External Validity in the Assessment of Intellectual Development in Adulthood

K. Warner Schaie

Ethel Percy Andrus Gerontology Center and Department of Psychology

University of Southern California

Paper presented as

part of a symposium on

External Validity in the Study of Human Development

Paul B. Baltes, Chair

at the 84th Annual Meeting of the

American Psychological Association

Washington, D.C., Sept. 6, 1976

[1] Preparation of this paper was facilitated by research and training grants Ni 400-04, AG 15-10 (National Institute of Aging), MH 70-2 (National Institute of Mental Health), and 93-P-57621/9 (Administration on Aging).

Abstract

The history of psychometric assessment of intelligence is characterized by intensive concerns with the internal validity of tasks and experimental design, but unreasonably strong assumptions with respect to external validity, frequently demanded to extend to the display of competence across many life situations. Beyond young adulthood, external validity problems are exacerbated since single criterion goals comparable to that of educational aptitude in work with the young are not available, and tasks do not retain ecological validity when the situational context of the individual under study changes as a function of developmental progression and idiosycratic modifications of individual life situation and roles.

This paper discusses the relation of intelligence and competence and examines external validity issues for the dimensions of settings, measurement variables, treatment variables and experimental units. It is argued that external validity across situations and life stages cannot be obtained for any single measure of intellectual ability. External validity must be addressed by examining task by person by situation interfaces separately for different life stages and across cohort groupings. A major test construction and validation program is outlined, and examples are given showing how some of the aspects of such a program can be operationalized.

External Validity in the Assessment of Intellectual Development in Adulthood

K. Warner Schaie

Ethel Percy Andrus Gerontology Center and Department of Psychology

University of Southern California

Introduction

The history of psychometric assessment of intelligence, achievement and competence is characterized by intensive concern with the internal validity of operations and research paradigms, but unreasonably strong assumptions, if not substantial question-begging with respect to external validity. The general presumptions seem to be that intellectual ability if properly defined has such all-pervasive characteristics that external validity must be considered a given. Yet, even the most superficial scrutiny would suggest that most tasks used for the assessment of intelldctual abilities have been designed to maximize educational aptitude as the explicit or implicit criterion variable. And further, the criterion situation within which such aptitude may be displayed is primarily that of an educational system which serves as a socialization agent towards the attainment or preservation of white middle-class American norms.

Much of the attack upon IQ testing seen in the past decade (cf. Kamin, 1974) has centered about the lack of external validity of conventional techniques with respect to non-white minority groups. As a result we have had substantial extension of our knowledge regarding the generalizability of various tasks for special populations, but thus far we have seen few systematic attempts to investigate alternate criterion systems. Even more unfortunate has been the failure to study the interaction between person characteristics, tasks and the situations in which intellectual competence is to be exhibited.

Campbell and Stanley (1963) in their classic presentation of threats to the validity of research design attend primarily to the dimensions of experimental units, settings, treatment and measurement variables. They do not emphasize specifically that threat to the validity of experiments which results from the changing relation between the construct validities of the predictor and criterion variables. Nevertheless, in their terminology this particular threat involves the interaction between measurement variables and situations, a matter of crucial concern in the assessment of intellectual ontogeny (also see Cook & Campbell, 1975).

When intellectual functioning is considered beyond young adulthood, external validity problems become exacerbated, as it is unlikely that one can suggest a plausible single criterion situation comparable to that of educational aptitutde, nor will tasks retain their ecological validity when the situational context of the individual to be assessed changes as a function of developmental progressions as well as idiosyncratic changes in individual life situations and roles (also see Fisher, 1973).

Five themes will be addressed in this paper: First we will draw the distinction between <u>intelligence and competence</u>, and show that the issues of external validity in the assessment of intellectual development refer to the manner in which formal tasks of intellectual abilities can be utilized to predict intellectual competence in real life situations. Second, we will consider the taxonomy of situations or <u>settings</u> within which intellectual competence might be displayed at different life stages. Third, we will address the ecological validity of tasks which can sample behavior relevant to the expression of intellectual competence in given life situations, the <u>measurement variables</u>. Fourth, we will discuss the motivational constructs which interact with task and subject characteristics, the <u>treatment variables</u>. And, finally we will reiterate some of the well-known concerns in developmental studies

frequently uncontrollable aspects of the <u>experimental units</u> (see also Baltes, Reese, & Nesselroade, in press; Schaie, 1976b).

