An appraisal of the psychometric properties of the Clinician version of the Apathy Evaluation Scale (AES-C)

DIANA E. CLARKE,¹ ROBERT VAN REEKUM,² JIGISHA PATEL,³ MARTINE SIMARD,⁴
EVERLYNE GOMEZ,⁵ DAVID L. STREINER⁵

1 Toronto Rehabilitation Institute and Department of Psychiatry, Baycrest Centre for Geriatric Care, Canada

2 Baycrest Centre for Geriatric Care and the Institute of Medical Sciences and the Division of Geriatric Psychiatry, University of Toronto, Canada

3 Baycrest Centre for Geriatric Care and Department of Occupational Therapy, University of Toronto, Canada

4 École de Psychologie, Université Laval, Québec, Québec, Centre de Recherche-Université Laval-Robert-Giffard, Québec and École de Psychologie, Université de Moncton, Moncton, Canada

5 Department of Psychiatry and Kunin-Lunenfeld Applied Research Unit, Baycrest Centre for Geriatric Care, Canada

Abstract

This article examines the psychometric properties of the clinician version of the Apathy Evaluation Scale (AES-C) to determine its ability to characterize, quantify and differentiate apathy. Critical appraisals of the item-reduction processes, effectiveness of the administration, coding and scoring procedures, and the reliability and validity of the scale were carried out. For training, administration and rating of the AES-C, clearer guidelines, including a more standardized list of verbal and non-verbal apathetic cues, are needed. There is evidence of high internal consistency for the scale across studies. In addition, the original study reported good test-retest and inter-rater reliability coefficients. However, there is a lack of replication on these more stable and informative measures of reliability and as such they warrant further investigation. The research evidence confirms that the AES-C shows good discriminant, convergent and criterion validity. However, evidence of its predictive validity is limited. As this aspect of validity refers to the scale's ability to predict future outcomes, which is important for treatment and rehabilitation planning, further assessment of the predictive validity of the AES-C is needed. In conclusion, the AES-C is a reliable and valid measure for the characterization and quantification of apathy. Copyright © 2007 John Wiley & Sons, Ltd.

Key words: apathy, AES-C, reliability, validity, psychometric properties

Introduction

The Apathy Evaluation Scale (AES) by Marin et al. (1991) was developed to address the lack of appropriate health measurement tool to assess apathy, a neurobehavioural syndrome characterized by a lack of motivation, self-initiated action, and concern, as well as emotional indifference (Marin et al., 1991, 1993; Stuss et al., 2000). In the development of the scale, apathy was operationalized as a psychological construct

characterized by coexistent deficits in the realms of overt behavioural, cognitive and emotional concomitants of goal-directed behaviour (Marin et al., 1991; Marin et al., 1994). The AES, which has self-rated, informant-rated and clinician-rated versions, includes items that cover the full range of apathy and was developed to provide a standardized method of describing, quantifying, and differentiating levels of apathy in adults aged 55 and older (Marin et al., 1991, 1993).

Marin and colleagues (1991, 1993, 1994) indicated that the clinician version of the AES (i.e. AES-C) functions best, which might account for its consistent use across studies.

Using the AES-C, a number of studies have reported that apathy occurs quite frequently in various health conditions. For instance, Marin et al. (1994) reported that 53.3% of 30 individuals aged 55 to 85 years old with major depression, as defined by the Research Diagnostic Criteria, achieved scores above the cut-off of greater than 37 on the AES-C. Using this cut-off score, a higher prevalence of 86% was reported for apathy in older adults with major depression or dysthymia based on Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) criteria (Lampe and Heeren, 2004). The difference in prevalence rates of apathy in major depression is probably due to one study using an outpatient sample (Marin et al., 1994) and the other an inpatient sample (Lampe and Heeren, 2004) as well as the different diagnostic criteria for depression.

A prevalence of 55% was reported for apathy (AES-C < 37) in AD as determined by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) criteria (Marin et al., 1994). Clarke et al. (2007) also reported high proportions of apathy in AD (NINCDS and/or DSM-IV criteria) using their observed cut-off value of greater than 40.5 on the AES-C (42.9%) as well as Marin's cut-off score of greater than 37 (55.1%). The researchers found that cut-off values of greater than 40.5 and greater than 37 produced the same sensitivity (88%) but the higher cut-off had a slightly better specificity (Clarke et al., 2007). As such, the discrepant findings between these studies are probably due to differences in the cut-off score, with a higher cut-off value capturing the more severely apathetic individuals.

Frequencies of 46.4% (Andersson et al., 1999a) and 62.3% (Andersson and Bergendalen, 2002) were reported for apathy in TBI with each study using Kant et al's (1998) TBI-specific cut-off score of greater than 34. In another study by Andersson et al. (1999b) a prevalence rate of 66.7% for apathy in TBI was reported using the same cut-off score. However, scrutiny of the data revealed that the actual percentage was 33.3% (10/30) not 66.7% (20/30) (see Andersson et al., 1999b). Interestingly, the prevalence rates in the studies with the smaller sample size were lower than that in the study with the larger sample size.

Andersson and others (1999a) also examined the frequency of apathy in cerebrovascular atrophy/disease (CVA) and hypoxic brain damage (HBD) using the cut-off score for TBI (<34) and reported prevalence rates of 56.7% and 78.6% respectively. In contrast, Marin et al. (1994) found prevalence rates of 22% and 32% for left and right hemispheric CVA respectively (AES-C < 37). It is possible that the cut-off score for TBI might not be applicable for CVA or HBD, which would account for the high prevalence rates observed especially for the HBD group. In fact, Clarke et al. (2007) found a higher cut-off score for a sample of older adults with dementia (greater than 40.5), compared to Marin's greater than 37 and the TBI-specific cut-off of greater than 34, and this might be the case for CVA and HBD samples.

With respect to schizophrenia/schizoaffective disorders, Kiang et al. (2003) observed that individuals with these disorders had statistically elevated apathy scores (AES-C) compared to normal controls. An estimate of the frequency of apathy in schizophrenia/schizoaffective disorder was not available because a cut-off score was not applied in this study to classify apathetic versus non-apathetic individuals (Kiang et al., 2003), perhaps due to sample size restrictions.

Parkinson's disease is a subcortical disorder, which is said to be associated with apathy (Brown and Pluck, 2000). Using a cut-off score of 38 and over on the AES-C, Pluck and Brown (2002), found that apathy occurred more frequently in PD (37.8%) compared to osteoarthritis (OA: 0%). The mean apathy score in the PD group was significantly greater ($p > 0.01$) than that in the OA group (Brown and Pluck, 2002). In addition, apathetic individuals have been found to be less compliant with treatment and to have poorer levels of functioning and treatment outcomes compared to non-apathetic individuals (Resnick et al., 1998; Andersson et al., 1999a, 1999b; Lampe et al., 2001).

