All items are coded as follows:

1. Not at all characteristic.
2. Slightly characteristic (trivial, questionable, minimal).
3. Somewhat characteristic (moderate, definite).
4. Very characteristic (a great deal, strongly). Requires verbal or nonverbal evidence of intensity.

Note: Very characteristic is the level obtained by normal individuals.

The criteria for applying these codes are quantified for several items (#1, #2, #4, #5, #12). These quantifiable items (labeled Q in Table 1) are rated by counting the number of instances cited by the subject for a particular item (e.g., number of interests, number of friends):

1. Not at all: 0 items
2. Slightly : 1-2 items
3. Somewhat: 2-3 items
4. Very: 3 or more

When there is difficulty in choosing between ratings, the following guidelines are used:

1. In general, rate toward the more apathetic score.

2. Consider the degree of differentiation of responses. For example, rate "Interest in things" as Slightly if a subject simply specifies "reading and television" as interests, but Somewhat if specific books or television programs can be specified. Similarly, if a subject is interested "only" in reading, but provides multiple examples of reading materials, rate Somewhat or Very, based on the number of examples given.

3. Consider the presence of verbal and nonverbal evidence of affect. For example, rate toward lower apathy if subject uses phrases such as "very much" or "tremendously," or uses facial expression, gesture, or vocal intonation to suggest affect.

4. If still in doubt, ask the patient whether, for example, "Somewhat" or "Very" is the more appropriate descriptor.

Steinberg, H.R.; Green, R.; and Durell, J. Depression occurring during the course of postpsychotic depression. *American Journal of Psychiatry,* 124:699-702, 1967.

Stuss, D.T., and Benson, D.F. *The Frontal Lobes.* New York: Raven Press, 1975.

Weiner, B. *Theories of Motivation: From Mechanism to Cognition.* Chicago: Rand-McNally, 1972.

Wildroe, H.J. Depression following acute schizophrenic psychoses. *Journal of Hillside Hospital,* 15:114-122, 1966.

Zung, W.W.K. A self-rating depression scale. *Archives of General Psychiatry,* 12:63-70, 1965.

Zung, W.W.K. A rating instrument for anxiety disorders. *Psychosomatics,* 12:371-379, 1971.

Zung, W.W.K.; Coppedge, H.M.; and Green, R.L. The evaluation of depressive symptomatology: A triadic approach. *Psychotherapy and Psychosomatics,* 24:170-174, 1974.

Appendix: Administration Guidelines

Instructions for the self- and informant-rated versions of the Apathy Evaluation Scale (AES): "For each question, circle the answer that best describes your (his/her) thoughts, feelings, and actions during the past 4 weeks."

For the clinician-rated version of the AES, the following additional instructions apply:

Instructions to patient: "I am going to ask you a series of questions about your thoughts, feelings, and activities. Base your answers on the last 4 weeks. To begin, tell me about your current interests. Tell me about anything that is of interest to you. r example, hobbies or work; activities you are involved in or that you would like to do; interests within the home or outside; with other people or alone; interests that you may be unable to pursue, but which are of interest to you—for example, swimming even though it's winter."

Interviewer then notes: (1) Number of interests reported; (2) degree of detail reported for each interest; (3) affective aspects of expression (verbal and nonverbal).

Interviewer then states: "Now I'd like you to tell me about your average day. Start from the time you wake up and go to the time you go to sleep." *How* the patient deals with this (and all other) questions is assumed to provide information about how other activities are dealt with (e.g., with initiative, exuberance, and energy). Therefore, prompting is indicated only if the subject seems not to understand what information is being sought or has forgotten the question.

Interviewer notes number of activities, degree of detail, intensity and duration of involvement, and affect associated with presentation of data.

Each item is now presented using the wording of the item itself. Additional information may be requested to clarify responses. Item 15, "Accurate understanding of problems," is rated by appraising subject's awareness and understanding of personal or, if present, clinical problems. Simple bridges between items may be used to preserve a conversational quality to the interview. Items are rated as they are esented using all information acquired. The response recorded is the clinician's assessment of the subject's response. Thus, if a subject states "a lot" but the clinician judges "somewhat," the latter is used. The only exceptions are the self-evaluation (SE) items in Table 1 (#3, #8, #13, #16). For these items, the subject specifies which response code to use (e.g., Not at All, Slightly); the clinician rater's appraisal is not considered for SE items.

Hamilton, M. The assessment of anxiety states by rating. *British Journal of Medical Psychology,* 32:50-55, 1959.

Hamilton, M. Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology,* 6:278-296, 1967.

Hecaen, H., and Albert, M. Disorders of mental functioning related to frontal lobe pathology. In: Benson, D.F., and Blumer, D., eds. *Psychiatric Aspects of Neurologic Disease.* New York: Grune & Stratton, 1975. pp. 137-149.

Hyland, M.E. Motivational control theory: An integrative framework. *Journal of Personality and Social Psychology,* 55:642-651, 1988.

Inamdar, S.C.; Siemopoulos, G.; Osborn, M.; and Bianchi, E.C. Phenomenology associated with depressed moods in adolescents. *American Journal of Psychiatry,* 136, 156-159, 1979.

Kohn, M., and Rosman, B.L. A social competence scale and symptom checklist for the preschool child. *Developmental Psychology,* 6:430-444, 1972.

Maller, O. Amotivational syndrome in chronic schizophrenics: A biophysiological model of schizophrenic impairment. *Neuropsychobiology,* 4:229-247, 1978.

Marin, R.S. Differential diagnosis and classification of apathy. *American Journal of Psychiatry,* 147:22-30, 1990.

Marin, R.S. Apathy: A neuropsychiatric syndrome. *Journal of Neuropsychiatry and Clinical Neuroscience,* 1991.

Marsden, C., and Parkes, J. Success and problems of long-term levodopa therapy in Parkinson's disease. *Lancet,* I:345-349, 1977.

McGlashan, T.H., and Carpenter, W.T., Jr. Postpsychotic depression in schizophrenia. *Archives of General Psychiatry,* 33:231-239, 1976.

McHugh, S.E., and Folstein, M.F. Psychiatric features of Huntington's disease: Recent approaches and findings. *Psychiatric Developments,* 2:193-206, 1983.

McKhann, G.; Drachman, D.; Folstein, M.; Katzman, R.; Price, D.; and Stadlan, E. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology,* 34:939-944, 1984.

Neugarten, B.L.; Havighurst, R.J.; and Tobin, S.S. Personality and patterns of aging. In: Neugarten, B.L., ed. *Middle Age and Aging.* Chicago, IL: University of Chicago Press, 1968. pp. 173-177.

Overall, J.E. Dimensions of manifest depression. *Journal of Psychiatric Research,* 1:239-245, 1963.

Raskin, A., and Sathananthan, G. Depression in the elderly. *Psychopharmacology Bulletin,* 15:14-16, 1979.

Reisberg, B. *Clinical Presentation, Diagnosis, and Symptomatology of Age-Associated Cognitive Decline and Alzheimer's Disease.* New York: The Free Press, 1983.

Robinson, R.G. Depression and apathy following stroke. Presented at the Annual Meeting of the American Psychiatric Association, New York, NY, 1990.