Throughout this discussion I will take a life-span development point of view. In line with my own experience, however, I will emphasize that larger portion of the life span past the stage of the initial acquisition and expression of intellectual abilities. Hopefully, I will not merely describe the sorry state of the art, but will in addition point to some recent approaches which may help move us forward.

<u>Intelligence and Competence</u>

In the introduction to their recent monograph, Connolly and Bruner (1973) suggest that the term competence refers to intelligence in its broadest sense; that is, in its aspect of "<u>knowing how</u> rather than simply <u>knowing that</u> (author's italics)". They distinguish between a narrow definition of intelligence as a passive structure of intellect a la Guilford (1967) (whether inherited or acquired), and a much broader delineation of competence as a construct which implies action which may change the environment as well as adapt to the environment. Three attributes of competence are said to involve first the ability to select those features from the total environment that are required information for initiating a course of action; second, initiating a sequence of movements designed to achieve the planned objectives and third, to learn from successes and failures to formulate new plans.

All these considerations suggest that competent behavior will involve the application of structure of intellect in specific situations, the attributes of which may well interact with the developmental level of the individual under study. When we speak of intelligence we must, of course, distinguish between the observed or phenotypic measures of a particular construct, and the latent trait or genotypic construct in itself. In the context of external validity in of unitary traits of intelligence, as represented by the more commonly used intelligence tests, ought to be situation-specific with respect to competence, even though within a given level of ontogeny they might be generalizable across different situations with respect to the intellectual process utilized. But no measure of intellectual ability which is thought to assess a single trait, no matter how elegant, will suffice fully to assess the expression of competence in a given situation. Hence, optimal combinations of such unitary traits will always be required to elicit competent behavior within as well as across situations (Schaie, Note 1).

For our purposes here then we may summarize the intelligence-competence distinction, by proposing that competence be viewed as the phenotypic expression of that combination of genotypic intelligence factors which, given minimally required levels of motivational incentives, will permit adaptive behavior in a specific situation or class of situations. Intelligence, on the other hand, would be viewed as that spectrum of genotypic factors which might be abstracted from phenotypic expressions of adaptive behavior measured across situations.

Settings in which Intellectual Competence is Exhibited

The above considerations suggest that we are concerned with issues of internal validity when we deal with intelligence as a set of mechanisms which permit us to understand the cognitive process, but that external validity issues become paramount when we examine how intelligence relates to our coping in a competent manner with the problems of daily living, i.e. the situations in which intellectual abilities are to be applied and assessed. The early pioneers in the assessment of intellectual ability, who were dealing primarily with children and young adults, were fortunate in that the competence-intelligence distinction may be relatively unimportant when criterion relevance is limited to narrow sets of situations in an organism where the effect of novel experience may suffice to provide the required minimal incentive for adaptive behavior.

But even when there is some limitations of socially prescribed situations there still remains a multiplicity of interactive settings. To paraphrase Bem and Allen (1974), the search for cross-situational consistencies in behavior can hope for no more than quite limited success in the best of circumstances, where if we are lucky it might indeed be possible to predict the behavior of some people some of the time.

What then are the settings within which intellectual development is to be validated? Both methodological and developmental criteria must here be considered. The first have been delineated for us by writers such as Weitz (1961) who provides formal characteristics of criterion relevance or Frederiksen (1972) who stresses the need to taxonomize the attributes of situations. Developmental considerations are more difficult to deal with since they evolve about the issue whether or not the situations in which competence is exhibited will change depending upon qualitative life stage characteristics. For example, we have recently argued that during childhood and adolescence, intellectual processes should be viewed in the context of skill acquisition and only secondarily in their application to life situations containing social consequences. By contrast, intellectual processes in adulthood are usually observed in the context of applications which primarily do have such consequences (Schaie & Marquette, 1975; Labouvie-Vief, 1976; Schaie, 1976a).

In order to assess the external validity of any set of measures of intellectual ability, an orderly investigation has recently been conducted in our laboratory by Rick Scheidt (Note 2, 3) with respect to settings characteristic for the elderly in the Los Angeles area.