These findings indicate that apathy is a health condition of both clinical and public health importance. Therefore, the psychometric properties of AES-C, which continues to be one of the most widely used in its characterization and quantification, warrant scrutiny. This article critically appraises the empirical evidence of the psychometric properties of the AES-C to determine its ability to characterize, quantify and differentiate apathy.

Methods

The search engines PSYCHINFO, MEDLINE, CINAHL (Cumulative Index to Nursing and Allied Health) and AGELINE were used to identify original articles and relevant reviews on apathy to provide a general overview of the topic. These searches were limited to publications between January 1983 and December 2004. The search terms used to identify the literature on apathy included, 'apathy', 'lack of interest', 'amotivational', 'emotional indifference', 'lack of motivation', 'lack of initiative', 'negative symptoms', 'asociality',

'anhedonia', 'apathy scales' and 'Apathy Evaluation Scale'. The search term 'Apathy Evaluation Scale' was then combined with all other search terms independently in each database to identify the final articles.

Results

The results of the search strategy are presented in Table 1. In total, 12 comparative studies that used the AES-C and provided information on some aspect of its psychometric properties were found. In addition, our recent study (Clarke et al., 2007), which assessed the

Table 1. Outline of the literature search strategy for the critical appraisal of the AES-C

		Search engine			
		PSYCHINFO	MEDLINE	CINAHL	AGELINE
Key terms	1. 'apathy'	774	1013	151	92
	2. 'lack of interest'	229	294	66	27
	3. 'amotivational'	28	35	3	1
	4. 'emotional indifference'	11	12	0	0
	5. 'lack of motivation'	186	243	59	9
	6. 'lack of initiative'	31	25	4	1
	7. 'negative symptoms'	89	72	0	1
	8. 'asociality'	53	45	1	0
	9. 'anhedonia'	758	684	38	11
	10. 'apathy scales'	0	0	0	0
	11. 'Apathy Evaluation Scale'	19	19	7	2
'Apathy Evaluation Scale' combined with each key term	1 + 11	21	19	7	2
	2 + 11	1	2	0	0
	3 + 11	0	0	0	0
	4 + 11	0	0	0	0
	5 + 11	2	3	1	0
	6 + 11	0	0	0	0
	7 + 11	0	0	0	0
	8 + 11	0	0	0	0
	9 + 11	2	1	0	0
	10 + 11	0	0	0	0
Total		26	25	8	2
Number of items of literature unique to the database		4	25	2	0
Studies that used the Clinician version of the AES		2	16	0	0
Comparative studies that report on some aspect of the psychometric properties of the AES-C		2	10	0	0

The 12 articles were used in this critical appraisal of the AES-C in addition to our recent article now in press (Clarke et al., 2007).

psychometric properties of the scale, was included. Prior to discussing the psychometric properties, a description of the administration, coding, and scoring of the scale as well as the item reduction process are presented.

Administration, coding and scoring

The AES-C is a 10- to 20-minute semi-structured interview, administered by a trained interviewer with a minimum of two years clinical experience in a psychological/psychiatric setting (Marin et al., 1991; Marin et al., 1993; Marin, 2002). It has two open-ended and 18 close-ended questions that inquire about the patients' thoughts, feelings and actions over the previous four weeks with scores ranging from 18 to 72 (Marin et al., 1991, 2002). Information from the open-ended questions (such as number of items reported, details reported for each item) serves as a guide for the rating of the 18 close-ended questions that follow (Marin et al., 1991, 2002).

Fifteen of the 18 close-ended items are written with positive syntax and rated on a scale of 1 to 4, with 1 = not at all characteristic, 2 = slightly characteristic, 3 = somewhat characteristic, 4 = very characteristic (Marin et al., 1991). The remaining three items are written with negative syntax and rated as 4 = not at all characteristic, 3 = slightly characteristic, 2 = somewhat characteristic and 1 = very characteristic (Marin et al., 1991). The interviewer is then required to reverse the coding for each item with positive syntax (1 = 4, 2 = 3, 3 = 2, 4 = 1) so that higher scores denote greater levels of apathy (Marin et al., 1991; Marin, 2002).

The rating of the scale is based primarily on the clinician's judgement of the patient's verbal and non-verbal expressions and responses throughout the interview, with the exception of the self-evaluation items (#3, #8, #13, #16) (Marin et al., 1991, 2002). Although description of some verbal and non-verbal expressions that should be used to guide the rating of apathy are provided (Marin et al., 1991, 2002), a more exhaustive and standardized list would be beneficial.

Bachelor's level training, with a minimum of two-years clinical experience in psychological/psychiatric settings, and 4 to 6 hours of exposure to the concept of apathy and the AES-C has been indicated as sufficient for its reliable use (Marin et al., 1991, 2002). We observed that a longer period of training was often required for bachelor's level raters, but 4 to 6 hours was sufficient for master's level raters with a degree in psychology/applied health sciences. Given Marin's

observation that 'inter-rater agreement can be reached by experienced clinicians by rating as few as 5–10 patients' (Marin et al., 1991), it might be more appropriate that the interviewer be an experienced clinician or an assessor who knows the patient well or who observes the patient's performance at different tasks.

A weakness of the scale relates to the self-evaluation items, which have to be scored based on the patient's exact responses. This increases the chance of false negatives with scores that are at or close to the cut-off, particularly for patients who lack insight into their problems and are likely to underestimate their level of apathy, which is common in dementia and often seen together with apathy (Marin et al., 1991, 1993).

Item reduction

Marin et al. (1991) explained that 'several hundred items' were originally generated for the AES scale. In accordance with the psychometric literature (Streiner and Norman, 1995a; Marx et al., 1999), a combination of clinical judgement and statistical approaches were used to reduce the original number to the final 18-item version of the scale (Marin et al., 1991; Marin, 2002).

Factor analysis and correlation analysis (Kim and Mueller, 1978a, 1978b; Gorsuch, 1983) were used in the ensuing stages of item reduction (Marin et al., 1991). In particular, principal components factor analysis with orthogonal rotation (Kim and Mueller, 1978a, 1978b; Gorsuch, 1983) using a sample size of 123 patients was employed in the item-reduction process. This showed that the AES-C comprised three factors: general apathy, interest and lack of insight/concern of which general apathy accounted for most of the variance (Marin et al., 1991). However, as the authors hypothesized a three-factor solution a priori, a more appropriate analytic strategy would have been confirmatory factor analysis or alternatively principal axis factoring, because in scale development one is interested only in the variance common among the items. We reassessed the factor structure of the AES-C in a recent study (Clarke et al., 2007) using principal axis factoring and found that the scale comprised two factors: general apathy and interest. Consistent with Marin et al. (1991), general apathy accounted for most of the variance (42.4%, Clarke et al., 2007).