Robinson, R.G.; Kubos, K.L.; Starr, L.B, Rao, K.; and Price, T.R. Mood disorders in stroke patients. *Brain,* 107:81-93, 1984.

Schulterbrandt, J.; Raskin, A.; and Reatig, N. Further replication of factors of psychopathology in the interview, ward behavior and self-reported ratings of hospitalized depressives. *Psychological Reports,* 34:23-32, 1974.

Sjögren, T.; Sjögren, H.; and Lindgren, A. Morbus Alzheimer and morbus Pick: A genetic, clinical and patho-anatomical study. *Acta Psychiatrica et Neurologica Scandinavica,* 82:1-152, 1952.

Smith, J.; Bright, B.; and McCloskey, J. Factor analytic composition of the geriatric rating scale (GRS). *Journal of Gerontology,* 32:58-62, 1977.

Sourander, P., and Sjögren, H. *The Concept of Alzheimer's Disease and Its Clinical Implications.* London: Churchill, 1970.

Spitzer, R.L.; Endicott, J.; and Robins, E. Research Diagnostic Criteria: Rationale and reliability. *Archives of General Psychiatry,* 35:773-782, 1978.

Acknowledgments. The research reported was supported in part by grants from the National Institute of Mental Health (MH-41930) and the National Institute of Aging (Academic Award AG00235). The late Norman Geschwind, M.D., made the clinical observations that brought the topic of apathy to the attention of the senior author. The following colleagues contributed generously to this research: Karen A. Matthews, Ph.D.; the late Robert E. Miller, Ph.D.; Anthony Nitko, Ph.D.; Paul Pilkonis, Ph.D.; Charles F. Reynolds III, M.D., Ralph E. Tarter, Ph.D.; Gerhard Werner, M.D.; and the late Joseph Zubin, Ph.D.

References

Abrams, R., and Taylor, M.A. A rating scale for emotional blunting. *American Journal of Psychiatry,* 135:226-229, 1978.

Albert, M.; Feldman, R.; and Willis, A. The "subcortical dementia" of progressive supranuclear palsy. *Journal of Neurology, Neurosurgery, and Psychiatry,* 37:121-130, 1974.

Andreasen, N.C. Affective flattening and the criteria for schizophrenia. *American Journal of Psychiatry,* 136:944-947, 1979.

Andreasen, N.C. Negative symptoms in schizophrenia. *Archives of General Psychiatry,* 39:784-795, 1982.

Atkinson, J.W., and Birch, D. *An Introduction to Motivation.* Princeton, NJ: Van Nostrand, 1978.

Benson, D.F. Subcortical dementia: A clinical approach. In: Mayeux, R., and Rosen, W.G., eds. *The Dementias.* New York: Raven Press, 1983.

Bleuler, E. *Dementia Praecox or the Group of Schizophrenias.* New York: International Universities Press, 1950.

Caine, E.; Hunt, R.; Weingartner, H.; and Ebert, M.H. Huntington's dementia: Clinical and neuropsychological features. *Archives of General Psychiatry,* 35:377-384, 1978.

Campbell, D.T., and Fisk, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin,* 56:81-105, 1959.

Crocker, L., and Algina, J. *Introduction to Classical and Modern Test Theory.* New York: Holt, Rinehart and Winston, 1986.

Crow, T.J. Molecular pathology of schizophrenia: More than one disease process? *British Medical Journal,* 12:66-72, 1980.

Deci, E.L., and Ryan, R.M. *Intrinsic Motivation and Self-Determination in Human Motivation.* New York: Plenum Press, 1985.

Depue, R.A., and Iacono, W.G. Neurobehavioral aspects of affective disorders. *Annual Review of Psychology,* 40:457-492, 1989.

Fishbein, M., and Agzen, I. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research.* Reading, MA: Addison-Wesley, 1975.

Fisk, D.W. The use of significant others in assessing the outcome of psychotherapy. In: Waskow, I.E., and Parloff, M.B., eds. *Psychotherapy Change Measures.* Washington, DC: Superintendent of Documents, U.S. Government Printing Office, 1975. pp. 189-201.

Floru, L.; Heinrich, K.; and Witteh, F. The problem of postpsychotic schizophrenic depressions and their pharmacological induction. *International Pharmacopsychiatry,* 10:230-239, 1975.

Folstein, M.F.; Folstein, S.E.; and McHugh, P.R. Mini Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research,* 12:189-198, 1975.

Friedman, A.S.; Cowitz, B, Cohen, H.; and Granick, S. Syndromes and themes of psychotic depression. *Archives of General Psychiatry,* 9:504-509, 1963.

Gainotti, G. Emotional behavior and hemispheric side of lesion. *Cortex,* 8:41-55, 1972.

Grinker, R.R.; Miller, J.; Sabshin, M.; Nunn, R.; and Nunnally, J.C. *The Phenomena of Depressions.* New York: Paul B. Hoeber, Inc., 1961.

measures. These correlations were all in the predicted directions and showed a consistent pattern across raters, in which the strength of correlations generally descended in the order AES-C, AES-I, and AES-S. Particularly noteworthy with respect to the discriminant validity of the AES was the fact that the depression ratings did not predict behavioral outcomes with the marginal exception of the informant depression rating, as discussed above.

The interpretation of our results will be influenced by variables that have not been considered here. The relationship between apathy and cognitive impairment or functional capacity is of particular relevance for individuals who suffer from neurological injury or who have functional impairments. Presumably, neurological variables are primarily responsible for apathy in some instances (e.g., RH stroke, frontal lobe injury, negative symptoms in schizophrenia, and postpsychotic depression), while psychological and socioenvironmental variables may dominate in others (e.g., role loss, institutionalism, or effector system impairments, such as spinal cord injury; Marin, 1990). When apathy represents a primary deficit in the central neurological mechanisms that are concerned with motivational capacity, the causes and treatments of apathy are likely to be quite different from those involved when lack of motivation arises from loss of effector mechanisms (e.g., amputations or paraplegia), from loss of cognitive abilities necessary to organize adaptive behavior (e.g., dementia), or from negative cognitions or emotional distress (e.g., some cases of depression or anxiety). It is also possible that multiple variables interact in some disorders to produce apathetic states. For example, the depressed individual might show apathy as a result of all of these considerations: neurogenic impairment in primary motivational systems (Depue and Iacono, 1989); neuropsychological deficits; low subjective expectancy of success or negative perceptions of potential sources of reward (Fishbein and Agzen, 1975); and maladaptive affective or behavioral responses that elicit reduction in environmental rewards.

Since apathy is, in essence, lack of motivation, the AES may be useful in exploring a variety of problems of relevance to psychiatry and neurology. The concept of apathy may link clinical disorders to a variety of theories and methodologies that have been developed to understand motivation (Weiner, 1972; Atkinson and Birch, 1978; Deci and Ryan, 1985; Hyland, 1988). For prognostic purposes, a measure of motivation may be a valid predictor of recovery from stroke, hip fracture, or other illnesses. For management purposes, the concept of apathy may facilitate differential diagnosis or guide families grappling with the performance difficulties of patients with dementia, schizophrenia, or focal brain diseases. Given a means to measure apathy, investigators may systematically evaluate the neurobiological or socio-environmental determinants of apathy in such varied circumstances as frontal lobe injury, stroke, Alzheimer's disease, Parkinson's disease, schizophrenia, and depression. In the case of role loss or institutionalism, it is also possible that apathy may serve as a variable sensitive to psychological, social, or environmental characteristics. Given the breadth of biological and psychosocial variables that may influence apathy, diverse strategies may be suitable in developing effective treatments for apathy.

intervention). The convergent validity coefficients for the AES suggest that under some circumstances, one version of the AES might be used as an alternative to the other (e.g., when severe impairment precludes self-ratings).