We began by talking to individuals over the age of 60 in naturalistic settings such as park benches, senior centers, shopping centers, church and senior volunteer groups, purposely avoiding the immobile or institutionalized elderly. Formats used to elicit descriptions of situations included an interview asking for descriptions or most frequently encountered situations in the respondents' daily life, a questionnaire asking respondents to evaluate such situations along dimensions such as the arousal and dominance of each situation, and diaries in which respondents logged their daily experience in term of situations encountered.

As a next step well-educated elderly raters evaluated each of the situations which had been acquired on a number of evaluative dimensions. Those situations and dimensions were retained for which reliable agreement as to classification was possible. These dimensions (see Scheidt, Note 3, for greater detail) classify situations as to whether they are common or uncommon, supportive or depriving, involve solitary or social activities, and whether they require a passive or active role on the part of the individual. A Q-sort sample, containing five situations in each of the 16 classes, was then constructed to serve as an objective means of defining situational attributes.

With the above technique in hand it will now be possible to define criterion situations by asking individuals to rate their perceived relative competence (respectively obtain ratings from observers who are familiar with the behavior of the individuals studied). We can then relate the assessment tools to each class of situations. Or, where we are faced with novel situations, we can ask raters to assign that situation within the now available classification system. Hopefully, we will thus avoid the trap of assuming that any assessment of intellectual development can have meaningful trans-situational generality. But keep in mind the impossibility of generalizing a situational taxonomy across the entire life-span. We will return to this issue when we discuss the role of the experimental unit, but let me insist here, that just as we will have to develop series of phenotypes for the measurement of intellectual abilities across the life-span, so must we develop series of taxonomies, in terms of their age and cohort appropriateness.

<u>Validity of Measurement Variables</u>

We began our discussion by stressing that phenotypic measures of intelligence would not possess cross-situational validity, even though they might be good estimates of a particular structure of intellect genotype. But in the developmental context we must ask whether such phenotypes, if constructed to fit the characteristics of individuals at one stage of ontogeny, are likely to possess validity with respect to <u>any</u> criterion situation at another stage. For example, a task which may be a perfectly good estimate of performance on a genotypic construct such as visuo-spatial behavior in adolescents or young adults, may turn out to be no more than a measure of test-taking behavior in the old (cf. Marquette, Note 4). Thus it becomes imperative to examine tasks with respect to motivational aspects and the anxiety produced by the conditions of instruction as well as by its construct validity for experimental groups with specified person characteristics (Krauss, Note 5).

None of these issues are exactly new, but the comprehensive effort required to deal with them is yet to be attempted. One of the concerns here, of course, is the need to construct tasks which are a better sample of phenotypes appropriat to different stages of the life-span. Deming and Pressey (1957) pioneered in suggesting a number of interesting procedures more appropriate to the life experience of older individuals. But there is a catch, in that measures must not only be relevant to the population to which they are applied; they must also be of known relevance to a structure of intellect model which is at least in the capable of extending across ontogeny. That is, we expect extensive changes in the phenotype-genotype relationships across the life-span, whether we subscribe to a differentiation-dedifferentiation model (Reinert, 1970), or to some structural reorganization paradigm (Schaie, 1976a; Schaie & Marquette, Note 6). Nevertheless, such changes are likely to refer to the hierarchy of importance of given traits of intellectual ability rather than to the emergence or disappearance of such traits at particular life stages.

We will have to accept the fact that construction must now begin for a series of tasks which cover the major dimensions of intellectual abilities but allow for the fact that there are individual differences by person type, and group differences by life stage and cohort membership, with respect to the phenotypic elements of measurement variables. Thus for each group of individuals with distinct characteristics we need tasks which are sufficiently meaningful and relevant to the life experience of the group being studied so as not to affect construct validity spuriously. Again, let me stress that the issue here is whether or not a test measures the individual's competence in performing with respect to the construct of interest rather than in coping with the mechanics of the test and the situation in which it is administered.

For example, our experience with traditional measures of visuo-spatial ability has led to the identification of substantial cohort differences and relatively early ontogenetic changes. Our observations of the test-taking behavior of the elderly on these tasks, however, suggest that at least part of the apparent decrement may involve the formal characteristics of the measures and their apparent lack of significant import for the individuals examined. Iseli Krauss in our laboratory has begun work with different styles of playing cards to replicate some of the findings on visuo-spatial behavior in the elderly (Krauss & Schaie, Note 7). These stimuli were selected because of their wide use in the adult population, their relatively non-threatening nature, <u>and</u> their suitability for a variety of classification and spatial rotation exercises. It is our intention to pursue similar novel test construction approaches for a representative array of intellectual abilities and we are not unaware that we are facing a task of a magnitude not less than that faced by Binet in his initial measurement of the intelligence of school children (Schaie, 1974; Note 1).