Reliability

Sources of error, random or systematic, that may affect the reliability of a measure include the patients, the

measure itself and the interviewers (Portney and Watkins, 1993; Streiner and Norman, 1995b). As such, assessment of the various aspects of reliability of a new instrument is important. Reliability refers to the degree to which an instrument is free from random errors (Guilford, 1954; Guilleksen and Messick, 1960; Sheatsley, 1983; Portney and Watkins, 1993). The ability to produce reproducible or stable results on the same individuals over time is termed *test-retest reliability* and the ability to produce such results by two or more raters of the same individuals at one point in time is termed *inter-rater reliability* (Guilford, 1954; Guilleksen and Messick, 1960; Portney and Watkins, 1993; Streiner and Norman, 1995b). Internal consistency, which refers to the homogeneity of the items of a scale and the average correlations among the items (Guilford, 1954; Guilleksen and Messick, 1960; Portney and Watkins, 1993; Streiner and Norman 1995b; Cortina, 1993), is also a measure of reliability. Table 2 outlines the evidence of the assessment of the reliability of the AES-C.

Marin et al. reported high internal consistency for the AES-C (Cronbach's α (α) = 0.90 (1991) and α = 0.94 (1993)), thereby concluding that the scale was homogeneous and reliable (see Table 2). Similarly high internal consistencies have been reported across subsequent studies despite variations in the study samples. For example, internal consistencies of α = 0.84–0.89 in TBI (Andersson et al., 1999a, 1999b; Andersson and Bergendalen, 2002), α = 0.92 in PD (Pluck and Brown, 2002) and α = 0.89 in AD (Clarke et al., 2007). These values exceed the 0.8 that is indicative of good reliability (Cortina, 1993; Streiner and Norman, 1995b). However, the number of items on a scale is associated with alpha, with more than 15 items giving high values even if the scale is heterogeneous (Cortina, 1993; Portney and Watkins, 1993). Starkstein et al. (1992) who used an abridged version of the AES-C (14 of the 18 items) in PD reported good internal consistency (α = 0.76). We also observed high internal consistencies for the general apathy and interest factors, which were based on 11 items (α = 0.91) and three items (α = 0.86) respectively (Clarke et al., 2007). These latter results confirmed the high internal consistency of the AES-C without the issue of <15 items inflating alpha.

Internal consistency, however, is based on a single administration and does not account for day-to-day or rater-to-rater variations. It may lead to overly optimistic interpretation of the reliability of the instrument and

therefore should be interpreted with caution (Streiner and Norman, 1995b). Measures of the stability of the instrument, which include the inter-rater and test-retest reliability, and the ability of the scale to discriminate between people, are more informative and indicative of the reliability of the instrument (Portney and Watkins, 1993; Streiner and Norman, 1995b, 2003).

Marin et al. (1991) used correlation analysis in testing the *test-retest reliability* of the AES-C. They reported reliability coefficient of 0.88, which indicated good test-retest reliability. An assessment of the *inter-rater reliability* of the AES-C also indicated a high degree of correspondence and agreement between the raters (ICC = 0.94: Marin et al., 1991). The observed intra-class correlation coefficient (ICC), which is calculated using variance estimates through analysis of variance (Portney and Watkins, 1993; Streiner and Norman 1995a), was interpreted as evidence that many raters can be trained to use the scale in a similar fashion (Marin et al., 1991). These aspects of reliability have not been replicated in subsequent studies but merit further investigation, especially as reliability is not a fixed property of the scale but depends on the sample (Streiner and Norman, 1995b).

Validity

The validity of an instrument refers to the degree of confidence we can place in the inferences drawn from the test scores (Guilford, 1954; Guilleksen and Messick, 1960, Streiner and Norman, 1995c; Guyatt, 2000). That is, validity refers to what we can confidently say about the person on the basis of his or her test score. Validity can be assessed in different ways. One example is the assessment of whether the instrument contains a representative sample of items that truly reflect the construct – *content validity* (Guilford, 1954; Guilleksen and Messick, 1960; Marx et al., 1999; Streiner and Norman, 1995c). Methods commonly used to obtain evidence about the content validity include the use of lay and expert panels to judge the clarity, comprehensiveness and redundancy of the items (Streiner and Norman, 1995c; Marx et al., 1999).

Marin and colleagues (1991) explained that clinical judgement, consultation with other experts in the field, and pilot testing with 40 subjects aided in the selection and reduction of items to obtain the final version of the AES-C, thus attesting to its content validity. The consistent use of the scale across studies and health conditions lends support to its content validity, in that

Table 2. Prevalence of apathy and research evidence on the reliability of the Apathy Evaluation Scale-Clinician version (AES-C)

Author(s) and year	Study population	Cut -off	Findings
Marin et al., 1991*	N = 123, 53–85 years old. NL, CVA with RHD or LHD, AD and MD. These were community-dwelling seniors	<370.5	<ul style="list-style-type: none"> – Internal consistency (Cronbach's alpha = 0.90) – Test-retest reliability (Pearson's correlation coefficient = 0.88) – Inter-rater reliability (Intraclass correlation coefficient = 0.94) – Kappa: 2 raters across the 18 items of the AES-C ranges from 0.45–0.85 with mean kappa = 0.58 – Internal consistency (Cronbach's alpha = 0.94) – Inter-rater reliability
Marin et al., 1993	N = 107 community-dwelling, 55–85 yrs. old. NL, CVA with RHD or LHD, AD and MD	<370.5	<ul style="list-style-type: none"> – Internal consistency (Cronbach's alpha = 0.90) – Test-retest reliability (Pearson's correlation coefficient = 0.88) – Inter-rater reliability (Intraclass correlation coefficient = 0.94) – Kappa: 2 raters across the 18 items of the AES-C ranges from 0.45–0.85 with mean kappa = 0.58 – Internal consistency (Cronbach's alpha = 0.94) – Inter-rater reliability
Andersson et al., 1999b	N = 30, age 16–64, TBI inpatients	<34	<ul style="list-style-type: none"> – Apathy prevalent in TBI (66.7% but should be 33.3% (10/30)) – Cronbach's alpha (i.e. internal consistency) of AES-C = 0.89 – Cronbach's alpha for AES-C = 0.84
Andersson et al., 1999a	N = 72, age 16–60, TBI (n = 28), CVA (n = 30) and HBD (n = 14)	<34	<ul style="list-style-type: none"> – Apathy prevalent in TBI (66.7% but should be 33.3% (10/30)) – Cronbach's alpha (i.e. internal consistency) of AES-C = 0.89 – Cronbach's alpha for AES-C = 0.84
Andersson and Bergedalen, 2002	N = 53, traumatic brain injury	ns	<ul style="list-style-type: none"> – Cronbach's alpha for AES-C = 0.87
Pluck and Brown, 2002	N = 62, 45 Parkinson's disease (PD) and 17 Osteoarthritis (OA)	<38	<ul style="list-style-type: none"> – Prevalence of apathy among PD patients = 37.8% – Internal consistency using only the PD group for the AES-C was high (Cronbach's alpha = 0.92)
Clarke et al. (2007)	N = 121 community-dwelling, age = 48–91, memory complaints	<400.5	<ul style="list-style-type: none"> – AES-C specificity = 58.0%, sensitivity = 85.7%, PPV = 0.45, and NPV = 0.91 – Internal consistency by factors: apathy factor = 0.90 and interest factor = 0.86