Validity. The elevated levels of apathy for AD, RH, and DP subjects indicate that the AES is able to discriminate among groups of subjects based on mean levels of apathy. The success of self- and informant-ratings in making these same distinctions is noteworthy. Since some individuals with clinically significant apathy are brain damaged, their insight and judgment in evaluating a construct as seemingly vague or subjective as apathy would be expected to be impaired. Furthermore, the multitrait-multimethod matrix analysis indicated that the AES-S discriminated apathy from depression almost as well as the AES-C. Self-ratings had weaker correlations with the behavioral outcome measures, however. This may be due to the narrower range of AES-S scores. Normals tended to rate themselves as slightly apathetic, while apathetic subjects tended to underestimate the severity of their apathy.

The multitrait-multimethod matrix procedure was generally supportive of the convergent validity and discriminant validity of the AES. Each of the convergent validity coefficients for the AES was strongly positive, although less so for the AES-I than for the AES-C and AES-S. Although higher convergence for the AES-I would be desirable, it should be noted that the informant convergent validity coefficients or depression and anxiety were also relatively low, suggesting an effect due to rater rather than construct. The multitrait-multimethod matrix supported the discriminant validity of the AES-C and AES-S since convergence coefficients were higher than the within-rater (heterotrait homomethod) discriminant validity coefficients. By this criterion, however, the AES-I did not appear to discriminate apathy from depression. It would not be surprising if the informants had difficulty discriminating apathy and depression. Despite their familiarity with the subjects, informants actually may not be acquainted with the subjective information, such as mood or subjects' appraisals of their own efforts and values, which are critical for discriminating apathy and depression. Fisk (1975) previously noted the limitations of informants in making evaluations that are dependent on subjective phenomena. On the other hand, several observations supported the ability of the AES-I to discriminate a dimension distinguishable from depression. First, the informants' ratings were effective in distinguishing between groups based on mean apathy levels and did so in a pattern similar to that of other apathy rather than other depression ratings. Second, the predictive validity of the informants' apathy ratings was nearly as strong as that of the clinicians' ratings. We suspect that the failure of the AES-I to discriminate depression according to multitrait-multimethod matrix criteria was not due to the AES-I but to the informant version of the Zung-D. When the HRSD or he self-rated version of the Zung-D were used to discriminate depression (heterotrait heteromethod comparisons), the AES-I did discriminate apathy from depression. This interpretation was supported by the predictive validity measures which showed that the only depression rating that had some correlation with the outcome variables was the informant version of the Zung-D.

Regarding the predictive validity of the AES, we found that the informant and clinician apathy ratings showed significant correlations with several of the outcome

For PACman, the average score per game correlated -0.45 ($p < 0.001$), -0.40 ($p < 0.001$), and -0.29 ($p = 0.014$) with the clinician-, informant-, and self-rated apathy scores, respectively. There also were significant correlations between apathy scores and the average response latency for signaling the decision to play again (AES-C: $r = 0.45$, $p < 0.001$; AES-I: $r = 0.35$, $p = 0.005$; AES-S: $r = 0.26$, $p = 0.036$). Number of cycles played was uncorrelated with apathy. The informant depression score correlated -0.33 ($p = 0.004$) with average score per game and 0.16 (NS) with response latency.

In the maze game, the number of times the subject played the game was again uncorrelated with apathy scores. However, apathy scores for the clinician-, informant-, and self-rated versions of the scale correlated -0.37 ($p = 0.001$), -0.26 ($p = 0.033$), and -0.31 ($p = 0.010$), respectively, with the difficulty level at which subjects chose to play. In the slot machine game, there were no significant correlations between apathy and the number of cycles played. Depression scores were uncorrelated with maze or slot machine outcomes, although the correlation of informant-rated depression with initial level of difficulty in the maze game was 0.23 ($p = 0.058$).

Discussion

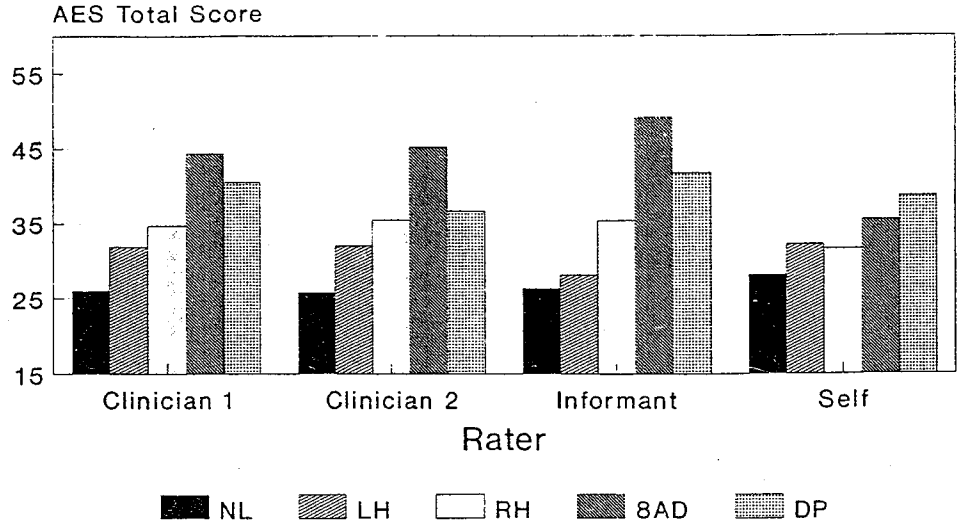
The reliability and validity studies reported here provide substantial support for the reliability and construct validity of the AES. Each version of the AES was found to have good to excellent reliability. Several types of validity were demonstrated for each version of the scale.

Subjects. Several aspects of the subjects in this study warrant comment. Although the subjects were predominantly elderly, the AES was not designed as a geriatric rating scale per se. The scale should be usable for adolescent and adult populations with a variety of clinical disorders. Modifications would undoubtedly be necessary to explore the utility of the concept for special populations, such as children or mentally retarded subjects, although the domains of inquiry would be similar.

The subjects in this study had mild to moderate levels of impairment. The ability of the AES to distinguish groups according to mean apathy levels and to predict behavior probably would be enhanced if a larger range of impairments were represented in the subjects. The fact that we failed to find a significant difference between the LH and RH stroke groups may also result from the moderate level of impairment in our subjects. However, a recent study by Robinson (1990) found that apathy was more strongly associated with LH than RH strokes.