<u>Effects of Treatment Variables</u>

What are the effects of experimental paradigms which in themselves serve as treatments and may serve to obscure or magnify developmental changes and differences in intellectual performance? A number of recent studies suggest that the implicit effects of incentive conditions in eliciting disparate or comparable performance may have been much underrated. For example, Birkhill and Schaie (1975) demonstrated substantial effects upon performance on the Primary Mental Abilities in a group of elderly subjects by introducing reinforcement schedules which minimized or maximized risk taking and cautious (reluctance to guess) behavior. Likewise, Hoyer, Labouvie and Baltes (1973) showed that the speed with which older adults performed on battery of ability tests could be markedly increased by rewarding subjects with trading stamps. And Labouvie-Vief and Gonda (1976) found that training in self-monitoring strategies not only improved performance on complex reasoning problems, but had transfer effects on other ability measures. But such effects are by no means universal. For example, Gribbin, Schaie and Parham (Note 8) did not find any significant differences in PMA performance at any point in the adult age range from 22 to 84, where subjects had been recruited either with the promise of a ten dollar subject fee or without promise of such fee.

Treatment variables may also be introduced covertly by the fact that the experimental paradigm may have differential anxiety-arousing attributes for persons at differential levels of autonomic nervous system integration (e.g. Eisdorfer, Nowlin & Wilkie, 1970), or for persons differing in personality traits related to learning parameters (Schaie & Goulet, in press). Just as the pharmacological literature is much concerned with the "side effects" of drugs, so ought the psychometric literature be concerned with the conditions under which we might expect non-standard performance on assessment devices. But the APA testing increasing literature on experimental artifacts which is clearly relevant (Rosenthal & Rosnow, 1969; see also Schaie & Schaie, 1976, reviewing problems of clinical assessment in the elderly).

There is no simple solution for these problems. In research efforts they demand careful controls as well as weighing the manner in which an increase in the external validity of a test will be paid for by introducing further threats to internal validity. For clinical practice, the lesson would be that it is quite dangerous to let loose individuals trained in assessing a limited subject population at one developmental stage upon any other population with different characteristics without further intensive preparation.

<u>Generalization Across Experimental Units</u>

We are often accused of having obtained our knowledge of behavioral principles via the study of the albino rat and the college psychology sophomore. In the assessment of intellectual abilities we stand similarly accused of basing much of our knowledge on the structure of intellect of white middle class school children and young adults. But the issues of generalizability in developmental studies go even further. Most of these issues have been discussed in much detail elsewhere (Schaie, 1973; 1976b). The most important ones for the assessment of intellectual development concern the difficulties in generalizing developmental functions across successive generations, generalizing from an attrited sample to the original population, generalizing from an intact sample to a population which itself has changed, and even generalizing across samples randomly drawn from a population which has experienced changes in its composition (e.g. Gribbin, Schaie & Stone, Note 9). Other issues of concern in studies of adult intellectual development involve the role of the interaction of pathology and normal aging, differential mortality across successive cohorts, and the effect of secular trends in introducing transient perturbances of the phenotype-genotype relations which may spuriously mimic suspected developmental trends.

And all the issues raised by inevitable utilization of volunteer subjects confound matters further (see Rosenthal & Rosnow, 1975).

I have tried in the past to suggest various methodologies that can h estimate the importance of these factors. None of these are simple or ine pensive and I know of no easy short-cuts if we wish to continue the system study of developing individuals in a rapidly changing environment.

But there are further complications. Not only may there be different patterns of ontogeny in successive generations, but in view of the great pl ticity of adult behavior we must consider different individual patterns by configurations of intellectual abilities may express competence across ontog Fortunately, there appears to be much greater stability in adult personality than has previously been suspected, albeit there are large differences for ma traits between successive generational groupings (Schaie & Parham, 1976). It seems likely therefore that we can properly generalize within personality type across the adult life-span, even though we would expect personality type by ontogenetic stage interactions when cross-sectional comparisons are involved.

