, a few times a week, or daily.” Or consider the “quarrelsome act” statement “Within the last 3 months, I accused my teacher for giving a bad lecture zero times, once, twice, once a month, twice a month, once a week, a few times a week, or daily.” Or the “submissiveness act” statement “I did not interrupt my neighbor’s long monologue.” There are many, many other examples.

I believe that many of the “act” statements, when conjoined with the response format imposed by the AFA, present confusing or unreal choices to subjects and that the subsequent re-
sponses of subjects cannot be accepted literally, as the AFA by its view of such responses as "appropriate criteria" requires. I note that, perhaps in partial recognition of the problems empirically encountered by use of the response format required by the AFA and perhaps for the purposes of analytic simplification, in the various data analyses reported by the AFA researchers the continuistic responses of subjects are generally dichotomized (into the categories, zero times versus nonzero times). However, this data analytic decision, of course, fundamentally distorts the meaning of the "act base rates" heavily emphasized by the AFA (Buss, 1985b). In sum, with regard to both statement content and format of statement presentation, the AFA appears to have disregarded long-understood personality assessment principles for the construction of self-report scales.

6. The research program of the AFA is formulaic, consists almost exclusively of what I call "internal analyses" (i.e., analyses of relations within and among the RSRAPS and between RSRAPS and semantically related, concurrently administered self-report measures), and envisions an endlessly busy, atheoretical, "strictly descriptive," stockpiling future for personality psychology.

Following the development of "act report" inventories for the disposition of dominance, the identical procedure was applied to generate RSRAPS for the dispositions of "aloofness," "gregari-ousness," and "submissiveness." Subsequently, additional RSRAPS were developed using the same protocol for the dispositions of "calculatingness," "ingenousness," "agreeableness," "quarrelsome ness," "extraversion," "introversion," and several dozen more to date. Analyses using these RSRAPS have emphasized (a) further demonstrations of "the internal structure of dispositional categories"; (b) evaluations of the temporal stability of "act trend indices" and related scores derived from RSRAPS; (c) analyses of "act density," "act extensity," and "act bipolarity," as indexed by the fixed RSRAPS; and (d) connections of RSRAPS to semantically related portions of the Wiggins (1979) circumplex model. RSRAPS and semantically related self-report measures administered to undergraduates provide the data primarily and often exclusively used.

In an especially enthusiastic article Buss and Craik (1985) look forward to the application of AFA procedures to literally thousands of dispositions. Their "research agenda" would, in each instance, begin with a possible or proposed personality disposition or trait, generally an adjective that is used to characterize individuals. Allport and Odbert (1936) offered a list of 17,954 such trait names. For each of these proposed traits or dispositions, the AFA would ask a number of criterion questions deemed a "distinctive," sufficient, and advancing basis for establishing the importance of a trait or disposition. In what follows, after each AFA criterion question, I respond with what I believe to be the likely and even certain answers to these questions.

Asks the AFA: For each proposed trait or disposition, can act nominals come up with a reasonable number of suggested act statements? Sure, why not? Given an uncertain conception of what an "act" is or how to delineate an "act," given industriousness, imaginativeness, and a willingness to create unlikely or trivially different or recursive scenarios, an infinite number of "act" statements can be generated for any disposition. The AFA concept of "act volume" (the number of acts that can be generated for an "act" category) does not have serious meaning. Does the disposition overlap with other dispositions in the prototypical acts they each contain? Sure, of course. There are many synonyms and antonyms in language and so overlap is guaranteed literally by definition. The problems consequent on such overlap already have been empirically observed (Angleitner & Demtroeder, 1988; Borkenau, 1986). This AFA question does not need to be asked. The larger issue is how to deal with different degrees and kinds of overlap. The AFA offers no help in this regard.

Is there consensuality in judgments of act prototypicality? There always is. Any set of reasonably intelligent and motivated judges, knowing the meaning of the trait or disposition involved or brought to a common understanding of the trait name and provided with a fixed set of "act" statements varying in their dispositional relevance, will, in the aggregate, produce composites judgments that are reproducible.

Does the act-trend index for that disposition display temporal stability? This is an empirical question, one not unique to the AFA. Personality psychologists have long been interested in such questions, have developed a great deal of information on such matters (cf. Block, 1971), and indeed have introduced theoretical considerations indicating what kinds of dispositions are likely to manifest temporal stability and what kinds will not. The AFA, because of its self-characterization as being "strictly descriptive," offers no theoretical anticipations regarding temporal stability; the AFA will simply have to crank out, disposition by disposition, its own answers to this question.

