Psychopathology 1989;22:42–49


0254–4962/89/0221–0042$2.75/0

# Measuring Subacute Mood Changes Using the Profile of Mood States and Visual Analogue Scales

*Karley Little, Elizabeth Penman*

Department of Psychiatry, University of Kentucky College of Medicine, Lexington, Ky., USA

**Abstract.** Mood lability and mixtures of moods are important clinical features that may contribute to diagnostic formulation and treatment response. Attempts at quantifying such lability and mixtures have been limited. In this preliminary study both the Profile of Mood States (POMS) and comparable Visual Analogue Scales were assessed for their feasibility and relative performance over a 3- to 5-day period in two groups of inpatients. Both demonstrated concurrent validity and feasibility. The performance of the POMS appeared somewhat superior. A major concern remains regression to the mean after multiple sampling. The authors suggest continued exploration of the relationships between POMS change scores and diagnosis, biological markers, and treatment outcome.

## Introduction

Stability or pervasiveness of depressed mood is a requisite feature for diagnosing the major mood disorders [American Psychiatric Association, 1987]. Conversely, atypical depression (or 'hysteroid dysphoria'), as well as more characterologically related mood syndromes, have been linked to unstable and mixed mood patterns. Regular diurnal mood variations have been seen as diagnostically significant, although their specificity has been questioned [Fahndrich, 1987]. Evidence suggests that stability of depressed mood is predictive of eventual response to an antidepressant [Nelson et al., 1984]. Recently, Lemus et al. [1987] have directly tied a physiological derangement, cortisol nonsuppression after dexamethasone to mood nonreactivity as quantitated on the SADS, an interview-based review of symptoms.

However, measuring mood mixtures and patterns of lability, as well as the effects of treatment on mood states, are challenging tasks. An instrument needs to be nonintrusive and sensitive, while providing a valid measure of an essentially subjective phenomenon. Multiple retesting is necessary, so the test must be reliable, simple, and practical enough to readminister, particularly at times when trained raters are not available. The most prominent tools currently used to quantitate moods in the clinical setting are Visual Analogue Scales (VAS; 100-mm line tests) and the Profile of Mood States (POMS; a 65-item adjective checklist). Other mood-adjective checklists exist, but the POMS is the

most often used. No standardized versions of 100-mm line tests exist. Whether or not these instruments might be useful in studying existing diagnostic categories, for identifying new clinical syndromes, and for clarifying the relationship between mood states, diagnosis, and treatment response, is unclear.

In this preliminary feasibility/validity study, we assessed these two instruments' sensitivity to diagnosis, clinical setting, and clinical change, as well as performance relative to each other over one subacute period (3–5 days). Two groups of psychiatric inpatients twice completed both the standard POMS and a 6-item VAS designed to match the factors on the POMS. If this study was encouraging it was felt future studies sampling several intervals at varying frequencies might produce even more meaningful lability and mixture measures.

## Methods

*Subjects.* Thirty-one adult inpatients were studied on a general admissions ward of a VA hospital. Another 31 adult subjects were patients on a university hospital ward. All patients were first given both the POMS and VAS on either hospital day 2, 3, or 4. Three to five days later these same test were given in reverse order. During this period treatment was often begun, but few patients had achieved their final therapeutic endpoint. Change scores probably reflected underlying psychopathology and the response to nonspecific treatments, rather than response to definitive psychopharmacological treatment. All new admissions were considered for the study. Inclusion criterion was a diagnosis of either mood disorder, personality disorder, or alcohol/drug abuse without signs of acute withdrawal. Subjects were excluded when judged on clinical grounds to have signs of organicity or psychosis that would clearly interfere with cooperation or understanding.

*Instruments.* The POMS were empirically formulated in the sixties by McNair and Lorr using results obtained in psychiatric outpatients and college students [McNair et al., 1981]. It has consistently been found sensitive to psychotherapy and drug effects. The POMS's 6 mood factors – anger, depression, tension, confusion, fatigue, and vigor – were identified by factor-analytic techniques. It is relatively quick and easy to understand, has normative data available, and evidence for its concurrent and predictive validity, and is regularly used experimentally. Standard POMS instructions were given to subjects including directions to mark how they felt 'right now'. Responses were scored using the overlay sheets.

Line tests have been popular since the 1920s [Aitken, 1969]. They have obvious appeal because of simplicity and directness. There has been little effort at standardizing the scales. Each user typically labels the scale extremes with his own construct. Reliability has been claimed because of stability of group means after multiple retesting [Folstein and Luria, 1973]. Evidence for more than face validity includes correlative relationships with the Zung Depression Scale, Clyde Mood Scale, and greater group extremes among affective disorder patients [Luria, 1975]. Our version had six 100-mm lines labeled respectively 'irritable', 'clear-headed', 'energetic', 'nervous', 'tired', and 'depressed'. These labels were felt to be vivid, understandable, and nonpejorative. One end was labeled 'least ever', the other 'most ever'. Subjects were told to recall actual times in their lives when they experienced these extremes and compare their present state to those times. They were directly reminded of the significance of an extreme or central response. Responses were measured to the nearest millimeter.

