Barrows, Paul David and Thomas, Shirley A. (2017) Assessment of mood in aphasia following stroke: validation of the Dynamic Visual Analogue Mood Scales (D-VAMS). Clinical Rehabilitation . ISSN 1477-0873

Assessment of mood in aphasia following stroke: validation of the Dynamic Visual Analogue Mood Scales (D-VAMS)

Paul D. Barrows[1] and Shirley A. Thomas[1]

[1] Division of Rehabilitation and Ageing, University of Nottingham

Corresponding author:

Paul Barrows, Division of Rehabilitation and Ageing, University of Nottingham, School of Medicine, B Floor, The Medical School, Queen's Medical Centre, Nottingham NG7 2UH. Email: paul.barrows@dvams.org.

Abstract

Objectives: To validate a nonverbal self-report measure of mood – the Dynamic Visual Analogue Mood Scales (D-VAMS) – against the Hospital Anxiety and Depression Scale (HADS), and assess its suitability as an outcome measure or screening measure for depressed mood following stroke.

Design: Cross-sectional observational cohort study.

Participants: Forty-six stroke survivors (24% with aphasia) recruited from online, from stroke clubs and via an NHS rehabilitation service.

Methods: A set of seven bipolar scales was developed enabling users to report mood by modifying facial expression images using a slider. Participants completed a tablet/computer task, reporting their mood on these scales mixed randomly with versions which used only words. The HADS was then completed, followed by a repeat run of the two versions in a different, random sequence.

Results: Exploratory factor analysis identified one factor consistent with pleasantness of mood accounting for 80% of the variance. Internal consistency of D-VAMS was high ($\alpha=0.95$) and there was a high correlation between face-only D-VAMS scores and HADS total scores ($r=-0.80$, $p<0.001$) as well as HADS-D/HADS-A subscale scores ($r=-0.73$, $p<0.001$; $r=-0.71$, $p<0.001$). D-VAMS showed good sensitivity and specificity against HADS, with means of 85%/77% (sensitivity/specificity) against the HADS-D and 80%/77% against the HADS-A across nine cut-offs.

Conclusions: D-VAMS is a valid and reliable measure likely suitable for assessment of depressed mood in aphasia following stroke. Though D-VAMS performed well as a screening measure in this study sample, further study is needed in the acute stage post-stroke.

Keywords

Aphasia, stroke, depression, assessment, non-verbal

Introduction

Stroke survivors are at high risk of depression,[1] and this may impact substantially upon their recovery. Evidence suggests that people with communication problems due to aphasia following stroke are particularly at risk,[2] yet there are few instruments to assess mood in this population. Though some adapted self-report measures exist,[3,4,5,6] they have been found to be limited. Questions have been raised about their suitability for people with aphasia and cognitive impairments,[7,8] their robustness as a screening measure for depression[9,10] and the quality of methodology employed in validation studies in which they were assessed.[11]

To circumvent these communication difficulties, observer-rated instruments have provided much needed, alternative means of assessing depression based on observable behaviours[12,13,14,15] and the Stroke Aphasic Depression Questionnaire has an evidence base supporting its utility as a screening measure for depression following stroke.[9,16] However correlations between observer-rated and self-reported mood have proven unreliable,[8,17,18] and self-report is an important source of information about mood. There is therefore a need for better self-report measures suitable for use with stroke patients with aphasia.

This article briefly details the construction of a nonverbal mood measurement instrument, and then describes a validation study examining the instrument's psychometric properties. Its suitability for use with people with aphasia following stroke was then assessed, both as a general outcome measure and as a screening instrument for depression or low mood. Ideally such a validation study should include only people with aphasia; however, people with aphasia often cannot use the language-based instruments that are required as a criterion measure. Since the ability to use an appropriate criterion measure is a priority for a validation study, stroke survivors whose language ability was largely intact were used in this sample. The implications of this will be discussed.

Methods

Development of Nonverbal Mood Scales

To address the problems of language difficulties following stroke, a novel, tablet/computer based instrument was developed. The instrument is based on a visual analogue scale (VAS), which is commonly used to rate subjective phenomena like mood or pain. A VAS consists of a 100-mm line between two dichotomous, end-point descriptors, which a rater marks to report a score.

The design of the instrument was guided by a Circumplex Model of Affect[19,20] in which mood is viewed as consisting of two main factors, valence and activation. Based on exploratory studies of judgements of facial expression,[21] seven labelled, bipolar scales were selected, each representing a different trajectory across affect space as described by plots of factor loadings from a principal component analysis (PCA). The scales are (1) *Miserable–Satisfied*, (2) *Sad–Happy*, (3) *Distressed–Peaceful*, (4) *Bored–Excited*, (5) *Afraid–Calm*, (6) *Angry–Peaceful*, and (7) *Sleepy–Alert* (Figure 1).

Photographs of posed facial expressions transitioning between the end-points of each scale (Figure 2) were then used to generate continua of morphed images each corresponding to a score in the range of 0–100, with 0 marking the negative valenced (or in the case of Scale 7, lowest activated) end of the scale and 100 marking the positive valenced (or for Scale 7, highest activated) end of the scale.