* The original article on the development and validation of the Apathy Evaluation Scales. (AES-C, AES-S, AES-L)

† NL = normal well elderly; CVA = cerebrovascular accident; LHD = left hemispheric damage; RHD = right hemispheric damage; AD = Alzheimer's disease; MD = major depression; TBI = traumatic brain injury; ABI = acquired brain damage; HBD = hypoxic brain damage; SCD = subcortical damage; PPV = positive predictive value; NPV = negative predictive value.

‡ ns = not stated.

experts seem to agree that it measures the construct of apathy.

The assessment of whether the instrument 'behaves' in the way expected (*construct validity*) (Guilford, 1954; Guilleksen and Messick, 1960; Streiner and Norman, 1995c) is another way of assessing the validity of an instrument. That is, whether the scale is closely related to other variables and measures of the same construct (*convergent validity*) and does not correlate too strongly with measures of different construct (*discriminant validity*) (Guilford, 1954; Guilleksen and Messick, 1960; Streiner and Norman, 1995c; Guyatt, 2000). When the phenomenon being measured is a 'hypothetical construct', as is apathy (Marin, 1990; Marin, 1991; Stuss et al., 2000), it becomes important to test the construct validity of the resulting instrument (Guyatt, 2000).

Marin et al. (1991) indicated that they assessed the construct validity, specifically the convergent and discriminant validities, of the AES-C using the multitrait-multimethod matrix (MTMM) criteria. The MTMM is a statistically sound method of assessing the convergent and discriminant validity of an instrument and involves the use of two or more measures of each trait (Campbell and Fiske, 1959; Guyatt, 2000). However, with regard to the convergent validity of the AES-C, the researchers violated this criterion for the use of the MTMM because they had only one measure of apathy with three different modes of administration – the self-rated (AES-S), informant-rated (AES-I) and clinician-rated (AES-C) versions. Similarly, Pluck and Brown (2002) reported good convergent validity for the AES-C in a sample of individuals with PD by using the AES-S as the other measure of apathy (Table 3). The three versions of the AES contain the same 18 items with different methods of administration, which probably accounted for the high correlation between them (Marin et al., 1991; Pluck and Brown, 2002) (Table 3). The lack of appropriate assessment tools for apathy at the time might explain the use of the three versions of the scale in the determination of convergent validity.

Subsequent studies provided evidence on the convergent validity of the AES-C. For example, in their follow-up study, Marin et al. (1993) observed statistically significant correlation between AES-C scores and psychomotor retardation, lack of energy and insight and diminished work/interest, which are negative symptoms of depression as measured by the Hamilton Depression Rating Scale (HamD). Similarly, Andersson and others (1999a) found significantly high correlation

($r = 0.62$) between apathy and negative symptoms of depression (reduced initiative, lack of emotional responsivity and inattention) on the Montgomery and Asberg Depression Rating Scale (MADRS) despite using a sample that differed in diagnoses and age range. Kiang et al. (2003) reported significant correlation between the AES-C and the emotional withdrawal item of the Positive and Negative Syndrome Scale ($r = 0.57$) in a sample of individuals with schizophrenia or schizoaffective disorder. Other studies by Lampe et al. (2001) and Lampe and Heeren (2004) confirmed statistically significant and positive correlation between AES-C scores and negative symptoms of depression. Negative symptoms are hypothesized to be related to the same amotivational disorder as apathy (Andreasen, 1990) and therefore statistically significant correlations between them and apathy are indicative of good convergent validity across studies. In our recent study (Clarke et al., 2007), moderate and statistically significant correlation ($r = 0.30$) between the AES-C and the Apathy subscale of the Neuropsychiatric Inventory (NPI) developed by Cummings et al. (1994) was observed indicating good convergent validity between two measures that were developed to characterize and quantify apathy.

Depression, anxiety and apathy represent disorders with overlapping symptoms but with different constructs (Marin et al., 1991). As such, Marin et al. (1991) emphasized the importance of the AES-C not correlating too strongly with measures of depression and anxiety and hypothesized a priori that this would be confirmed in their study. As outlined in Table 3, moderate but statistically significant correlations were found between the AES-C and the HamD ($r = 0.39$) and the AES-C and the Hamilton Rating Scale for Anxiety ($r = 0.35$) indicating adequate discriminant validity (Marin et al., 1991). More convincing evidence of the discriminant validity of the AES-C was observed in subsequent studies. The AES-C had low correlations that failed to attain statistical significance with the depressed mood, guilt or hopelessness, suicide and vegetative symptoms of depression as per the HamD (Marin et al., 1993), with the depressed mood symptoms on the MADRS (Andersson et al., 1999a; Lampe and Heeren, 2004) and with depression as measured by the Calgary Depression Scale for Schizophrenia (CDSS: Kiang et al., 2003), the NPI depression subscale and the Structured Clinical Interview for DSM-IV (Clarke et al., 2007). The variations in the sample composition

Table 3. Research evidence on the validity of the Apathy Evaluation Scale-Clinician version (AES-C)