Are there individual differences in the act trend index manifesting the disposition? If a measure is reliable, then one knows there are individual differences. If an "act trend index" or "multiple-act criterion" is reliable, as more ordinary traditional and contemporary personality measures tend to be, then of course dependable individual differences will be observed.

Do act trends show a reasonable base rate of occurrence? The answer to this question depends on just how the set of "act" statements being used was selected. One can create RSRAPS that have unreasonably high or unreasonably low or just right "act-trend indices;" it is up to the "act" statement selector.

It is recognized by the AFA that "the screening of thousands of dispositional constructs by the act frequency approach is a daunting enterprise" (Buss & Craik, 1985, p. 945), but it is contended that this industry will advance progress on "the larger taxonomic task that faces personality psychology" (p. 945). Buss and Craik (1985) hope for the AFA that At some future date, perhaps visitors to a personality research institute could be shown videotapes depicting 50 or more prototypical acts for each of hundreds of dispositions of systemically established significance. These demonstrations would not be final ends or goals of personality research, of course, but they would offer a vivid and compelling way to communicate the nature of dispositional concepts and their rich and varied instantiation. (p. 944)

Applying the entailed arithmetic, this sought-for AFA future would involve looking at 15,000 or 30,000 or 60,000 videotaped prototypical "acts." One must ask, how long is an "act," especially an "act-in-situ" if "context, style, and inferences about intentionality" are to be perceived? What about the multiple category memberships of acts? Will such a collection of
video-taped prototypical "acts" appreciably advance the science of personality psychology? Can "the larger taxonomic task that faces personality psychology" (Buss & Craik, 1985, p. 945) be met by an approach that is, although arduous, so conceptually unguided?

Concluding Remarks

There are many more aspects of the AFA that warrant consideration and debate. The concerns I have expressed focus primarily on the way the AFA has been implemented. When these matters have been dealt with well enough, then it will become pertinent to consider more closely conceptual and strategic aspects of the AFA. With no or little elaboration, I indicate some of the other problems that I believe will beset the AFA.

1. The very notion of an act must be seriously considered and an attractive definition offered that remains sufficiently psychological rather than amusingly behavioristic. In this endeavor, many complexities and complications will be encountered. In view of previous efforts (Rommetveit, 1987; Searle, 1969), this task can be expected to be formidable and likely intractable.

2. The usefulness or tenability of the AFA to personality can be strongly challenged. As many personality psychologists conceptualize the study of personality, an approach solely in terms of acts is certainly insufficient and is in many respects inaccurate in the relations it would provide. Thus, the AFA ignores individual differences in regard to temperament or stylistic variables (e.g., activity level, reactivity, ruminativeness, etc.). Furthermore, the AFA ignores or cannot know about unobservable behaviors. There are nonacts of central relevance for the study of personality (e.g., unobservable emotions and thoughts). These experiential occurrences cannot be considered to be covert acts because the AFA by its original stance must deal with the behaviorally observable.

3. The AFA, by its summation of "acts" over time, chooses to ignore the distribution of "acts" over time. A bipolar depressed person, oscillating between manic and depressive states, may well have the same frequency count of dominant "acts" as a more normal person but with a different underlying pattern of "acts" not registered via an "act frequency index." Instances abound of personality-relevant behaviors and important regularities that are not discernible by the summation logic of the AFA.

4. Buss and Craik (1984) proudly declared that "the act frequency concept of disposition deals primarily with what the individual is like . . . and not at all with why the person tends to behave that way" (p. 283). The "act frequency approach provides actuarial predictions about future trends in conduct as well as identifies regularities of act patterns. The usefulness of explanatory concepts . . . must be determined subsequently" (p. 245). The AFA thus ignores the explanatory endeavor of scientific personality psychology. It is less than inspiring, speaking as a scientist, to aspire only to providing descriptions of phenomena. In addition, is it wise and indeed possible to be so atheoretical? Can one provide deep or "thick" scientific description without having in mind explanatory concepts in terms of which to organize observations? I do not think so.

References


CRITIQUE OF THE ACT FREQUENCY APPROACH


Received December 4, 1987
Revision received April 5, 1988
Accepted June 6, 1988