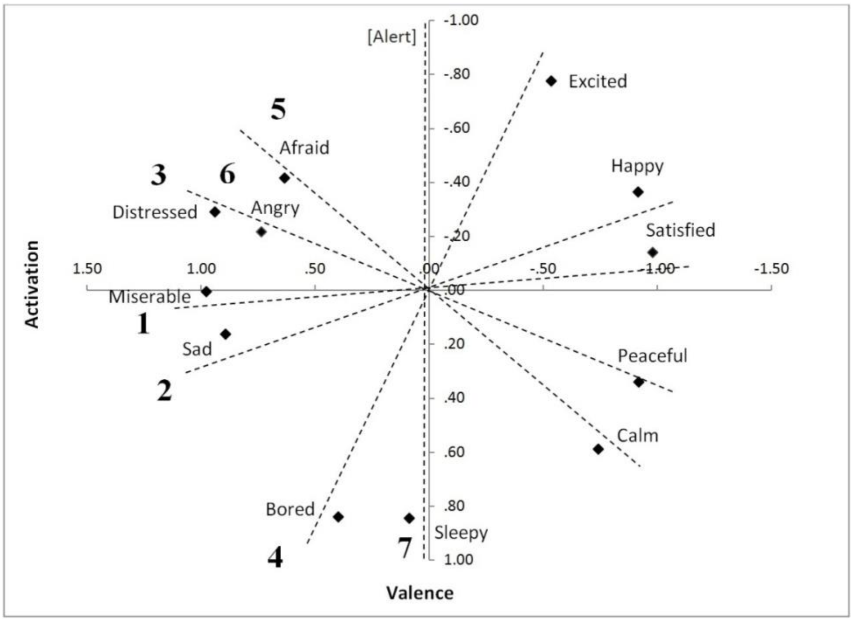

Figure 1. The seven bipolar scales of the D-VAMS charted as trajectories across affect space, delineated by a plot of factor loadings from a judgement study of facial expressions [21]. The first six scales have endpoints loading both on valence and activation, whereas the seventh appears valence-neutral, loading primarily on the activation dimension.

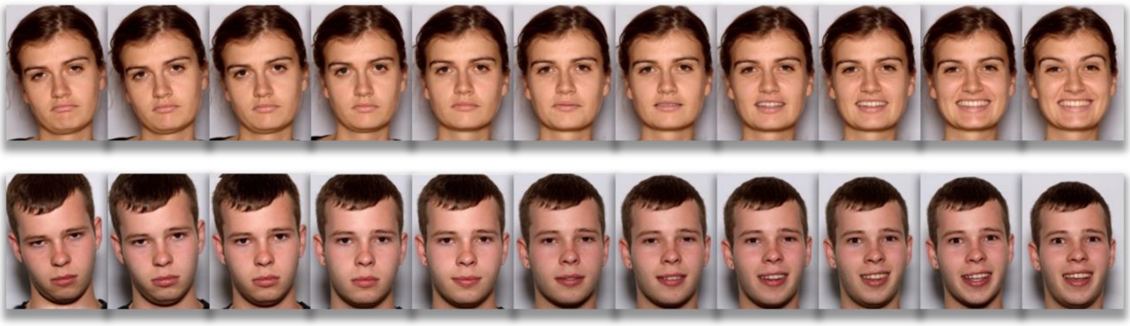

Figure 2. Images of posed transitional facial expressions along the *Bored-Excited* scale (top:female; bottom: male). These and others were used to generate morphed images along each scale in increments of 0–100 using data from a scaling study [21].

An interface was devised that would allow a respondent to report their mood by modulating a facial expression image on a web page (Figure 3(a)). On the left side of the page, a large image is displayed which is initially set to a 'neutral' facial expression corresponding to the midpoint of a given scale. To the right of the page is a vertical slider marked with graduated increments like those along the edge of a ruler, with the slider position initially set to the midpoint. At the top and bottom of the page are the words for the scale endpoints. As the slider is moved up or down, the facial expression image changes to the one corresponding to the current slider position, so that the expression appears to change smoothly from one extreme of the scale to the other. In this way, a respondent can adjust the image, selecting a face that most closely reflects their own mood on a given scale. For this prototype of the instrument – Dynamic Visual Analogue Mood Scales (D-VAMS) – each scale is presented on seven consecutive pages, finishing with a results page in which scores and summary statistics are charted and the selected faces displayed (Figure 3(b)). One version of the scales was created using a female face, and a second using a male face.

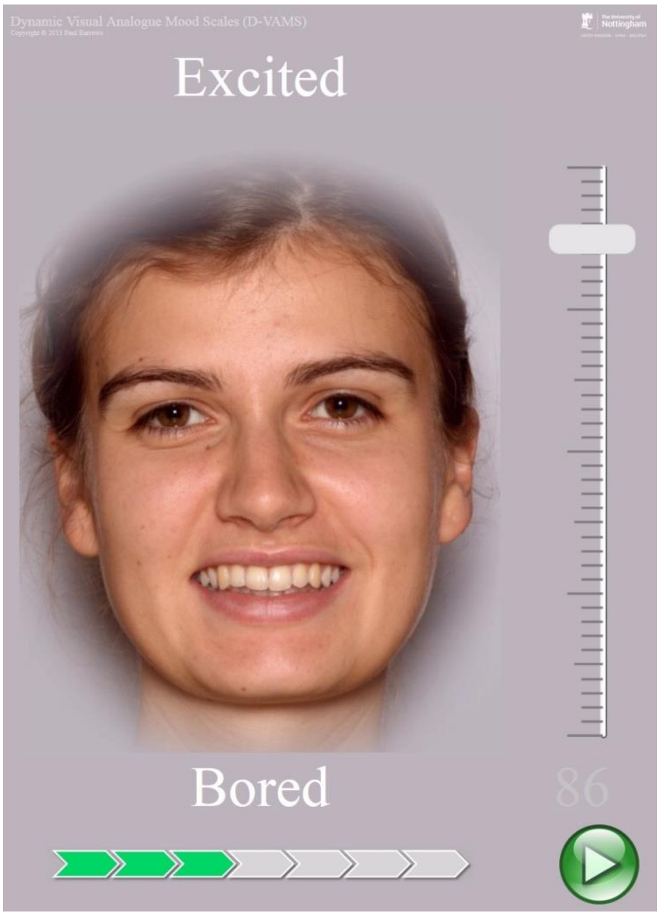

Figure 3(a). The D-VAMS interface (Scale 4: ‘Bored-Excited’; female face). A slider (right) is adjusted to select an expression that most closely reflects a user’s mood during the previous week.