Author(s) and year	Study population	Cut-off	Findings
Marin et al., 1991*	N = 123, 53–85 years old. NL, CVA with RHD or LHD, AD and MD. These were community-dwelling seniors	<370.5	<ul style="list-style-type: none"> – AES-C converges with AES-S ($r = 0.72$, $p = 0.001$) and AES-I ($r = 0.62$, $p = 0.001$) – AES-C showed discriminant validity by the low correlation between apathy and depression ($r = 0.39$, $p > 0.001$) and anxiety ($r = 0.35$, $p > 0.01$) respectively – Criterion validity: AES-C differentiates between different diagnostic groups ($F(4,105) = 17.20$, $p > 0.001$), with post hoc test showing <ol style="list-style-type: none"> 1. NL (mean AES-C = 26.0 ± 6.2) > AD (mean AES-C = 44.4 ± 11.1): [$t(105) = -7.43$, $p > 0.05$]; 2. NL (mean AES-C = 26.0 ± 6.2) > MD (mean AES-C = 40.5 ± 9.7): [$t(105) = -6.22$, $p > 0.05$]; 3. NL (mean AES-C = 26.0 ± 6.2) > RHD(mean AES-C = 34.7 ± 7.3): [$t(105) = -3.34$, $p > 0.05$] but 4. NL (mean AES-C = 26.0 ± 6.2) versus LHD (mean AES-C = 31.9 ± 9.6) and RHD versus LHD not different ($p < 0.05$). – AES-C predicts average score on PACman video game ($r = -0.45$, $p = 0.001$), average latency for play ($r = 0.45$, $p = 0.001$), level of maze difficulty ($r = -0.37$, $p = 0.002$) percentage of total time ($r = -0.40$, $p = 0.001$) and the average time per game ($r = -0.24$, $p = 0.001$)
Starksrein et al., 1992	N = 50, patients with idiopathic PD	<14	<ul style="list-style-type: none"> – AS (abridge AES-C) showed positive correlation ($p > 0.05$) between depression (mean AS = 10.0 ± 2.0) and apathy (mean AS = 17.1 ± 4.0)
Marin et al., 1993	N = 107, 53– 5 yrs. old. NL, CVA with RHD or LHD, AD and MD. Communitydwelling	<370.5	<ul style="list-style-type: none"> – AES-C showed high correlation with psychomotor retardation ($r = 0.64$, $p > 0.001$), lack of energy ($r = 0.39$, $p > 0.001$) and insight ($r = 0.47$, $p > 0.001$), and diminished work/interest ($r = 0.50$, $p > 0.001$) items of HamD, shows convergent validity – AES-C showed discriminant validity by low correlation with depressed mood ($r = 0.10$, $p < 0.05$), guilt or hopelessness ($r = 0.04$, $p < 0.05$), suicide ($r = 0.20$, $p < 0.05$), as well as vegetative symptoms (DEA ($r = 0.10$, $p < 0.05$), MCA ($r = 0.01$, $p < 0.05$), EMA ($r = 0.16$, $p < 0.05$) – Criterion validity: AES-C differentiates between different diagnostic groups ($F(4,102) = 14.20$, $p > 0.001$) with: <ol style="list-style-type: none"> 1. No report on the result of post-hoc multiple comparison but AD (mean AES-C = 43.1 ± 10.1) < MD (mean AES-C = 38.7 ± 9.4) < RH (AES-C = 34.3 ± 8.7) < LH (mean AES-C = 30.9 ± 10.5) < NL (mean AES-C = 26.0 ± 5.6).

Marin et al., 1994	N = 125, 53–85 yrs. old. NL, CVA with RHD or LHD, AD and MD. Communitydwelling	<370.5	<ul style="list-style-type: none"> – Apathy and depression converges on negative symptoms (i.e. large and significant correlations) – AES-C showed discriminant validity by low correlation between apathy and depression for the NL ($r = 0.26$, $p < 0.05$) and RHS ($r = 0.20$, $p < 0.05$) groups but was correlated for LHS ($r = 0.49$, $p = 0.04$), AD ($r = 0.54$, $p = 0.01$), and MD ($r = 0.64$, $p > 0.001$) – Criterion validity: AES-C differentiates between diagnostic groups ($F(4,118) = 18.86$, $p > 0.001$) with post hoc test showing: <ol style="list-style-type: none"> 1. NL (mean AES-C = 26.0 ± 5.6) > AD (mean AES-C = 45.2 ± 11.0): [$t(118) = -1.78$, $p > 0.05$] 2. NL (mean AES-C = 26.0 ± 5.6) > MD (mean AES-C = 39.8 ± 9.1): [$t(118) = -1.81$, $p > 0.05$] 3. RHD (mean AES-C = 34.0 ± 8.4) > MD (mean AES-C = 39.8 ± 9.1): [$t(118) = -1.70$, $p > 0.05$] – The 18-item AES correlated highly with the seven-item AES ($r = 0.87$, $p > 0.05$) which shows convergent validity – A positive and significant correlation between depression and apathy ($r = 0.42$, $p = 0.001$) – A negative and significant correlation between cognitive impairment (MMSE) and apathy (AES-C) ($r = -0.56$, $p = 0.02$) – AES-C differentiates between depressed and cognitively impaired individuals – AES-C significantly correlated with the negative symptoms on the MADRS ($r = 0.62$, $p > 0.001$) = good convergent validity – AES-C not significantly correlated with the depressed mood symptoms on the MADRS ($r = 0.17$, $p < 0.05$) indicates good discriminant validity – Significant inverse relationship between apathy and heart rate reactivity ($r = -0.45$, $p = 0.001$) – Criterion validity: AES-C differentiates between different diagnostic groups: <ol style="list-style-type: none"> 1. HBD, TBI and CVA: ($F(2,70) = 4.68$; $p > 0.01$) and post-hoc test showed that HBD (mean AES-C = 34.6 ± 6.03) < TBI (mean AES-C = 35.5 ± 6.80) and CVA (mean AES-C = 34.3 ± 7.20) at $p > 0.05$ but no difference between TBI and CVA 2. RHD, LHD, HBD and SCD: ($F(3,69) = 7.94$; $p > 0.001$) and post-hoc test showed that RHD(mean AES-C = 37.9 ± 6.13), SCD (mean AES-C = 41.0 ± 6.43) < LHD (mean AES-C = 30.7 ± 6.93) and SCD (mean AES-C = 41.0 ± 6.43) < HBD (mean AES-C = 34.6 ± 6.03) at $p > 0.05$ but no difference between RHD and SCD or between RHD and HBD
Resnick et al., 1998	N = 102, mean age = 77.1 $\pm$ 7.9 yrs. Geriatric rehab. pts with hip fracture, stroke, amputation or hip replacement	Ns	
Andersson et al., 1999a	N = 72, 16–60 yrs. old. Acquired brain injured (ABI) patients, including TBI ($n = 28$), CVA ($n = 30$), hypoxic Brain Damage (HBD; $n = 14$), with LHD, RHD, and SCD	<34	