Reliability. For each version of the AES, three types of reliability have been evaluated. Each has practical implications for future investigations or clinical applications. The AES-C's interrater reliability suggests that multiple raters can be trained to use the scale in a similar fashion. The correlation coefficients for test-retest reliability suggest that the AES can be used to evaluate the stability of apathy over time. A corollary is that the AES can be used to evaluate the extent to which apathy changes in concert with, or independently of, other clinical variables (e.g., mood, cognition, functional impairment, environmental manipulation, or pharmacological

Fig. 1. Mean apathy scores on the Apathy Evaluation Scale



NL=Well Elderly; LH & RH=Left and Right Hemisphere Stroke, respectively; AD=Alzheimer's Disease; DP=Major Depression

Table 6. Correlation (Pearson's *r*) of apathy and depression ratings with predictive validity measures

	PACman average	PACman latency	Maze level difference	% Total time	Average time per game
No. cases	72	63	65	82	76
AES-C	-0.45	0.45	-0.37	-0.40	-0.24
	(< 0.001)	(< 0.001)	(0.002)	(< 0.001)	(< 0.001)
AES-I	-0.40	0.35	-0.26	-0.33	-0.34
	(< 0.001)	(0.005)	(0.033)	(0.004)	(0.003)
AES-S	-0.29	0.26	-0.31	-0.15	-0.13
	(0.014)	(0.036)	(0.010)	(NS)	(NS)
HRSD	-0.08	-0.05	-0.13	-0.03	-0.09
	(NS)	(NS)	(NS)	(NS)	(NS)
Dep-I	-0.33	0.16	-0.23	-0.10	-0.10
	(0.004)	(NS)	(0.058)	(NS)	(NS)
Dep-S	-0.20	-0.01	-0.04	-0.05	-0.04
	(0.072)	(NS)	(NS)	(NS)	(NS)

Note. PACman average = average score per game of PACman. PACman latency = average latency for decision to play PACman 2 or more times. AEC = Apathy Evaluation Scale (C = Clinician, I = Informant, and S = Self). HRSD = Hamilton Rating Scale for Depression. Dep-I and Dep-S = Zung ratings of depression by informant and self, respectively. Tests of significance are 2-tailed.

Table 5. Convergent and discriminant validity of Apathy Evaluation Scale

	Clinician			Informant			Self	
	Apathy	Depression	Anxiety	Apathy	Depression	Anxiety	Apathy	Depression
Clinician								
Apathy								
Depression	0.39							
Anxiety	0.35[3]	0.87						
Informant								
Apathy	0.62	0.23[2]	0.15[1]					
Depression	0.56	0.53	0.41	0.65				
Anxiety	0.46	0.67	0.45	0.43	0.74			
Self								
Apathy	0.72	0.35[3]	0.37	0.43	0.46	0.42		
Depression	0.41	0.74	0.71	0.27[3]	0.53	0.56	0.42	
Anxiety	0.36	0.56	0.63	0.23[2]	0.42	0.55	0.42	0.78

Note. Intercorrelations are shown among clinician-, informant-, and self-rated measures of apathy, depression, and anxiety. Apathy scores are self, informant, and clinician (R.C.B.) scores on the Apathy Evaluation Scale. Clinician measures of depression and anxiety are from the Hamilton Rating Scales for Depression and the Hamilton Rating Scale for Anxiety, respectively. Informant- and self-rated measures for depression and anxiety are both from the Zung Rating Scale for Depression and the Zung Rating Scale for Anxiety, respectively. Underlined items are convergent validity coefficients.

All correlations are significant at $p < 0.001$ except as noted by superscripts:
1. NS.
2. $p < 0.05$, 2-tailed.
3. $p < 0.01$, 2-tailed.

Table 4. Reliability

Full scale	Internal consistency	Test-retest	Interrater
AES-C	0.90	0.88	0.94
AES-I	0.94	0.94	
AES-S	0.86	0.76	

Kappa coefficients for items

1.	0.53	10.	0.50
2.	0.60	11.	0.50
3.	0.85	12.	0.63
4.	0.64	13.	0.73
5.	0.49	14.	0.49
6.	0.65	15.	0.45
7.	0.53	16.	0.73
8.	0.73	17.	0.53
9.	0.56	18.	0.35

Note. Reliability coefficients represent: Cronbach's α for internal consistency. Pearson's r for test-retest reliability. Intraclass correlation coefficient for interrater reliability. AES-C measures of internal consistency and test-retest reliability are based on the ratings of clinician 2 (R.C.B.).

nificantly elevated apathy levels for each rater for AD (Clinician 1: $tD = -7.43$, $df = 105$, $p \leq 0.05$; Clinician 2: $tD = -7.40$, $df = 104$, $p \leq 0.05$; Informant: $tD = -7.56$, $df = 103$, $p \leq 0.05$; Self: $tD = -3.30$, $df = 118$, $p \leq 0.05$) and for DP (Clinician 1: $tD = -6.22$, $df = 105$, $p \leq 0.05$; Clinician 2: $tD = -4.26$, $df = 104$, $p \leq 0.05$; Informant: $tD = -5.32$, $df = 103$, $p \leq 0.05$; Self: $tD = -5.18$, $df = 118$, $p \leq 0.05$). In addition, for the Informant and the two Clinician raters, RH was significantly higher than NL (Clinician 1: $tD = -3.34$, $df = 105$, $p \leq 0.05$; Clinician 2: $tD = -3.53$, $df = 104$, $p \leq 0.05$; Informant: $tD = -3.03$, $df = 103$, $p \leq 0.05$; Self: $tD = -1.56$, $df = 118$, NS). RH was not significantly different from LH for any of the raters. An analysis of covariance showed no significant effects of age or education on mean levels of apathy.

Predictive validity. The novelty toy/waiting room experiment provided a naturalistic way of observing freely initiated behavior. Regarding the video games, most subjects were readily engaged and learned them easily. Several of the more impaired AD subjects, however, were unable to learn them. Table 6 summarizes the correlations (Pearson's r) of apathy and depression ratings with the dependent variables.

In the novelty toy/waiting room experiment, there were no significant correlations between AES scores and the number of games (including repeated uses of the same me) subjects used. However, ratings of apathy by clinicians and informants correlated -0.40 ($p \leq 0.001$) and -0.33 ($p = 0.004$), respectively, with percentage of total time (PTT). Self-rated apathy scores showed nonsignificant, although negative, correlations with PTT. Average time per game, calculated as the ratio of total time subjects used games and the number of games used (including repeated uses of the same game), also showed significant correlations with clinician and informant apathy. Depression ratings showed no significant correlations with the dependent variables.

Table 3. Clinical characteristics of subject groups

	MMS		HRSD		Dep-I		Dep-S		AES-C1		AES-C2		AES-I		AES-S	
Group	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
NL	29.1	1.1	3.9	3.9	28.4	4.4	30.4	6.2	26.0	6.2	25.8	5.8	26.3	7.5	28.1	6.4
LH	25.0	4.6	6.4	6.2	37.7	8.2	37.4	9.4	31.9	9.6	32.0	11.7	28.1	6.9	32.2	8.6
RH	26.9	2.3	9.2	4.7	42.1	9.9	39.3	7.6	34.7	7.3	35.4	9.6	35.4	10.9	31.6	6.7
AD	19.1	6.5	5.1	5.0	43.7	6.5	35.0	7.2	44.4	11.1	45.2	11.7	49.1	9.9	35.5	8.1
MD	28.0	1.7	19.8	5.8	46.9	9.7	49.2	6.5	40.5	9.7	36.6	8.3	41.7	15.0	38.7	9.8

Note. NL = well elderly. LH and RH = left hemisphere and right hemisphere stroke, respectively. AD = probable Alzheimer's disease. MD = major depression. MMS = Mini-Mental State Examination. HRSD = Hamilton Rating Scale for Depression. Dep-I and Dep-S = Zung depression scale, informant and self-rated versions, respectively. AES-C1, AES-C2, AES-I, and AES-S = scores for Apathy Evaluation Scale as rated by the 2 clinician raters (R.S.M., R.C.B.), informants, and subjects' self-ratings, respectively.