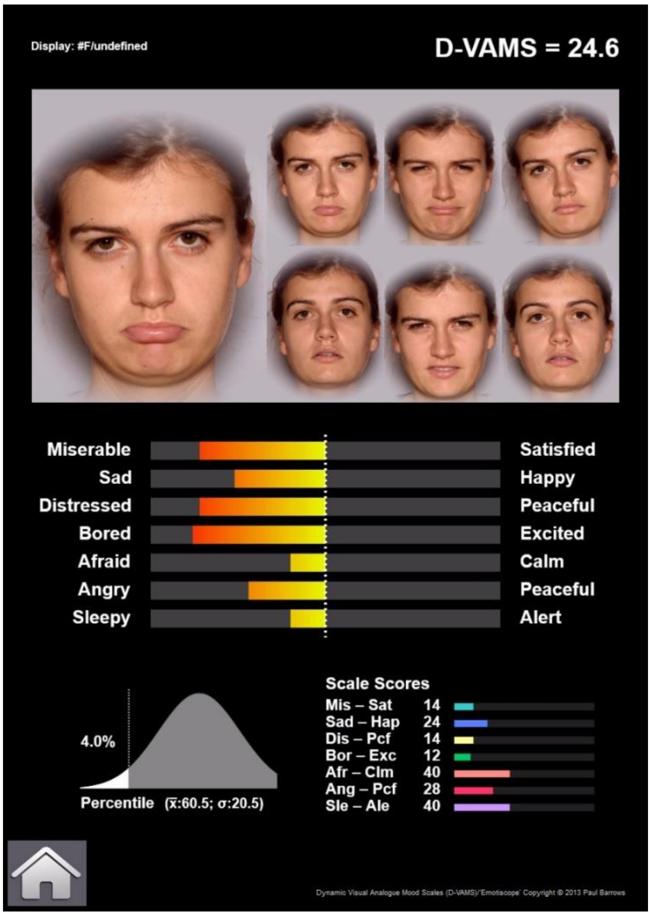

Figure 3(b). The D-VAMS interface results page. The mean score is presented top right, underneath which faces selected from the seven scales are displayed. Scores are represented on bipolar bar displays with coloured gradations from red (negative) through to yellow (neutral) and green (positive). Unipolar score values are displayed as bar charts (bottom right). A bell curve offers normative guidance based on configurable population means and SD.

Validation Study Task

To examine the psychometric properties of these scales, a task was devised employing two versions of the scales, and using the Hospital Anxiety and Depression Scale (HADS) as a comparison measure. For the first, *face-only* version of the scales, the mood words (such as ‘Happy’ and ‘Sad’) were omitted from the top and bottom of the page, so that no language cues were present. For the second, *word-only* version, no facial expression image was included; there were only the mood words for each scale presented at either end of a vertical slider in the middle of the page.

The task consisted of three parts in which participants were asked to rate their own mood during the last week (Figure 4). For the first part of the task, participants responded to both versions of the seven scales presented on 14 consecutive pages in random sequence (‘Run 1’). For the second part of the task, participants responded to 14 items comprising the HADS.[22] For the final part of the task, the first part of the task was repeated, but with both versions of the seven scales presented in a different, random sequence (‘Run 2’). The gender of faces used in the scales was alternated with each consecutive participant.

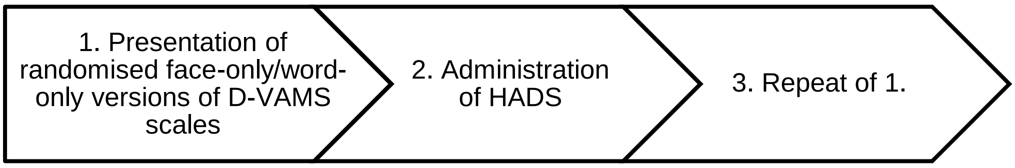

Figure 4. The three stage experimental protocol. Administration of face-only and word-only versions of Dynamic Visual Analogue Mood Scales in random sequence (1) is followed administration of the Hospital Anxiety and Depression Scales (2). This is then followed by a repeat of (1), but with the scales presented in a different random sequence.

Design and Participants

This was a cross-sectional observational cohort study. Power calculations were performed using G*Power.[23] Sample size was calculated based on a medium to large effect size (dz=0.35–0.45) for a power of 0.8 and an α of P<0.05 for an analysis of correlations (Pearson’s r, one-tailed); this gave an estimated target sample size in the range of 26–46. Ethical approval was granted by the Faculty of Medicine and Health Sciences Research Ethics Committee, University of Nottingham (Ref : I10102013) and by West of Scotland Research Ethics Service (WoSRES) (Ref: 15/WS/0239).

Participants were stroke survivors who were recruited via the internet, local stroke clubs, and an NHS rehabilitation service. All were English-speakers over 18 years of age who were capable of giving informed consent. For stroke club and rehabilitation group participants, the task was performed on a tablet supervised by a researcher at a stroke club meeting or at the participant’s home. For participants recruited online, the task was completed via the internet without supervision. Prior to beginning the experimental task, participants reported their gender, age, and the time elapsed since their stroke, and whether they had experienced any significant aphasia. Detailed medical information about participants was not available.

Statistical Analysis

SPSS version 22 (SPSS Inc., Chicago, IL, USA) was used to analyse the data. Two summary scores of the D-VAMS were examined: the mean score and the mean score excluding the activation-oriented seventh ‘*Sleepy-Alert*’ scale. Tests of distribution revealed no significant skew or kurtosis, and a Q-Q plot indicated approximately normal distribution, so parametric tests were used. Construct validity was determined by performing Pearson’s r correlations between the scores for the face and word versions of the scales. To examine convergent and discriminant validity between the scales, cross-scale correlations (Pearson’s r) were computed for face scale scores from Run 2. Factor structure was examined by performing principal axis factoring on scale scores, and internal consistency was assessed by computing their Cronbach’s α.

Criterion validity was determined by performing Pearson’s r correlations between face/word scale scores and HADS total and subscale scores. Test-retest reliability was assessed by intraclass correlations (two-way mixed) of scale scores and means between

Run 1 and Run 2. A comparison of means (related *t*-test, two-tailed) was performed between Run 1 and Run 2 scores to test for proportional bias.