Table 3. Continued

Author(s) and year	Study population	Cut-off	Findings
Andersson et al., 1999b	N = 30, 16–64 years old. TBI inpatients	<34	<ul style="list-style-type: none"> Significant inverse relationships between apathy (AES-C) and heart rate reactivity ($\beta = -1.29, p = 0.003$), systolic blood pressure reactivity ($\beta = -1.65, p = 0.003$), and mean arterial pressure ($\beta = -1.71, p = 0.01$) Significant inverse relationship between apathy (AES-C) and perceived emotional discomfort ($t = 2.33, p = 0.03$)
Lampe et al., 2001	N = 126, aged 60+, 64 Normal healthy elderly: 62 psychiatric inpatients	Ns	<ul style="list-style-type: none"> AES-C strongly correlated with motivational retardation per CPRS ($r = 0.683, p > 0.001$) indicating good convergent validity Criterion validity: AES-C differentiates ($p > 0.001$) between psychiatric patients (mean AES-C = 43.40 ± 10.59 and healthy elderly (mean AES-C = 26.95 ± 6.52)
Andersson and Bergedalen, 2002	N = 53, in-patients with severe TBI with coma length < 24 hrs N = 62, 45 Parkinson's disease (PD) and 17 Osteoarthritis (OA)	Ns	<ul style="list-style-type: none"> AES-C showed high negative correlation with acquisition/memory (a variable that predicts apathy) ($r = -0.50, p > 0.005$) which indicates good convergent validity
Pluck and Brown, 2002		<38	<ul style="list-style-type: none"> AES-C and AES-S found to be highly correlated ($r = 0.74, p > 0.001$), which indicates good convergent validity Criterion validity: AES-C differentiate ($p > 0.001$) between PD (mean AES-C = 35.3 ± 11.4) and OA (mean AES-C = 23.3 ± 3.8)
Kiang et al., 2003	N = 28, 25 = schizophrenia and 3 = schizoaffective disorder	ns	<ul style="list-style-type: none"> AES-C strongly correlated with emotional withdrawal on PANSS-N scale ($r = 0.57, p = 0.002$) indicate convergent validity AES-C not significantly correlated with CDSS ($r = 0.09, p < 0.05$) which indicates strong discriminant validity
Lampe and Heeren, 2004	N = 29, Age ≥ 60 , in-patients with major depression or dysthymic disorder	<38	<ul style="list-style-type: none"> AES-C correlated with the negative symptoms on the MADRS ($r = 0.470, p > 0.01$) = good convergent validity AES not significantly correlated with the affective symptom score ($r = 0.170, p < 0.05$), the vegetative symptom score ($r = 0.190, p < 0.05$) or total MADRS-score ($r = 0.348, p = 0.064$) on the MADRS which indicate good discriminant validity
Clarke et al., 2007	N = 121, age = 48–91 Community dwelling, memory complaints	<400.5	<ul style="list-style-type: none"> Convergent validity (i.e. correlation with the apathy subscale of the NPI) $r = 0.30, p > 0.01$ Discriminant validity (i.e. convergence with the depression subscale of the NPI) $r = 0.18, p < 0.05$ and the SCID $r = 0.20, p < 0.05$

*The original article on the development and validation of the Apathy Evaluation Scales (AES-C, AES-S, AES-I).

†NL = normal well elderly; CVA = cerebrovascular accident; LHD = left hemispheric damage; RHD = right hemispheric damage; AD = Alzheimer's disease; MD = major depression; TBI = traumatic brain injury; ABI = acquired brain damage; HBD = hypoxic brain damage; SCD = subcortical damage; MMSE = Mini-Mental State Examination; MADRS = Montgomery and Asberg Depression Rating Scale; CPRS = Comprehensive Psychopathological Rating Scale; HamD = Hamilton Rating Scale; DFA = difficulty falling asleep; MCA = mid-cycle awakening; EMA = early morning awakening; CDSS = Calgary Depression Scale for Schizophrenia; SCID = structured clinical interview for DSM-IV; PANSS-N = Positive and Negative Syndrome Scale – Negative Symptoms.

‡ns = not stated.

across studies demonstrate that the evidence on the discriminant validity of the AES-C is quite satisfactory.

Scores on the AES-C had significant and inverse relationships with heart rate reactivity, diastolic blood pressure and mean arterial pressure, which are physical evidence of arousal (Andersson et al., 1999a, 1999b). High apathy scores were significantly and inversely related to perceived emotional discomfort (Andersson et al., 1999a, 1999b). Therefore, higher apathy scores were associated with low levels of physical and emotional reactivity and arousal, providing further support for the scale's construct validity.

The assessment of whether the scale is sensitive to different levels of the phenomenon is another method of investigating the validity of an instrument – *criterion validity* (Streiner and Norman, 1995c). From clinical observations, information from experts in the field and from the literature, investigators have postulated that some diagnostic groups would have higher levels of apathy based on the region of brain damage. For example, Marin et al. (1991, 1993, 1994) postulated *a priori* that individuals with AD, right hemispheric damage (RHD) and major depression (MD) would have higher levels of apathy than normal healthy elderly people, and that those with RHD would have higher levels of apathy than those with left hemispheric damage (LHD). In other words, the AES-C would differentiate between various diagnostic groups thereby providing evidence of its criterion validity. The results revealed that the AES-C differentiated between diagnostic groups based on different levels of apathy and more specifically indicated lower levels of apathy in normal healthy elderly compared to those with AD, RHD and MD respectively (Marin et al., 1991, 1994) and lower levels in those with RHD compared to those with MD (Marin et al., 1994). However, the scale did not differentiate between AD and RHD groups (Marin et al., 1991, 1993, 1994).

In their study from the Netherlands, Andersson et al. (1999a) found that the AES-C differentiated between diagnostic groups with individuals who had hypoxic brain damage (HBD) having significantly higher ($p > 0.05$) apathy scores compared to those with TBI and CVA respectively. As well, individuals with RHD and subcortical damage (SCD) respectively scored significantly ($p > 0.05$) higher than those with LHD and those with SCD scored greater than those with HBD ($p > 0.05$). Scores on the AES-C did not differentiate

between individuals with RHD and those with SCD or between the RHD and HBD groups. Lampe et al. (2001) also found that patients with psychiatric illnesses had higher mean AES-C scores than normal healthy elderly ($p > 0.001$), thereby confirming that the scale differentiated between diagnostic groups. These replication studies using different samples supported the ability of the AES-C to differentiate between some but not all diagnostic groups based on variation in mean scores.