Table 2. Demographic characteristics of subject groups

Group	No.	Sex		Age		Education	Income
		M	F	Mean	SD		
Well elderly	31	14	17	68.3	5.7	3.8	3.7
Left hemisphere stroke	19	11	8	66.2	6.6	2.6	3.0
Right hemisphere stroke	22	12	10	70.1	5.0	2.1	1.7
Probable Alzheimer's disease	21	10	11	70.8	7.6	3.1	3.1
Major depression	30	3	27	71.6	5.7	2.3	2.0
Total	123	50	73				

Note. Education: 1 ≤ 12 yr; 2 = completed high school; 3 > 12 yr; 4 = graduated from college. 5 = at least some graduate training. Income: 1 < $10,000 annual income; 2 = $10,000-<$20,000; 3 = $20,000-<$30,000; 4 = $30,000-<$40,000. 5 > $40,000.

tions of 0.62 and 0.72 of clinician-rated apathy with informant-rated apathy and self-rated apathy, respectively. Similarly, self-rated apathy correlated 0.42 with self-rated depression, by comparison to a correlation of 0.72 with clinician-rated apathy. On the other hand, informant ratings of apathy showed almost the same correlation with informant ratings of depression (0.65) as they did with clinician-rated apathy (0.62) and actually showed better correlations with informant-rated depression than with self-rated apathy (0.43). These observations suggested that the clinician-rated nd self-rated versions of the AES discriminated apathy from depression, but the informant ratings may not have. When the informant apathy rating was compared with clinician-rated depression or self-rated depression scores, however, good discrimination was found between apathy and depression: informant apathy correlated only 0.23 ($p = 0.03$) with the HRSD and 0.27 ($p < 0.01$) with self-rated depression. Thus, the failure of the AES-I to meet the multitrait-multimethod matrix criteria for discriminant validity might have been due to the informant version of the Zung Depression Rating Scale whose validity as a measure of depression has received less attention than that of the others (Zung et al., 1974). That possibility was supported by the predictive validity data below.

Results were similar when apathy and anxiety intercorrelations were evaluated. The within-rater (heterotrait homomethod) discriminant validity coefficients for apathy vs. anxiety showed lower intercorrelations between apathy and anxiety for the clinician- ($r = 0.35$) and self-ratings ($r = 0.42$), compared with the convergent validity coefficients for apathy. As with depression, there was no discrimination for the informant ratings ($r = 0.43$). For anxiety, too, use of the clinician- and self-rated anxiety measures as the criteria for discriminating apathy and anxiety revealed marked discrimination: informant-rated apathy correlated 0.15 ($p = 0.15$) with the Hamilton Rating Scale for Anxiety and 0.23 ($p = 0.03$) with self-rated anxiety.

Group differences in apathy levels. Fig. 1 depicts the mean apathy levels for the five diagnostic groups. One-way analyses of variance for effect of diagnosis were highly significant for each rater source (Clinician 1: $F = 17.20$, $df = 109$, $p \leq 0.0001$; Clinician 2: $F = 14.37$, $df = 108$, $p \leq 0.0001$; Informant: $F = 18.01$, $df = 107$, $p \leq 0.0001$; Self: $F = 7.46$, $df = 122$, $p \leq 0.0001$). Dunn's test for a priori linear comparisons indicated that, in comparison with ratings for normals, there were

Table 1. Items and factor structure of the Apathy Evaluation Scale

1.	S/he is interested in things.	+	C		Q
2.	S/he gets things done during the day.	+	B		Q
3.	Getting things started on his/her own is important to him/her.	+	C	D	SE
4.	S/he is interested in having new experiences.	+	C		Q
5.	S/he is interested in learning new things.	+	C		Q
6.	S/he puts little effort into anything.	−	B		
7.	S/he approaches life with intensity.	+	E		
8.	Seeing a job through to the end is important to her/him.	+	C	D	SE
9.	S/he spends time doing things that interest her/him.	+	B		
10.	Someone has to tell her/him what to do each day.	−	B	D	
11.	S/he is less concerned about her/his problems than s/he should be.	−	C	D	
12.	S/he has friends.	+	B		Q
13.	Getting together with friends is important to her/him.	+	C	D	SE
14.	When something good happens, s/he gets excited.	+	E	D	
15.	S/he has an accurate understanding of her/his problems.	+	O	D	
16.	Getting things done during the day is important to her/him.	+	C		SE
17.	S/he has initiative.	+	O		
18.	S/he has motivation.	+	O		

	Clinician 1		Clinician 2		Informant		Self	
	Item	Loading	Item	Loading	Item	Loading	Item	Loading
F 1	16	0.76	16	0.78	16	0.81	16	0.77
	17	0.74	8	0.75	7	0.79	3	0.77
	18	0.72	18	0.68	2	0.79	8	0.67
	7	0.66	7	0.58	18	0.78	17	0.62
	8	0.63	6	0.57	17	0.71	18	0.60
	3	0.63	2	0.52	1	0.70	2	0.59
	6	0.55	3	0.52	3	0.64	1	0.58
	2	0.50	14	0.51	4	0.63	7	0.56
	14	0.32			8	0.59		
					5	0.56		
					6	0.54		
					14	0.53		
					9	0.52		
F 2	5	0.79	4	0.83	12	0.88	13	0.78
	4	0.79	5	0.74	13	0.63	4	0.68
	1	0.65	9	0.66	15	0.58	9	0.60
	9	0.62	1	0.63			5	0.58
	12	0.43	17	0.63			14	0.56
			13	0.44			12	0.53
F 3	11	0.81	15	0.87	11	0.77	10	0.74
	15	0.80	11	0.82	10	0.72	6	0.58
	10	0.50	10	0.63			11	0.53
	13	0.41	12	0.53			15	0.38

Note. Items that have positive vs. negative syntax are identified by +/−. C,B,E,O: Type of item as discussed on p. 145. C = cognitive; B = behavior; E = emotional; O = other. The definitions of self-evaluation (SE) items and quantifiable items (Q) are discussed in the administration guidelines (see Appendix). D = items whose correlation (Pearson's *r*) with the Hamilton Rating Scale for Depression is nonsignificant.

correlations with depression. The items for impaired insight and dependency on others for structuring daily activities suggest that these items may be endorsed preferentially by the more severely impaired subjects. Further data analysis will be required to evaluate this possibility.

Reliability. All measures of reliability were satisfactory for each version of the AES (Table 4). Internal consistency reliability, measured as coefficient α, was 0.86-0.94 for the different raters. Test-retest reliability (mean test-retest interval 25.4 days) varied from 0.76 to 0.94, with the lower value being attributable to a test-retest reliability coefficient of 0.44 for the AD group's self-ratings. The intraclass correlation coefficient for the two clinician raters was 0.94. Mean kappa for the AES-C, calculated as the average kappa for the 18 items included in the AES, was 0.58 (Table 4).