To assess sensitivity and specificity, Receiver Operator Characteristics (ROC) were computed against HADS total score, and depression and anxiety subscale scores. Lincoln et al.[10] recommend cut-offs for stroke patients ranging from 4/5 to 7/8 for the depression subscale, and 4/5 to 5/6 for the anxiety subscale. However, higher cut-offs may be desirable in some circumstances, and this instrument may find use in other populations, so a larger range of cut-offs was analysed. Depression and anxiety subscale cut-offs ranging from 4–12 were therefore examined, along with even-numbered total cut-offs ranging from 8–24.

Results

Study sample characteristics for aphasia, age, time since stroke, HADS subscale scores and D-VAMS mean scores (with and without Scale 7) are shown in Table 1.

Table 1. Principle characteristics and test scores for three groups of stroke patients, separately and combined

Variables	Online	Stroke Club	Rehabilitation	ALL
Sample size, *n* (%) of total	15 (33)	20 (43)	11 (24)	46 (100)
With aphasia, *n* (%) of total*	5 (11)	3 (7)	3 (7)	11 (24)
Age in years, mean ±SD	48.1 ±9.3	72.2 ±9.7	70.5 ±11.4)	63.8 ±14.7
Gender, n (%) female	5 (33)	7 (35)	6 (54)	18 (39)
TSS (years), mean ±SD	3.0 ±1.6	5.2 ±4.5	0.6 ±0.4	3.4 ±3.6
HADS total, mean ±SD	18.1 ±6.5	12.1 ±7.2	16.8 ±7.8	15.3 ±7.2
HADS-D, mean ±SD	8.7 ±3.4	5.2 ±3.3	7.5 ±4.4	6.9 ±4.0
HADS-A, mean ±SD	9.4 ±4.8	7.0 ±4.2	9.3 ±3.7	8.4 ±4.7
D-VAMS[1], mean ±SD	52.5 ±18.9	71 ±18.1	52.1 ±19.4	60.5 ±20.5
D-VAMS[2], mean ±SD	54.5 ±20	71 ±18.2	50.9 ±20.4	60.8 ±21.0

TSS: Time since stroke; HADS-D: Hospital Anxiety and Depression Scale, depression subscale score; HADS-A: Hospital Anxiety and Depression Scale, anxiety subscale score; D-VAMS[1]: Dynamic Visual Analogue Mood Scales scores; D-VAMS[2]: Dynamic Visual Analogue Mood Scales scores excluding 'Sleepy-Alert' scale. * Total sample size.

Construct validity

Construct validity was reflected in the correlations between the face-only and word-only versions of the scales. Correlations by scale and by run are detailed in Table 2. The correlation for all of the scales combined was r=.59 for Run 1, and r=.76 for Run 2. The correlation for scores from both runs combined was r=0.66. Though construct validity by this measure was good, high intercorrelations were noted between face scale responses, and discriminant validity between the separate scales was poor (Table 3).

Table 2. Correlations (Pearson's *r*) between face-only and word-only versions of the Dynamic Visual Analogue Mood Scales, by scale and combined.

Run	Scale 1 *Miserable-Satisfied*	Scale 2 *Happy-Sad*	Scale 3 *Distressed-Peaceful*	Scale 4 *Bored-Excited*	Scale 5 *Afraid-Calm*	Scale 6 *Angry-Peaceful*	Scale 7 *Sleepy-Alert*	All
Run 1	.78	.67	.68	.62	.66	.40	.40	.59
Run 2	.76	.79	.76	.70	.79	.79	.73	.76

All correlations significant to 0.01 level (1-tailed)

Table 3. Intercorrelations (Pearson's *r*) between face-only versions of the Dynamic Visual Analogue Mood Scales, Run 2.

Scale	Scale 1 *Miserable-Satisfied*	Scale 2 *Happy-Sad*	Scale 3 *Distressed-Peaceful*	Scale 4 *Bored-Excited*	Scale 5 *Afraid-Calm*	Scale 6 *Angry-Peaceful*	Scale 7 *Sleepy-Alert*
Scale 1	1.00	.77	.73	.77	.76	.70	.73
Scale 2		1.00	.74	.88	.77	.73	.73
Scale 3			1.00	.71	.79	.81	.58
Scale 4				1.00	.70	.72	.77
Scale 5					1.00	.75	.62
Scale 6						1.00	.72
Scale 7							1.00

All correlations significant to 0.01 level (1-tailed)

Principal axis factoring of scores revealed one factor consistent with pleasantness of mood. The variance accounted for was similar for both face-only and word-only versions of the scale, but was consistently higher for the version including all seven scales. Variance accounted for was also consistently higher for the second run, with this factor accounted for 79.5% of the variance for the full version and 77.4% for the version without the 'Sleepy-Alert' scale (Table 4(a)). Internal consistency was very high with Cronbach's α values exceeding 0.9 for all versions of the scales (Table 4(b)); though these values were uniformly better for the second run, they were similar for both versions of the scales.

Table 4(a). Principal Axis Factoring analysis: Percentage of variance accounted for by factor 1

	Run 1		Run 2	
	D-VAMS[1]	D-VAMS[2]	D-VAMS[1]	D-VAMS[2]
Face-only	75.6	68.8	79.5	77.4
Word-only	71.2	68.2	81.6	78.0

D-VAMS[1]: Dynamic Visual Analogue Mood Scales scores; D-VAMS[2]: Dynamic Visual Analogue Mood Scales scores excluding 'Sleepy-Alert' scale.

Table 4(b). Cronbach's Alpha of D-VAMS scales

	Run 1		Run 2	
	D-VAMS[1]	D-VAMS[2]	D-VAMS[1]	D-VAMS[2]
Face-only	.933	.920	.948	.950
Word-only	.918	.920	.954	.951

D-VAMS[1]: Dynamic Visual Analogue Mood Scales scores; D-VAMS[2]: Dynamic Visual Analogue Mood Scales scores excluding 'Sleepy-Alert' scale.