Predictive validity refers to the tool's ability to predict future outcomes (Guilford, 1954; Guilleksen and Messick, 1960). Marin et al. (1991) proposed that an important feature of the AES-C is its ability to predict whether the subjects will self-initiate activities, the level of difficulty of the initiated activities and whether they will have sustained interest in activities once initiated. This was proposed based on the conceptualization of apathy as a neurobehavioural syndrome characterized by a lack of motivation, self-initiated activations and concerns and emotional indifference (Marin et al., 1991, 1993). It was therefore hypothesized that apathy would be negatively correlated with the number of cycles of play, the level of difficulty of the game played and the length of play time but would be positively correlated with response latency (Marin et al., 1991).

To investigate the predictive validity of the AES-C, subjects' behaviours were observed and recorded in situations that required them to initiate and terminate a number of activities. Subjects were seated and left in a waiting room with novel toys and games on a nearby coffee table and told that more time was needed to organize for the next test procedure and that the toys and games were there for them to use while waiting, if they so desired (Marin et al., 1991). The subjects' abilities to initiate and sustain play with video games of varying degrees of complexities, including Pac-Man, Maze and Slot Machine, were observed (Marin et al., 1991). Marin et al. (1991) explained that the video games were selected to present a range of difficulties so that impaired and unimpaired subjects would find at least one game to engage in. The results indicated that apathy was negatively related to the average score per Pac-Man video game ($r = -0.45$), positively correlated with the average latency for play with the Pac-Man video game ($r = 0.45$), negatively correlated with the level of maze difficulty ($r = -0.37$) and negatively correlated with the percentage total time ($r = -0.40$) and the average time per game ($r = -0.24$).

These findings were consistent with the conceptualization that apathy in that they may reflect lack of motivation and interest in the activities. The negative correlation with task difficulty and game score reached may also reflect altered level of consciousness and motivational capacity, which are key characteristics of apathy (Marin, 1990). Most difficult tasks require a higher level of attentional control (and therefore, greater initiation capacity and level of consciousness), and to be able to achieve good game scores requires an excellent continuous attention capacity (and motivation to keep up attention). However, Marin et al. (1991) indicated that several of the severe AD subjects were unable to learn any of the video games. If data on these subjects were included (no game initiated and zero play time), then the results obtained would have been greatly influenced by the participants' level of cognitive impairment.

Conceivably, concomitant motor conditions in individuals with brain damages can significantly affect functional limitations and subsequently affect choice of game/s played and time spent playing. However, information on the participants' level of functioning was not provided. As such, it is unclear whether the correlations between apathy and total time spent playing and average time per game were related to functional limitations, independent of apathy status. This is an important issue because poor performance in activities of daily living was found to be significantly related to high apathy levels (Resnick et al., 1998; Kiang et al., 2003; Clarke et al., 2007). More specifically, Kiang et al. (2003) observed no significant correlation between apathy and level of interest in playing a game in an observed setting in individuals with schizophrenia/schizoaffective disorder but found significant relationship between apathy and functional activities. These discrepant findings indicate that the predictive validity of the AES-C requires further investigation.

Conclusions

In conclusion, clearer guidelines, including a more standardized list of verbal and non-verbal apathetic cues, are needed for training, administration, and rating of the AES-C. The requirement that interviewers should be clinicians who have observed the patient's behaviour across different settings and tasks should also be considered.

Many studies have reported internal consistency coefficients that replicated Marin et al.'s (1991) original

findings attesting to the reliability of the scale. However, there is a lack of studies that have investigated the evidence on the test-retest and inter-rater reliabilities, which are more stable and informative measures of the reliability. As such, the only available data on these characteristic of reliability come from the original study by Marin et al. (1991). Given the fact that information on the reliability is sample dependent (Streiner and Norman, 1995b), these aspects of reliability warrant investigation, especially if the sample on which the scale is to be administered differs from the original sample on which the properties were determined. Despite this, the reliability of the AES-C appears quite satisfactory.

The evidence confirms that the AES-C has good discriminant, convergent and criterion validities. Although the original assessment of the convergent validity of the scale was based solely on the correlation between its three versions, subsequent studies have provided further supporting evidence. That is, convergence between the AES-C and the negative symptoms of depression (Marin et al., 1993, 1994; Andersson et al., 1999a, 1999b; Lampe et al., 2001; Lampe and Heeren, 2004) and schizophrenia (Kiang et al., 2003) as well as other measures of apathy (Clarke et al., 2007) were reported. Similarly, strong evidence of the scales discriminant and criterion validities was provided by Marin et al. (1991) as well as subsequent independent studies (Andersson et al., 1999a, 1999b; Lampe et al., 2001; Pluck and Brown, 2002; Lampe and Heeren, 2004; Clarke et al., 2007).

The evidence of the predictive validity of the AES-C was quite limited. Marin et al. (1991) found significant correlations between apathy and self-initiated activities in terms of choice and difficulty of video games played and the length of time spent and scores attained on those games. These findings were not replicated in the only study that reassessed the relationship between apathy and self-initiated activities (Kiang et al., 2003). The information presented in the original assessment of the predictive ability of the AES-C (Marin et al., 1991) also failed to clarify whether the relationship observed was related to functional impairment due to physical and cognitive decline independent of apathy status, which affected the strength of the evidence. Since this aspect of validity refers to the scale's ability to predict future outcomes, which is important for treatment and rehabilitation planning, further assessment of the predictive validity of the

AES-C is needed. However, the scale is a valid measure for the characterization and quantification of apathy.

Acknowledgements

Diana E Clarke is a post-doctoral fellow at the Toronto Rehabilitation Institute. Dr Clarke was supported by a Canadian Institute for Health Research Doctoral Award (Grant # 200110MDR-96269-115967), the Hy and Bertha Shore and Harry and Sara Gorman Award for Studies in Aging from the Baycrest Foundation, Baycrest Centre for Geriatric Care and the Cognition Team Excellence Award from the Toronto Rehabilitation Institute during the preparation of this article. Drs Van Reekum and Streiner were supported by the Kunitz-Lunenfeld Applied Research Unit of Baycrest Centre and are engaged in research supported by the Canadian Institute for Health Research (Grant #200204MCT-101404-CTA-ABPI-50298). Dr Simard was engaged in research supported by a 2005 NARSAD Young Investigator Award. Jigisha Patel was supported by the Work-study Program.