Validity. Table 1 presents the 18 items used to evaluate the reliability and validity of the AES. The items have face validity for such pertinent motivational variables as productivity, initiative, effort, emotional responsivity, novelty seeking or curiosity, perseverance, and social engagement. Although we intended for theoretical reasons to have equal numbers of items to represent the overt behavioral, cognitive, and emotional aspects of goal-directed behavior, the data did not support such a selection of items. The paucity of emotional items is probably compensated for by the fact that these domains were not dissociated in our subjects. Also, the administration guidelines explicitly specify that emotional information be considered in evaluating overt behavioral and cognitive items. As shown in Table 1 and based on the test administration guidelines for the AES, five items dealt with overt behavior, seven dealt with cognitive aspects of motivation, two dealt with emotional responsivity, and three were not readily categorized in these respects. Two of the latter (motivation and initiative) constituted abstract characterizations of behavior in explicitly motivational terms. In fact, these items were sufficiently abstract that patients with lower levels of education occasionally questioned their meaning. Item 15 (accurate understanding of one's problems) was originally included in the scale along with other items to evaluate insight and other features (e.g., irritability and euphoria) that might be useful in distinguishing subtypes of apathy syndromes (e.g., frontal lobe syndromes). Item 15 was retained because it was subsequently found to have nonsignificant correlation with the HRSD (see **Methods**).

Convergent and discriminant validity. Convergent and discriminant validity coefficients (Pearson's r) were generally consistent with the multitrait-multimethod matrix criteria (Table 5). Convergent validity coefficients for apathy, depression, and anxiety (underlined) were all positive and significant ($p < 0.001$). For apathy, the convergent validity coefficients (i.e., intercorrelatons among AES-C, AES-I, and AES-S) were between 0.43 and 0.72. The AES-C and AES-S convergent validity coefficients were moderately higher than the within-rater correlations between apathy and depression (heterotrait homomethod comparisons), suggesting that the raters did distinguish between the two constructs. For example, clinician-rated apathy and clinician-rated depression correlated 0.39 by comparison with correla-

which the PACman figures moved through the display maze. After subjects understood the object of the game and could control the PACman figure with the joystick, they were instructed that the game could be played as many times as desired. The computer recorded the score for each game, the number of times the game was played (number of cycles of play), and the response latency for deciding to play again.

Video Game 2: Maze. In this game (modified with permission from an original game developed by Linda Plesko), subjects used the joystick to move a dog through a path which a bird had just flown through without leaving a trail. The maze consisted of connected horizontal and vertical pathways representing the spaces created by four rows of four square blocks. The dog moved horizontally or vertically one block at a time. Each move required the joystick to be moved once in the appropriate direction. After the training period, subjects chose the level of difficulty for the game. In other words, would the bird make 2, 4, 6, 8, or 10 moves to reach its destination? Subjects played the game as many times as desired (number of cycles).

Video Game 3: Slot Machine. This game was a computer-based version of the familiar one-armed bandit slot machine game. Subjects played as many times as desired. They did not have to spend money and could not lose money. To play, subjects just pulled the slot machine manipulandum (joystick). The pattern of winning was predetermined so that in the first 25 cycles subjects won $5.00, and after that there were no wins. Thus, subjects all received the same pattern and level of reenforcement in winning $5.00, after which they received no further reenforcement—a simple operant paradigm. The hypothesis was that apathy would be negatively correlated with the number of times subjects chose to play the game once reward ended.

Results

Subjects. Results are reported for 123 subjects, with 19-31 subjects per diagnostic group (Table 2). Except for the 40 pilot subjects, development and validation procedures were evaluated in the same subjects.

As summarized in Table 2, there were significant differences between groups for sex ($\chi^2 = 16.29$, $df = 4$, $p = 0.003$), age ($F = 3.31$; $df = 4, 118$; $p = 0.013$), education ($F = 7.73$; $df = 4, 118$; $p < 0.0001$), and income ($F = 6.90$; $df = 4, 118$; $p < 0.0001$). Age was slightly lower in the LH group. The group differences in sex reflect primarily the predominance of women in the MD group. Education and income differences were higher in the NL and AD groups, and lower in the RH and DP groups. Table 3 summarizes the groups' MMS, depression, anxiety, and apathy ratings.

Factor Structure. Principal components factor analysis of the AES identified three similar factors in the factor analyses carried out on each of the rater sources. Together, these three factors accounted for 50-65% of the total variance for the different raters. However, the scale was predominantly a single factor scale. Factor 1 represented a general apathy factor accounting for 32-53% of the variance of the scale for the four rater sources. Factor 2 accounted for 5-10% of the variance for the four rater sources. It included the items that dealt with curiousity or novelty seeking: interest in things (in general), learning, new experiences, and spending time in interesting activities. Factor 3 included the items for insight, lack of concern about one's problems, and needing someone to provide structure for daily activities. It accounted for 7-8% of the scale's variance. These items all had nonsignificant

the Department of Psychiatry, University of Pittsburgh School of Medicine.Tables 2 and 3 summarize demographic and clinical characteristics of the subjects.

Procedure. The subjects, who were paid $10 per hour for their participation, completed self-rating scales for apathy, depression, and anxiety immediately before the administration of the clinician-rated scales. Informants completed their scales without consultation with subjects or interviewers.

Instructions for the AES indicate that items should be answered based on the subject's thoughts, emotions, and actions over the previous 4 weeks. The AES-S and AES-I are administered as paper-and-pencil tests. The AES-C is administered as a semistructured interview. Instructions for administering the AES are included in the Appendix.

The clinician raters in this study were a board-certified psychiatrist with fellowship training in neurology and geriatric psychiatry (R.S.M.), and a research associate with a master's degree in counseling education and 2 years' experience in personal and career counseling (R.C.B.). The research associate was trained during the testing of 30 pilot subjects.

To evaluate test-retest reliability, the three versions of the AES were administered on a second occasion at least 2 weeks after the first administration. To evaluate interrater reliability, simultaneous ratings of subjects were carried out by the clinician raters on one of these two occasions. The raters alternated in the role of interviewer. The observer was permitted to ask additional questions. However, to obtain a more stringent test of interrater reliability, the primary interviewer was not permitted to modify ratings based on additional information elicited in this way.

In addition to the AES, clinician-, informant-, and self-rated measures of depression and anxiety were also obtained. Clinician-rated measures of depression and anxiety were drawn from the HRSD (first 17 items) and the Hamilton Rating Scale for Anxiety (Hamilton, 1959). Self- and informant-ratings of depression and anxiety were obtained using the Zung Self-Rating Scale for Depression (Zung-D; Zung, 1965) and the Zung Anxiety Scale (Zung, 1971). These scales were developed as self-rated instruments, but, as reported by Zung et al. (1974) for the Zung-D, are easily converted to informant-rated devices.