Criterion validity

There were strong and highly significant correlations between D-VAMS mean scores and all HADS scores (Table 5). Correlations for the face-only scales were consistently strongest against HADS total scores, slightly weaker against the depression subscale scores, and weakest against the anxiety subscale scores.

Correlations for the face-only scale scores were uniformly higher than correlations for the means without the 'Sleepy-Alert' scale; correlations for the word-only scale scores followed a similar but less pronounced pattern. Correlations for the face-only – but not word-only – scales consistently improved from one run to the next. Word-only correlations were generally better than face-only correlations for the first run, but not for the second.

Table 5. Correlations (Pearson's *r*) between D-VAMS face-only and word-only score means ($\bar{x}$) and HADS total and subscale scores

Version	HADS score	Run 1		Run 2	
		D-VAMS^1 $\bar{x}$	D-VAMS^2 $\bar{x}$	D-VAMS^1 $\bar{x}$	D-VAMS^2 $\bar{x}$
Face-only	HADS Total	-.73	-.71	-.80	-.77
	HADS-D	-.71	-.69	-.73	-.70
	HADS-A	-.63	-.59	-.71	-.69
Word-only	HADS Total	-.82	-.80	-.80	-.77
	HADS-D	-.83	-.84	-.77	-.77
	HADS-A	-.67	-.62	-.69	-.64

D-VAMS^1: Dynamic Visual Analogue Mood Scales scores; D-VAMS^2: Dynamic Visual Analogue Mood Scales scores excluding 'Sleepy-Alert' scale; HADS-D: Hospital Anxiety and Depression Scale, depression subscale score; HADS-A: Hospital Anxiety and Depression Scale, anxiety subscale score.

All correlations significant to 0.01 level (1-tailed)

Test-retest reliability

Table 6 shows intraclass correlations of D-VAMS scores and means between Run 1 and Run 2. Correlations for face-only and word-only versions of the scales were similarly high (.62 to .89), with correlations of score means exceeding 0.9 for both the seven-item and six-item versions. Correlations for the seven-item version were slightly higher than those for the six-item version. Figure 5 shows Bland-Altman diagrams for scale scores and score means for the face-only version of the scales. Comparison of means (paired values *t*-test, two-tailed) between Run 1 and Run 2 scores for face-only scales revealed proportional bias for six-item scale means (+3.4; *P*=.024), Scale 1 scores (+5.1; *P*=.048), and Scale 4 scores (+7.5; *P*=.002) but not for any of the others. For word-only scales, proportional bias was noted only for Scale 7 (+5.6; *P*=0.018).

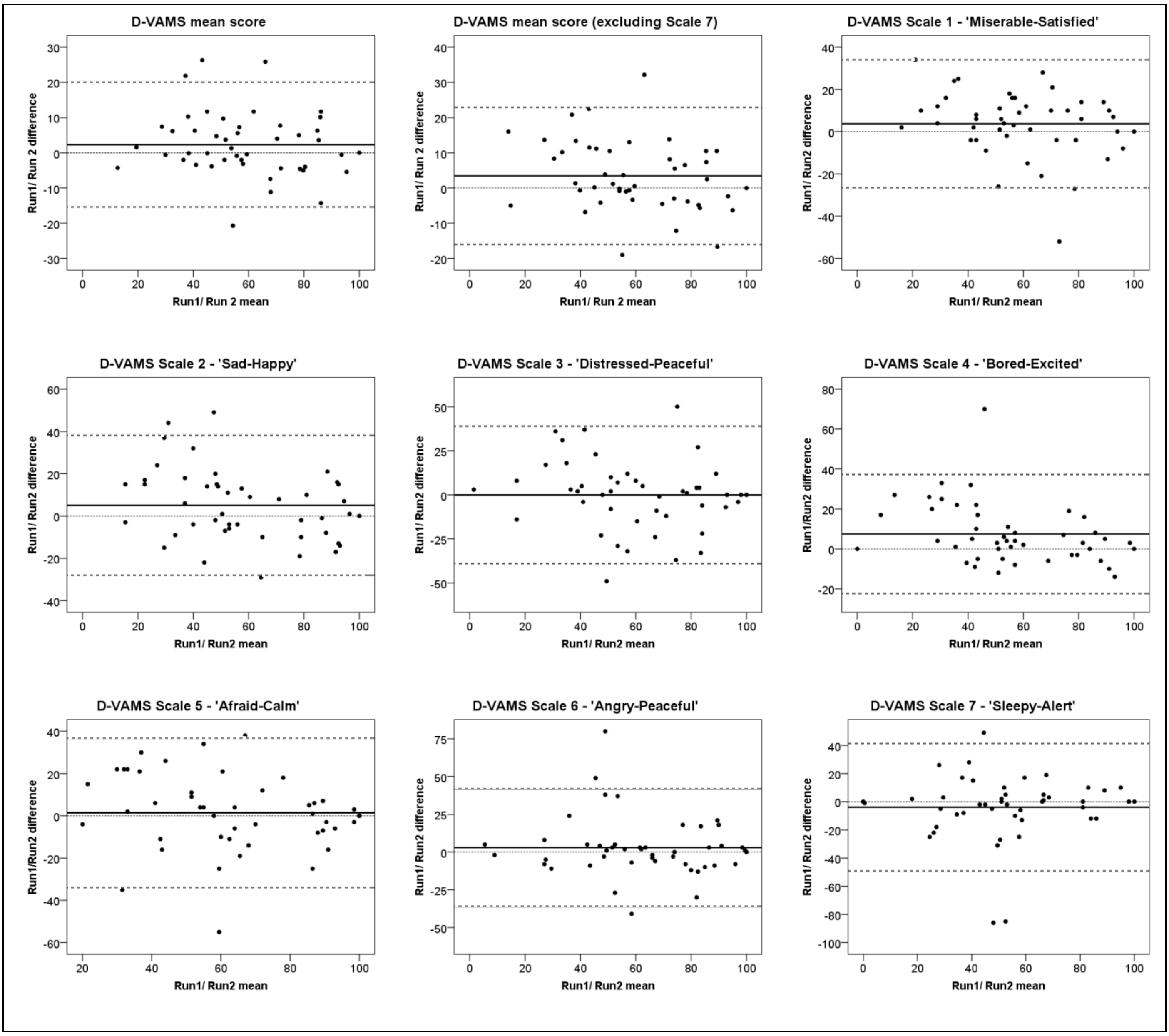

Figure 5. Bland-Altman plots for D-VAMS mean and scale scores. Score differences between runs (black line) with 95% confidence interval (dashed line). The dotted line at Y=0 indicates perfect agreement between scores from one run to the next.