Whether we sort individuals by age-group or cohort membership (and note that we suspect that sorting by cohort will be much more useful [cf. Schaie, 1974; Schaie & Gribbin, 1975]) or by life-style person types (cf. Schaie & Gribb Note 10), we must expect to have differential validity for each of such grouping. Our concern then clearly must be with a task by person by person by situation interface, which moves along time frames within individual ontogeny, and no doubt will exhibit shifting patterns across the succession of generations.

<u>Summary and Concluding Remarks</u>

We began our analysis of external validity in the assessment of intellectual development by stressing that psychologists have in the past attempted to search for intellectual genotypes which would have cross-sectional generality in predicting behavioral consequences in specific situations. Our detailed analysis suggests that such approach is a sisyphan task which should be abandoned. Instead, we have pointed to the possibility of contrasting alternate systems of phenotypic measurement for a basic set of major dimensions of intellect. Such measures would take into account the fact that different life stages require differential incentives (such as novelty for the young but meaningfulness for the old), and different test-taking paradigms. Moreover, we have stressed that differential combinations of intellectual genotypes may be involved in the multiplicity of criterion situations facing most individuals throughout life and particularly during adulthood and old age. The proper assessment of intellectual ontogeny therefore requires greater attention to the criterion situations within which intellectual abilities are expressed as behavioral competencies. In addition we have suggested that there may be a variety of personality and/or life-style typologies across which generalization may be fraught with error.

As a consequence of these situations it becomes clear that the measurement of intelligence always involves a task by person by situation interface, but that such interfaces may have different attributes and relations within the ontogeny of the individual and across groups representing different cohort or other socio-cultural or genetic attributes. From the view of the developmental psychologist interested in the assessment of intellectual functioning it will therefore be necessary to engage in both instrument development and major efforts at instrument validation. Such a program will involve the development of parallel tasks for the assessment of intellectual genotypes, but possessing phenotypic characteristics appropriate for different life stage and/or cohort groups. Next, we will have to develop adequate taxonomies of life situations in which competence can be expressed, separately for major portions of the life-span. And finally, we will need to study how intellectual genotypes, measured by person and age/cohort appropriate phenotypes express themselves in life-stage-appropriate criterion-situations.

Some of the examples of work in our laboratory described above repres initial attempts to operationalize these issues. I hope others will join efforts because the program outlined in this paper is a formidable one. D it is not impossible, if addressed by many innovative researchers, and it be pursued if the concept of external validity is to have any meaning in t developmental assessment of intelligence.

Reference Notes

1. Schaie, K. W. The assessment of competence in adulthood and old age. Paper read at the annual meeting of the Western Psychological Association, Los Angeles, 1976.
2. Scheidt, R. J. Situational factors in the assessment of adult competence. Paper presented at the annual meeting of the Western Psychological Association, Los Angeles, 1976. (a)
3. Scheidt, R. J. Taxonomy of situations for the aged: Generating situational criteria. Paper read at the 84th Annual Meeting of the American Psychological Association, Washington, 1976. (b)
4. Marquette, B. W. Limitations on the generalizability of adult competence across situations. Paper read at the annual meeting of the Western Psychological Association, Los Angeles, 1976.
5. Krauss, I. K. Criteria for predictors of competence in the elderly. Paper read at the annual meeting of the Western Psychological Association, Los Angeles, 1976.
6. Schaie, K. W., & Marquette, B. W. Stages in transition: A bio-social analysis of adult behavior. Paper read at the satellite meeting of the International Society for the Study of Human Development, Kiryat Anavim, Israel, 1975.
7. Krauss, I. K., & Schaie, K. W. Errors in spatial rotation in the elderly. Paper read at the 84th Annual Meeting of the American Psychological Association, Washington, 1976.
3. Gribbin, K., Schaie, K. W., & Parham, I. A. Monetary incentive, age and personality. Paper read at the annual meeting of the Gerontological Society, Louisville, 1975.

9. Gribbin, K., Schaie, K. W., & Stone, V. Ability differences between establis and redefined populations in sequential studies. Paper read at the 84th Annual Meeting of the American Psychological Association, Washington, 1976.
10. Schaie, K. W., & Gribbin, K. The impact of environment complexity upon adult cognitive development. Paper read at the biennial meeting of the International Society for the Study of Behavioral Development, Guildford, England, 1975.