The following procedures were used to evaluate predictive validity: In the novelty toy/waiting room procedure, subjects were seated in a waiting room immediately behind a coffee table on which a variety of novelty toys and games were placed—for example, an Etch-a-Sketch, a Slinky, a kaleidoscope, a magnetic toy, and a 3-inch square number puzzle with 15 numbered, movable tiles. Subjects were told that we needed to organize the next test procedures and that the toys on the table were for them to use if desired while waiting. Subjects' behavior was videotaped. The variables calculated were the percentage of the total time that subjects spent on any of the games and the number of games used.

The video games were selected to present a range of difficulties so that impaired and unimpaired subjects might find at least one of the games engaging. All games used an Apple IIe computer fitted with a joystick and an ECHO+ voice synthesizer. The games are described below in order from the most complex (PACman) to the least complex (slot machine). The sequence in which subjects played the games was randomized within diagnostic groups. Because of the total length of testing (8-9 hours/subject) and the fact that subjects were not all able to return for multiple visits, not all subjects played all games. In contrast to the novelty toy/waiting room procedure, subjects were encouraged to play the video games since, as the investigator was leaving the room at the conclusion of the training period for each game, subjects were told, "You can start the game now." However, no termination time was specified. Rather, after each cycle of a particular game, subjects were prompted by the computer to "Pull the joystick down if you want to play again." Therefore, subjects elected to stop playing by not pulling the joystick down at this choice point. When the response latency for this decision reached 1 min, the computer informed subjects that the game was over. This termination procedure was explained for each game.

Video Game 1: PACman. An Apple version of the familiar PACman was modified (by Greg Autry, HAL Labs, Perris, CA) so that during the training period subjects chose the rate at

To develop the present 18-item scale, we calculated the item-total correlations for each item in the 56-item scale, evaluating each version of the AES separately. As a preliminary step in scale reduction, we retained only items that had item-total correlations of at least 0.4 for at least three of the four rater sources (2 clinician ratings, 1 informant rating, and 1 self-rating for each subject). This step yielded 27 items that were then subjected to rational and statistical criteria for reaching the 18-item scale. As a rational criterion, we evaluated the correlation of each item with the total score of the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1967). This enabled us to identify seven items that had nonsignificant correlations with the HRSD, and thus would contribute to the discriminability of apathy and depression. As a statistical criterion, we sought preliminary information about the factor structure of these 27 items. Using SPSSpc+ Factor Analysis, we carried out a principal components factor analysis with orthogonal rotation, although we recognized that our sample size was marginal in size to yield a stable factor structure. This permitted us to select items representative of a primary factor and three smaller factors.

Responses to items were recorded on a 4-point Likert-type scale with the following categories: Not at All True, Slightly True, Somewhat True, and Very True. For the AES-C, the comparable terms are: Not at All Characteristic, Slightly Characteristic, Somewhat Characteristic, and Very Characteristic. Reliability and validity data given below were calculated based on a total score for these 18 items. All items were coded so that a higher score represented greater apathy.

Subjects. Subjects were 55 to 85 years old. They resided in a private or community dwelling that did not restrict their activities and were able to identify a friend or family member familiar with their daily activities who would complete the AES-I. All subjects were ambulatory, although some required a cane to walk and assistance to climb steps. Exclusionary criteria were: history of alcoholism; drug abuse; central nervous system disease other than stroke or probable Alzheimer's disease; history of bipolar affective disorder, psychotic disorder, or systemic disease disposing to organic behavioral change; electroconvulsive therapy within the previous 6 months; use of neuroleptic medications within the previous 2 weeks; history of closed head injury with failure to resume previous level of functioning; head injury within 12 months of onset of present illness; insulin-requiring diabetes mellitus; and hypertension not adequately controlled with medication.

Additional criteria were as follows: (1) Stroke patients were required to have history, neurological findings, and computed tomographic (CT) head scan report indicating a single infarction involving the left or right hemisphere (LH or RH, respectively). Basal ganglia infarction was included ($n = 1$ in LH and RH groups). Transient ischemic attacks (TIAs) excluded subjects if the laterality of the TIA indicated involvement of the otherwise intact hemisphere. Aphasic patients were included if their language impairments permitted them to complete the AES with minimal or no assistance. Stroke subjects were studied 3 months to 3 years after their strokes. LH and RH subjects were recruited through the rehabilitation programs of Harmarville Rehabilitation Center, Highland Drive and Oakland VA Medical Centers, the Rehabilitation Institute, and St. Francis Hospital, all located in Pittsburgh. (2) AD subjects met NINCDS-ADRDA criteria (McKhann et al., 1984) for probable Alzheimer's disease and had a score ≥ 10 on the Folstein Mini-Mental State Examination (MMS; Folstein et al., 1975). Subjects were excluded if there was historical or CT evidence of coexisting cerebral infarction. AD subjects had Hachinski Scale Ratings < 5. AD subjects were recruited from the Alzheimer's Disease Research Center of the University of Pittsburgh School of Medicine. (3) Depressives (DP) met Research Diagnostic Criteria (Spitzer et al., 1978) for major depressive disorder (unipolar, nonpsychotic) and had MMS scores > 24. DP subjects were recruited from the inpatient and outpatient programs of Geriatric Health Services, University of Pittsburgh School of Medicine. (4) Normal control subjects (NL) did not meet criteria for any axis I psychiatric diagnosis at the time of evaluation. Inclusion also required that they have an HRSD score < 7 (1 rater, 17 items) and MMS score ≥ 26. NL subjects were obtained through the pools of normal elderly volunteers for research conducted by faculty of

addition, subjects were evaluated for the level of difficulty at which they chose to play one of the video games and for the latency of their decisions to continue playing at these games. These measures were then correlated with AES ratings and, to evaluate further the discriminability of apathy and depression, with depression ratings as well. We predicted that apathy, but not depression, would be negatively correlated with number of cycles of play and with level of difficulty, and positively correlated with response latency.

Methods

Scale Development.

General aims and definition. The aims in developing the AES were to: (1) discriminate apathetic patients from normals; (2) discriminate depression from apathy; and (3) distinguish between different diagnostic groups using only items related to apathy. In other words, instead of relying on diagnosis-specific features, such as sleep disturbance in depression or cognitive impairment in dementia, the goal was to distinguish among apathetic subjects considering domains such as lack of interests, productivity, initiative, perseverance, and affect. These domains are consistent with the definition of apathy as lack of motivation. Our operational definition is that apathy is a state characterized by simultaneous diminution in the overt behavioral, cognitive, and emotional concomitants of goal-directed behavior (see Marin, 1991, for discussion). Examples of these three domains are, respectively, diminished productivity, diminished goals, and diminished emotional responses to success or failure.

Scale structure. AES-C, AES-I, and AES-S each include the same (core) items. The AES-C is supplemented by items that address various aspects of nonverbal behavior and emotionality. These supplementary items are excluded from consideration in this article, which is concerned only with the reliability and validity of the core items present in all three versions of the AES.