Table 6. Intraclass correlations (2-way mixed) between Run 1 and Run 2 scores for face-only and word-only versions of D-VAMS

	Scale 1 *Miserable-Satisfied*	Scale 2 *Happy-Sad*	Scale 3 *Distressed-Peaceful*	Scale 4 *Bored-Excited*	Scale 5 *Afraid-Calm*	Scale 6 *Angry-Peaceful*	Scale 7 *Sleepy-Alert*	D-VAMS[1] $\bar{x}$	D-VAMS[2] $\bar{x}$
Face-only	0.78	0.79	0.70	0.82	0.74	0.71	0.62	0.91	0.90
Word-only	0.76	0.69	0.89	0.66	0.86	0.71	0.88	0.93	0.92

D-VAMS[1]: Dynamic Visual Analogue Mood Scales scores; D-VAMS[2]: Dynamic Visual Analogue Mood Scales scores excluding 'Sleepy-Alert' scale.

All correlations significant to 0.01 level (1-tailed).

Sensitivity and Specificity

ROC analyses were performed for the face-only score means using data from the retest run. Optimal cut-offs for D-VAMS scores against the HADS total score, and depression and anxiety subscale scores are shown in Table 7, along with sensitivity, specificity, and percentage of area under the curve.

Table 7(a). D-VAMS[1] mean cut-offs for optimal sensitivity/specificity against HADS total score (/42)

HADS cut-off	ROCS AUC %	D-VAMS cut-off	Sens %	Spec %
≥ 8	91.9	≤ 77	87	88
≥ 10	94.3	≤ 77	94	91
≥ 12	95.3	≤ 71	97	87
≥ 14	92.6	≤ 69	93	74
≥ 16	90.9	≤ 66	91	63
≥ 18	83.9	≤ 59	83	64
≥ 20	82.7	≤ 59	81	60
≥ 22	80.6	≤ 47	67	85
≥ 24	86.7	≤ 41	83	93

HADS: Hospital Anxiety and Depression Scale total score; D-VAMS[1]: Dynamic Visual Analogue Mood Scales scores; ROC, AUC: Receiver Operating Characteristics, Area Under Curve; Sens: Sensitivity; Spec: Specificity.

Table 7(b). D-VAMS[1] mean cut-offs for optimal sensitivity/specificity against HADS-D score (/21)

HADS-D cut-off	ROCS AUC %	D-VAMS cut-off	Sens %	Spec %
≥ 4	87.8	≤ 74	89	91
≥ 5	89.5	≤ 74	97	77
≥ 6	91.0	≤ 69	96	79
≥ 7	83.6	≤ 63	83	65
≥ 8	82.1	≤ 59	81	68
≥ 9	86.0	≤ 59	85	69
≥ 10	83.2	≤ 57	73	87
≥ 11	88.2	≤ 52	81	68
≥ 12	82.4	≤ 44	80	88

HADS-D: Hospital Anxiety and Depression Scale depression subscale score; D-VAMS[1]: Dynamic Visual Analogue Mood Scales scores; ROC, AUC: Receiver Operating Characteristics, Area Under Curve; Sens: Sensitivity; Spec: Specificity.

Table 7(c). D-VAMS[1] mean cut-offs for optimal sensitivity/specificity against HADS-A score (/21)

HADS-A cut-off	ROCS AUC %	D-VAMS cut-off	Sens %	Spec %
≥ 4	90.3	≤ 78	87	86
≥ 5	86.8	≤ 71	81	80
≥ 6	90.6	≤ 74	83	80
≥ 7	85.4	≤ 69	89	78
≥ 8	84.7	≤ 60	81	75
≥ 9	83.5	≤ 59	81	68
≥ 10	83.5	≤ 59	80	65
≥ 11	82.3	≤ 51	69	83
≥ 12	79.1	≤ 51	67	81

HADS-A: Hospital Anxiety and Depression Scale anxiety subscale score; D-VAMS[1]: Dynamic Visual Analogue Mood Scales scores; ROC, AUC: Receiver Operating Characteristics, Area Under Curve; Sens: Sensitivity; Spec: Specificity.

Linear regression using HADS total scores and D-VAMS (seven-item) mean cut-off scores yields a β of 2.23 and an intercept of 98.6, very close to the theoretical β of 2.38 and intercept of 100 predicted by an inverse, linear relationship between two scales in the range of 0–100 (D-VAMS) and 0–42 (HADS) respectively. The relationship between the two scales can therefore be approximated as

$$DVAMS = 100 - (HADS \times 2.4)$$

or

$$HADS = \left(\frac{100 - DVAMS}{2.4}\right)$$

where *DVAMS* is the Dynamic Visual Analogue Mood Scale score mean and *HADS* is the Hospital Anxiety and Depression Scale total score.

Discussion

The high cross-correlations between the scales, poor discriminant validity, single-factor structure, and high Cronbach's α suggest that the D-VAMS serve as a valid measure of pleasantness of mood on a scale of 0–100. Construct validity, criterion validity against the HADS, and test–retest reliability all appear good to excellent, and sensitivity and specificity across a range of HADS cut-offs was generally very good, though better against the HADS total score than the depression and anxiety subscale scores.

The improved construct validity from test to retest runs, the uniformly higher percentage of variance accounted for, the improved criterion validity against the HADS and the higher Cronbach's α all suggest a practise effect, so a test run is recommended for patients using the scales for the first time. This also suggests that the results from Run 2 better reflect the scales' performance under optimal conditions, where a respondent is already familiar with their use.