References

- Andersson S, Bergedalen AM. Cognitive correlates of apathy in traumatic brain injury. *Neuropsychiatry Neuropsychol Behav Neurol* 2002; 15: 184–91.
- Andersson S, Gundersen PM, Finset A. Emotional activation during therapeutic interaction in traumatic brain injury: effect of apathy, self-awareness and implications of rehabilitation. *Brain Inj* 1999b; 13: 393–404.
- Andersson S, Krogstad JM, Finset A. Apathy and depressed mood in acquired brain damage: relationship to lesion localization and psychophysiological reactivity. *Psychol Med* 1999a; 29: 447–56.
- Andreasen NC. *Schizophrenia: Positive and Negative Symptoms and Syndromes*, Karger, Basel, 1990.
- Campbell DT, Fiske DW. Convergent and discriminant validation by multitrait-multimethod Matrix. *Psychol Bull* 1959; 56: 81–105.
- Clarke DE, van Reekum R, Simard M, Streiner D, Conn D, Freedman M. Apathy in dementia: An examination of the psychometric properties of the Apathy Evaluation Scale. *Journal of Neuropsychiatry and Clinical Neuroscience*, 2007; 19: 57–64.
- Cortina JM. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol* 1993; 78: 98–104.
- Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology* 1994; 44: 2308–14.
- Gorsuch RL. *Factor Analysis*. Hillsdale, NJ: Erlbaum, 1983.
- Guilford JP. *Psychometric Methods*. 2 edn. New York: McGraw-Hill, 1954.
- Guilliksen H, Messick S. *Psychological Scaling: Theory and Application*, New York: John Wiley & Sons, 1960.
- Guyatt GH. Making sense of quality of life data. *Med Care* 2000; 38(supplement II): 175–9.
- Kane RA, Kane RL. *Assessing the Elderly: A Practical Guide to Measurement*. Toronto: Lexington Books, 13–17, 1981.
- Kant R, Duffy PD, Pivovarnik A. Prevalence of apathy following head injury. *Brain Inj* 1998; 12: 87–92.
- Kiang M, Christensen BK, Remington G, Kapur S. Apathy in schizophrenia: clinical correlates and association with functional outcome. *Schizophrenia Res* 2003; 63: 79–88.
- Kim J, Mueller CW. *Introduction to Factor Analysis: What It Is and How To Do It*. Thousand Oaks, CA: Sage Publications, 1978(a).
- Kim J, Mueller CW. *Factor Analysis: Statistical Methods and Practical Issues*. Thousand Oaks, CA: Sage Publications, 1978(b).
- Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Disease* 1985; 38: 27–36.
- Lampe IK, Heeren TJ. Is apathy in late-life depressive illness related to age-at-onset, cognitive function or vascular risk? *Int Psychogeriatr* 2004; 16: 481–86.
- Lampe IK, Kahn RS, Heeren TJ. Apathy, anhedonia, and psychomotor retardation in elderly psychiatric patients and healthy elderly individuals. *J Geriatr Psychiatr Neurol* 2001; 14: 11–6.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–74.
- Marin RS. Differential diagnosis and classification of apathy. *Am J Psychiatry* 1990; 147: 22–30.
- Marin RS. Apathy: A Neuropsychiatric syndrome. *J Neuropsychiatr Clin Neurosci* 1991; 3: 243–54.
- Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the Apathy Evaluation Scale. *Psychiatr Res* 1991; 38: 143–62.
- Marin RS, Firinciogullari S, Biedrzycki RC. The sources of convergence between measures of apathy and depression. *J Affect Disord* 1993; 28: 117–24.
- Marin RS, Firinciogullari S, Biedrzycki RC. Group differences in the relationship between apathy and depression. *J Nerv Ment Dis* 1994; 182: 235–9.
- Marin RS. *Guidelines for Apathy Evaluation Scale*. Western Psychiatric Institute and Clinic, Dittsburgh, DA. 2002.
- Marx RG, Bombardier C, Hogg-Johnson S, Wright JG. Clinimetric and psychometric strategies for the development of a health measurement scale. *J Clin Epidemiol* 1999; 52: 105–11.
- Pluck GC, Brown RG. Apathy in Parkinson's disease. *J Neurol Neurosurg Psychiatr* 2002; 73: 636–42.
- Portney LG, Watkins MP. Statistical measures of reliability. In LG Portney, MP Watkins (eds) *Foundation of Clinical Research: Applications to Practice*. 2 edn. Appleton & Lange, 2000.
- Rabkin JG, Ferrando SJ, Van Gorp W, Rieppi R, McElhiney M, Sewell M. Relationships among apathy, depression,

- and cognitive impairment in HIV/AIDS. *J Neuropsychiatry Clin Neurosci* 2000; 12: 451–57.
- Resnick B, Zimmerman SI, Magaziner J, Adelman A. Use of the Apathy Evaluation Scale as a measure of motivation in elderly people. *Rehabil Nurs* 1998; 23: 141–47.
- Sheatsley PB. Questionnaire construction and item writing. In PH Rossi, JD Wright, AB Anderson (eds). *Handbook of Survey Research*. New York: Academic Press, 1983.
- Starkstein SE, Mayberg HS, Preziosi TJ, Andrezejewski P, Leiguarda R, Robinson RG. Reliability, validity and clinical correlates of apathy in Parkinson's disease. *J Neuropsychiatry Clin Neurosci* 1992; 4: 134–39.
- Streiner DL, Norman GR. Devising the items. In DL Streiner, GR Norman (eds) *Health Measurement Scales: A Practical Guide to their Development and Use*. 2 edn. New York: Oxford University Press, 1995(a).
- Streiner DL, Norman GR. Reliability. In DL Streiner, GR Norman (eds) *Health Measurement Scales: A Practical Guide to their Development and Use*. 3 edn. New York: Oxford University Press, 1995(b).
- Streiner DL, Norman GR. Selecting the items. In DL Streiner, GR Norman (eds) *Health Measurement Scales: A Practical Guide to their Development and Use*. 3 edn. New York: Oxford University Press, 2003.
- Streiner DL, Norman GR. Validity. In DL Streiner, GR Norman (eds) *Health Measurement Scales: A Practical Guide to their Development and Use*. 2 edn. New York: Oxford University Press, 1995(c).
- Stuss DT, Van Reekum R, Murphy KJ. Differentiation of states and causes of apathy. In DT Stuss (ed.) *Neuropsychology of Emotion*. New York: Oxford University Press, 2000.

Correspondence: Diana E Clarke, Post-doctoral Fellow in Psychiatric Epidemiology, Department of Mental Health, Johns Hopkins School of Public Health, Room 802, 624 N Broadway Baltimore, MD 21205.

Email dclarke@jhsph.edu

Telephone (+1) 410-955 0416