Item development. Items were developed through consultation with colleagues, evaluation of pertinent publications, and through the authors' observations and conceptualizations of apathetic patients. The literature review identified clinical descriptions and a variety of scales and factor analytic studies dealing with apathy or closely related concepts. They dealt with a variety of age groups (Neugarten et al., 1968; Kohn and Rosman, 1972; Inamdar et al., 1979; Raskin and Sathananthan, 1979; Smith et al., 1977) and clinical concepts. Diagnoses and symptom types included schizophrenia (Bleuler, 1950; Abrams and Taylor, 1978; Maller, 1978; Andreasen, 1979); depression (Grinker et al., 1961; Friedman et al., 1963; Overall, 1963; Schulterbrandt et al., 1974); negative symptoms (Andreasen, 1982); flat affect (Andreasen, 1979); emotional blunting (Abrams and Taylor, 1978); postpsychotic depression (Wildroe, 1966; Steinberg et al., 1967; Floru et al., 1975; McGlashan and Carpenter, 1976); right hemisphere stroke (Gainotti, 1972; Robinson et al., 1984); frontal lobe injury (Hecaen and Albert, 1975; Stuss and Benson, 1975); Alzheimer's disease (Sjögren et al., 1952; Sourander and Sjögren, 1970; Reisberg, 1983); basal ganglia disorders, including Parkinson's disease (Marsden and Parkes, 1977), progressive supranuclear palsy (Albert et al., 1974), and Huntington's disease (Caine et al., 1978; McHugh and Folstein, 1983); and others (Benson, 1983; Marin, 1990).

Item selection. An original set of several hundred items was reduced to a preliminary scale of 70 items based on the investigators' judgments that items were unambiguous, easily understood, and representative of the domains of interest. Preliminary evaluation of the 70-item scale was carried out on 40 subjects, ages 55 to 85, with primary diagnoses of major depression or dementia. At this point, 14 additional items were eliminated on the basis of qualitative judgments as to their simplicity and clarity. This revised 56-item scale was then administered to the subjects described below. Items were written with both positive and negative syntax, although for purposes of clarity in cognitively impaired subjects, a predominance of positively worded items was found preferable.

ubiquity. The Apathy Evaluation Scale (AES) was developed to quantify and characterize apathy in adult patients. It treats apathy as a psychological dimension that may be evaluated in patients whose apathy characterizes their overall clinical state, and those in whom it is a symptom of some other syndrome, such as delirium, dementia, or depression. This report describes the development, reliability, and validity of the AES. The AES was developed for multiple rater sources: clinician, informant, and self-rated versions (AES-C, AES-I, and AES-S, respectively). Using multiple sources of information permitted evaluation of potentially complementary sources of information. Apathy is often associated with impaired insight—for example, because of its association with frontal lobe injury (Hecaen and Albert, 1975; Stuss and Benson, 1975) or dementing disease (Sjögren et al., 1952; Sourander and Sjögren, 1970; Reisberg, 1983). For this reason, we developed AES versions for a clinician and informant (family member, friend, or caregiver). The informant version complements the clinician version since it is based on direct observation of subjects' behavior in a home environment. By contrast, the clinician version is based on clinical observations and subjects' self-reports during an interview. Despite the obstacles posed by cognitive impairment, insight, or denial of illness, we also tested a self-rated version since it was expected that self-ratings might have at least some validity.

Internal consistency, test-retest, and, for the AES-C, interrater reliability are reported. Regarding validity, three questions are addressed.

1. Can apathy be discriminated from depression? This question is of interest because clinicians who are unfamiliar with the differential diagnosis of apathy (Marin, 1990) often infer that patients who show apathy are depressed. It was approached with the multitrait-multimethod matrix procedure (Campbell and Fisk, 1959), which has been used widely in construct validation (Crocker and Algina, 1986). According to the multitrait-multimethod matrix procedure, validity assessment requires measuring two or more constructs and then evaluating each construct by two or more methods to demonstrate: (a) reliability; (b) convergent validity—different methods used to measure the same construct should show strong positive intercorrelations (homotrait heteromethod correlations); (c) discriminant validity: correlations between different constructs measured by the same methods (heterotrait homomethod correlations) should be substantially less than the convergent validity coefficients. In this study, the constructs evaluated are apathy, depression, and anxiety. Each is separately evaluated by interview and by paper-and-pencil procedures.

2. Does the AES differentiate between groups according to levels of apathy? On the basis of clinical descriptions (see **Methods**), we hypothesized that subjects with probable Alzheimer's disease (AD), right hemisphere stroke (RH), and major depression (DP) would have higher mean levels of apathy than normal subjects (NL). We also hypothesized that RH subjects would have higher levels of apathy than left hemisphere stroke (LH) subjects (Robinson et al., 1984).

3. Does the AES predict behavior in appropriate observational settings? Predictive validity is probably the most important aspect of construct validation. To evaluate predictive validity, subjects were observed in situations that permitted them to initiate and terminate their play at a variety of novelty toys and video games. In

Psychiatry Research, 38:143-162
Elsevier

Reliability and Validity of the Apathy Evaluation Scale

Robert S. Marin, Ruth C. Biedrzycki, and Sekip Firinciogullari

Received January 4, 1991; revised version received July 16, 1991; accepted July 28, 1991.

Abstract. This article presents evidence for the reliability and construct validity of the Apathy Evaluation Scale (AES). Conceptually, apathy is defined as lack of motivation not attributable to diminished level of consciousness, cognitive impairment, or emotional distress. Operationally, the AES treats apathy as a psychological dimension defined by simultaneous deficits in the overt behavioral, cognitive, and emotional concomitants of goal-directed behavior. Three versions of the AES (clinician, informant, and self-rated) were evaluted for 123 subjects, ages 53-85, meeting research criteria for right or left hemisphere stroke, probable Alzheimer's disease, major depression, or well elderly control. Multiple forms of reliability (internal consistency, test-retest, and interrater) were satisfactory. Several types of validity evidence are presented for each version of the scale, including the following: ability of the AES to discriminate between groups according to mean levels of apathy, discriminability of apathy ratings from standard measures of depression and anxiety, convergent validity between the three versions of the scale, and predictive validity measures derived from observing subjects' play with novelty toys and videogames. Guidelines for the administration of the AES are presented, along with suggestions for potential applications of the scale to clinical and research questions.

Key Words. Apathy, depression, dementia, geriatrics, motivation, negative symptoms, organic personality disorder, poststroke affective disorders, stroke.

Apathy has been described in a variety of clinical disorders and is an important psychological response to many major life stressors. For clinical purposes, apathy means lack of motivation that is not attributable to diminished level of consciousness, cognitive impairment, or emotional distress (Marin, 1990). The behavioral changes associated with right hemisphere stroke (Gainotti, 1972; Robinson et al., 1984), frontal lobe injury (Hecaen and Albert, 1975; Stuss and Benson, 1975), and negative symptoms in schizophrenia (Crow, 1980; Andreasen, 1982) are examples of such apathetic syndromes. Apathy also occurs in association with a variety of other clinical problems and may complicate both assessment and treatment (Marin, 1990).

No instrument has been developed specifically to measure apathy, despite its

---

An earlier version of this article was presented at the Annual Meeting of the American Psychiatric Association, New York, NY, May 1990.

Robert S. Marin, M.D., is Assistant Professor of Psychiatry; Ruth C. Biedrzycki, M.Ed., was Senior Research Associate; and Sekip Firinciogullari, M.A., M.S., is Systems Analyst, Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA. (Reprint requests to Dr. R.S. Marin, Western Psychiatric Institute & Clinic, 3811 O'Hara St., Pittsburgh, PA 15213, USA.)