The comparison of two D-VAMS score means relates to questions raised by the anomalous nature of the Scale 7 (*Sleepy-Alert*), which was charted as close to valence-neutral in PCA plots from preliminary studies[21] (Figure 1). Since only a single valence factor was observed in scores for this study, it was reasoned that the mean of the first six scales might provide a more valid total score for pleasantness of mood. The pattern of data, however, suggests that omitting Scale 7 does not improve its psychometric qualities. Scale 7 scores were highly correlated with those of the other scales (Table 3), and the variance accounted for by the seven-item scale scores was consistently higher than those of the six-item version (Table 4(a)). Furthermore, the HADS correlations for the former were generally better than those of the latter (Table 5). Scale 7 should therefore be retained.

Employing versions of the faces scales in which verbal labels were completely absent offered a strong test of the utility of the scales, as respondents were reliant on images of facial expressions alone, thereby simulating the conditions for a profoundly aphasic respondent. It is reasonable to expect that the live version of the D-VAMS – in which the scale endpoints are also accompanied by corresponding mood words – would perform better, as it is rare for people with aphasia to have no language comprehension at all.

Though adapted mood measures are frequently used as an indirect measure of depression following stroke, it is important to bear in mind the distinction between depressed mood and depression as a clinical entity. Though depressed mood and lack of pleasure are central features of the phenomenology of depression, there are other features of depression, such as the complex cognitions and physical symptoms, that are neglected by this simplified definition. The D-VAMS can therefore only be considered an outcome measure of depressed mood rather than depression *per se*.

Similarly, the HADS, though widely used as a measure of depression and anxiety, also relies on a somewhat narrower definition of these constructs. Since the HADS was created to assess these symptoms in the context of physical illness, somatic symptoms are avoided because they are commonly confounded by the effects of physical illness. This necessarily confines its scope to non-somatic symptoms, of which mood is a central feature, and this narrower focus may partly explain the high correlations between scores from these instruments. However, the HADS does cover cognitive as well as affective features of depression, and the two self-report methods – one using faces and sliders, and the other using a brief set of questions – are very different in the way that information of their respective domains is communicated; such high correlations underline the importance of mood as a predominating feature of depression.

Turning to criterion validity, it seems clear that the D-VAMS are a better correlate of the HADS total scores than its depression or anxiety subscale scores individually. The higher correlations against depression subscale scores compared to anxiety subscale scores are understandable, given the placement of descriptors for depressed and anxious mood within the two-factor arrangement of the circumplex model of affect. Mood words like ‘depressed’ and ‘sad’ – characterised predominately as negative or unpleasant mood – typically load heavily on the valence dimension. ‘Anxious’ mood, however, along with related mood descriptors such as ‘tense’, ‘afraid’, and ‘nervous’ is generally found to load less on the valence dimension and much more on the activation dimension.[21,24] The single factor structure of the D-VAMS would be expected to favour correlations with constructs loading heavily on the valence dimension.

This strong criterion validity suggests that the D-VAMS may prove useful as an indirect measure of depression for those whose communication is too poor to allow language-based measures or clinical interview. This and the excellent test–retest reliability suggest that this instrument is a valid and reliable outcome measure suitable for clinical assessment and research purposes.

Bennett et al.[9] recommend a sensitivity of at least 0.8 and a specificity of at least 0.6 as an acceptable cut-off for a screening measure for depression following stroke. A number of cut-offs against the HADS were found that consistently exceeded these criteria, suggesting that the D-VAMS may also be suitable for this purpose.

Though these findings are encouraging, it is important to acknowledge the limitations of the present study. First, this validation study mostly comprised stroke patients without significant aphasia, which was necessary because participants had to be able to respond to questions on the HADS. Though the inclusion of more people with aphasia would be desirable, it is impossible to escape the problem of using a language-based measure with people who are severely language impaired. It is also difficult to know what conclusion to draw from the results were more people with aphasia to have been included as a comparison group. We would expect language impairment to impact upon their ability to complete the HADS, and therefore a decreased correlation between this and the D-VAMS would seem likely. But it would be difficult to establish how much how much of this reduced correlation was due to an impaired ability to used the criterion measure and how much was due to any impaired ability to use the face scales in this group. However, there is a body of neuropsychological evidence suggesting that the recognition of emotion in facial expressions is mediated primarily by right hemisphere processes, with only right-brain injuries being associated with corresponding impairment.[25-30] Since most aphasic patients have left hemisphere lesions, recognition of facial expression is unlikely to be impaired in this group, so the use of faces may be a particularly suitable means of enabling mood to be reported.

Second, this study sample comprised participants for whom a generally long period of time had elapsed since their stroke (mean=3.4 years). The lack of participants in the period immediately following stroke means that these findings should be treated with caution. Further study is needed in patients in the acute stage following stroke.

Third, the relatively small sample size warrants caution. Though adequate for the purposes of the correlational analyses performed, results of factor analysis can be less reliable on data from sample sizes below one hundred.

Finally, responsiveness to change could not be evaluated by this study. These scales need to be assessed in the context of an intervention in order for this property to be evaluated.

Granting these limitations, however, the D-VAMS show promise and should prove to be an improvement upon the limited instrumentation currently available to assess mood in this population. This instrument is free to use and available on the internet at DVAMS.ORG. A downloadable version that can be run offline is also available.

Clinical Messages

- Dynamic Visual Analogue Mood Scales used in people with stroke showed good validity, internal consistency, and reliability, and show promise as a measure for depression following stroke.
- They do not require the use of language and early data suggest they may be used in people with aphasia.