References

Baltes, P. B., Reese, H. W., & Nesselroade, T. R. Life Span Developmental Psychology: An Introduction to Research Method. Monterey, California: Brooks-Cole, in press.

Bem, D. J., & Allen, A. On predicting some of the people some of the time: The search for cross-sectional consistencies in behavior. Psychological Review, 1974, 81, 506-520.

Birkhill, W. R., & Schaie, K. W. The effect of differential reinforcement of cautiousness in the intellectual performance of the elderly. Journal of Gerontology, 1975, 30, 578-583.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research in teaching. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand-McNally, 1963. Pp. 171-246.

Connolly, K. J., & Bruner, J. S. Competence: its nature and nurture. In K. J. Connolly & J. S. Bruner (Eds.), The growth of competence. New York: Academic Press, 1973, Pp. 3-10.

Cook, T. C., & Campbell, D. T. The design and conduct of quasi-experiments and true experiments in field settings. In M. D. Dunette (Ed.), Handbook of industrial and organizational research. Chicago: Rand-McNally, 1975.

Deming, J. A., & Pressey, S. L. Tests indigenous to the adult and older years. Journal of Counseling Psychology, 1957, 4, 144-148.

Eisdorfer, C., Nowlin, J., & Wilkie, F. Improvement of learning by modification of autonomous nervous system activity. Science, 1970, 170, 1327-1328.

Fisher, J. Competence, effectiveness, intellectual functioning and aging. Gerontologist, 1973, 13, 62-68.

Frederiksen, N. Towards a taxonomy of situations. American Psychologist, 1972, 27, 114-123.

Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Hoyer, W. J., Labouvie-Vief, G. V., & Baltes, P. B. Modification of response speed deficits and intellectual performance in the elderly. Human Development, 1973, 16, 233-242.

Kamin, L. J. The science and politics of IQ. Hillsdale, N. J.: Eribaum, 1974.

Labouvie-Vief, G. Toward optimizing cognitive competence in later life. Educational Gerontology, 1976, 1, 75-92.

Labouvie-Vief, G., & Gonda, J. N. Cognitive strategy training and intellectual performance in the elderly. Journal of Gerontology, 1976, 31, 327-332.

Reinert, G. R. Comparative factor analytic studies of intelligence throughout the human life-span. In L. R. Goulet & P. B. Baltes (Eds.), Life-span developmental psychology: Research and theory. New York: Academic Press, 1970. Pp. 468-485.

Rosenthal, R., & Rosnow, R. L. (Eds.). Artifact in behavioral research. New York: Academic Press, 1969.

Rosenthal, R., & Rosnow, R. L. The volunteer subject. New York: Wiley, 1975.

Schaie, K. W. Methodological problems in descriptive developmental research on adulthood and aging. In J. R. Nesselroade & H. W. Reese (Eds.), Life-span developmental psychology: Methodological issues. New York: Academic Press 1973, Pp. 253-280.

Schaie, K. W. Translations in Gerontology--From lab to life: Intellectual functioning. American Psychologist, 1974, 29, 802-807.

Schaie, K. W. Toward a stage theory of adult cognitive development. International Journal of Aging and Human Development, 1976, in press. (a)

Schaie, K. W. Quasi-experimental designs in the psychology of aging. In J. E. Birren & K. W. Schaie (Eds.), Handbook of the psychology of aging. New York: Van Nostrand-Reinhold, 1976. Pp. 39-58. (b)

Schaie, K. W., & Goulet, L. R. Trait theory and verbal learning processes. In R. M. Dreger (Ed.), _Handbook of modern personality theory_. New York: Hemisphere, in press.

Schaie, K. W., & Gribbin, K. Adult development and aging. _Annual Review of Psychology_, 1975, _26_, 65-96.

Schaie, K. W., & Parham, I. A. Stability of personality traits: Fact or fable? _Journal of Personality and Social Psychology_, 1976, _34_, 146-158.

Schaie, K. W., & Schaie, J. P. Clinical assessment and aging. In J. E. Birren & K. W. Schaie (Eds.), _Handbook of the psychology of aging_. New York: Van Nostrand-Reinhold, 1976.

Weitz, J. Criteria for criteria. _American Psychologist_, 1961, _16_, 228-231.