References

1. Kutlubaev M, Hackett ML. Part II: predictors of depression after stroke and impact of depression on stroke outcome: an updated systematic review of observational studies. *International Journal of Stroke*. 2014;9(8):1026-36.
2. Thomas SA, Lincoln NB. Factors relating to depression after stroke. *British Journal of Clinical Psychology*. 2006;45(Pt 1):49-61.
3. Stern RA, Arruda JE, Hooper CR, Wolfner GD, Morey CE. Visual analogue mood scales to measure internal mood state in neurologically impaired patients : description and initial validity evidence. *Aphasiology*. 1997;11(1):59-71.
4. Arruda JE, Stern RA, Somerville JA. Measurement of mood states in stroke patients: validation of the visual analog mood scales. *Archives of Physical Medicine and Rehabilitation*. 1999;80(6):676-80.
5. Brumfitt SM, Sheeran P. The development and validation of the Visual Analogue Self-Esteem Scale (VASES). *British Journal of Clinical Psychology*. 1999;38:387-400.
6. Turner-Stokes L, Kalmus M, Hirani D, Clegg F. The Depression Intensity Scale Circles (DISCs): a first evaluation of a simple assessment tool for depression in the context of brain injury. *Journal of Neurology, Neurosurgery and Psychiatry*. 2005;76(9):1273-8.
7. Townend E, Brady M, McLaughlan K. A systematic evaluation of the adaptation of depression diagnostic methods for stroke survivors who have aphasia. *Stroke*. 2007;38(11):3076-83.

8. Berg A, Lonnqvist J, Palomaki H, Kaste M. Assessment of depression after stroke: a comparison of different screening instruments. *Stroke*. 2009;40(2):523-9.

9. Bennett HE, Thomas SA, Austen R, Morris AM, Lincoln NB. Validation of screening measures for assessing mood in stroke patients. *British Journal of Clinical Psychology*. 2006;45(Pt 3):367-76.

10. Lincoln NB, Kneebone II, Macniven AB, Morris RC. Psychological Management of Stroke. Chichester: John Wiley & Sons, Ltd. ; 2012.

11. Van Dijk MJ, de Man-van Ginkel JM, Hafsteinsdóttir TB, Schuurmans MJ. Identifying depression post-stroke in patients with aphasia: A systematic review of the reliability, validity and feasibility of available instruments. *Clinical Rehabilitation*. 2016;30(8):795-810.

12. Benaim C, Cailly B, Perennou D, Pelissier J. Validation of the aphasic depression rating scale. *Stroke*. 2004;35(7):1692-6.

13. Benaim C, Decavel P, Bentabet M, Froger J, Pelissier J, Perennou D. Sensitivity to change of two depression rating scales for stroke patients. *Clinical Rehabilitation*. 2010;24(3):251-7.

14. Watkins C, Leathley M, Daniels L, Dickinson H, Lightbody CE, van den Broek M, et al. The signs of depression scale in stroke: How useful are nurses' observations? *Clinical Rehabilitation*. 2001;15:447.

15. Sutcliffe LM, Lincoln NB. The assessment of depression in aphasic stroke patients: the development of the Stroke Aphasic Depression Questionnaire. *Clinical Rehabilitation*. 1998;12(6):506-13.

16. Hacker VL, Stark D, Thomas S. Validation of the stroke aphasic depression questionnaire using the brief assessment schedule depression cards in an acute stroke sample. *British Journal of Clinical Psychology*. 2010;49(Pt 1):123-7.

17. Williams L, Bakas T, Brizendine E, Plue L, Wanzhu T, Hendrie H, et al. How valid are family proxy assessments of stroke patients' health-related quality of life? *Stroke*. 2006;37(8):2081-5.

18. Cobley CS, Thomas SA, Lincoln NB, Walker MF. The assessment of low mood in stroke patients with aphasia: reliability and validity of the 10-item Hospital version of the Stroke Aphasic Depression Questionnaire (SADQH-10). *Clinical Rehabilitation*. 2012;26(4):372-81.

19. Russell JA. A circumplex model of affect. *Journal of Personality and Social Psychology*. 1980;39(6):1161-78.

20. Yik MSM, Russell JA, Steiger JH. A 12-point circumplex structure of Core Affect. *Emotion*. 2011;11(4):705-31.

21. Barrows PD. *The development and validation of mood scales suitable for use with stroke patients with aphasia* (available at University of Nottingham ePrints). PhD Dissertation, University of Nottingham, Nottingham. 2016.

22. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*. 1983;67(6):361-70.

23. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*. 2007;39(2):175-91.

24. Russell JA and Pratt G. A description of affective quality attributed to environments. *Journal of Personality and Social Psychology*. 1980;38(2):311-22.

25. Borod JC, Koff E, Perlman Lorch M, Nicholas M. The expression and perception of facial emotion in brain-damaged patients. *Neuropsychologia*. 1986;24(2):169-80.

26. Moreno CR, Borod JC, Welkowitz J, Alpert M. Lateralization for the expression and perception of facial emotion as a function of age. *Neuropsychologia*. 1990;28(2):199-209.

27. Adolphs R, Damasio H, Tranel D, Damasio R. Cortical systems for the recognition of emotion in facial expressions. *The Journal of Neuroscience*. 1996;16(23):7678-87.

28. Borod JC, Cicero BA, Obler LK, Whalen JR. Right hemisphere emotional perception: evidence across multiple channels. *Neuropsychology*. 1998;12(3):446-58.

29. Kucharska-Pietura K, Phillips ML, Gernand W, David AS. Perception of emotions from faces and voices following unilateral brain damage. *Neuropsychologia*. 2003;41(8):1082-90.

30. Philippi CL, Mehta S, Grabowski T, Adolphs R, Rudrauf D. Damage to association fiber tracts impairs recognition of the facial expression of emotion. *The Journal of Neuroscience*. 2009;29(48):15089-99.

Acknowledgements

We thank Nottingham CityCare Partnership Community Stroke Discharge and Rehabilitation Service, Nottingham Stroke Club, Sherwood Stroke Club and the many service users and members who agreed to take part in this study. We also thank Professor

Nadina Lincoln for her supervisory help and the Nottingham Stroke Research Partnership for their support of this project.

Conflict of interest

None declared.

Funding

The author(s) received no financial support for this research.

Supplementary Materials

D-VAMS can be accessed from the internet via DVAMS.ORG. A downloadable version and an Android app are also available. Study data is available upon request from Paul Barrows at paul.barrows@dvams.org.