Each subject was objectively rated on a global clinical improvement/worsening scale (from +3 to −3) at the second test date. For each individual, a total distress score was calculated for the POMS and VAS by summing all 5 negative mood scales and subtracting the vigor score.

*Statistical Procedures.* Differences in change scores between each diagnostic group were tested for significance by one-way ANOVA. Change scores from day 1 to day 2 on each subscale were analyzed for each hospital using the t test. Correlations between POMS versus VAS subscales on the same day, between day 1 versus day 2 for each scale, and between POMS change scores versus VAS change scores were calculated using the Pearson product-moment. Changes in individual total distress scores were also correlated with the objective clinical change score using the Pearson product-moment.

**Table I.** Change scores in different diagnosis

| Diagnosis | Day | Tension | | Anger | | Vigor | | Depression | |
|---|---|---|---|---|---|---|---|---|---|
| | | POMS | VAS | POMS | VAS | POMS | VAS | POMS | VAS |
| Major depression (n = 27) | 1 | 18.5 | 53.5 | 12.1 | 40.6 | 10.0 | 33.1 | 27.5 | 63.3 |
| | 2 | 13.7* | 43.8 | 8.6* | 39.3 | 13.0 | 47.1* | 20.6* | 45.2 |
| Bipolar disorder (n = 10) | 1 | 17.8 | 38.2 | 18.0 | 38.3 | 12.4 | 27.2 | 20.9 | 36.8 |
| | 2 | 16.3 | 41.9 | 11.6* | 30.8 | 12.3 | 35.7 | 17.7 | 31.7 |
| Minor depression (n = 8) | 1 | 18.1 | 62.2 | 16.1 | 63.5 | 10.1 | 32.4 | 24.2 | 69.1 |
| | 2 | 12.6* | 46.1* | 9.4* | 33.5* | 15.4* | 51.6* | 12.5 | 35.0 |
| Personality disorder (n = 5) | 1 | 15.6 | 65.4 | 6.8 | 38.4 | 7.4 | 33.8 | 20.4 | 56.6 |
| | 2 | 11.0* | 37.6* | 3.2* | 20.4* | 15.6* | 61.0* | 13.8* | 40.6* |
| Alcohol abuse, chronic (n = 7) | 1 | 22.7 | 56.7 | 15.9 | 36.9 | 14.4 | 48.0 | 34.7 | 66.1 |
| | 2 | 21.2 | 61.1 | 14.3 | 44.0 | 13.2 | 49.8 | 31.7 | 61.7 |
| Schizoaffective disorder (n = 5) | 1 | 20.0 | 50.6 | 10.2 | 50.4 | 11.8 | 28.2 | 26.4 | 42.4 |
| | 2 | 17.4 | 41.8 | 13.0 | 49.8 | 14.6 | 39.0 | 16.0 | 22.0 |

* $p < 0.05$.

**Table II.** Change scores at different hospitals

| Hospital | Tension | | Anger | | Vigor | | Depression | |
|---|---|---|---|---|---|---|---|---|
| | POMS | VAS | POMS | VAS | POMS | VAS | POMS | VAS |
| VA (n = 31) | 22.3 | 60.0 | 15.4 | 44.4 | 10.4 | 30.1 | 31.6 | 60.4 |
| | 20.9 | 55.7 | 14.6 | 42.6 | 11.3 | 36.0 | 28.6 (p = 0.08) | 53.1 |
| UN (n = 31) | 15.1 | 46.5 | 11.5 | 42.3 | 11.3 | 36.8 | 20.8 | 55.3 |
| | 9.04**** | 34.51* | 4.84**** | 31.4 | 15.84**** | 57.23*** | 10.24**** | 29.64**** |

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

## Results

Table I lists by diagnostic category initial and repeat mood scores. Patients diagnosed minor depression and primary personality disorder improved on most measures. Unipolar major depressives showed some improvement, while bipolars, alcoholics and schizoaffective patients demonstrated little change. Personality disorders rated themselves low on anger initially, and even significantly lower on repeat. Minor depressives initially scored themselves at high levels comparable to major depressives (somewhat

| Fatigue | | Confusion | |
|---|---|---|---|
| POMS | VAS | POMS | VAS |
| 14.7 | 64.8 | 14.3 | 52.2 |
| 9.4* | 44.4* | 10.9* | 41.8* |
| 11.4 | 37.4 | 13.3 | 42.7 |
| 11.7 | 46.0 | 12.2 | 44.7 |
| 11.6 | 51.0 | 13.9 | 59.6 |
| 4.5* | 33.4* | 7.4* | 36.6* |
| 13.2 | 66.0 | 11.8 | 63.8 |
| 8.0* | 33.2* | 8.6 | 40.2 |
| 15.4 | 61.3 | 9.2 | 20.3 |
| 12.8 | 34.6 | 9.7 | 41.6 |
| 19.4 | 55.4 | 14.4 | 62.6 |
| 16.8 | 62.6 | 9.4 | 52.6 |

| Fatigue | | Confusion | |
|---|---|---|---|
| POMS | VAS | POMS | VAS |
| 15.5 | 54.7 | 14.2 | 51.4 |
| 14.8 | 51.5 | 12.9 | 48.0 |
| 12.4 | 60.1 | 12.4 | 48.2 |
| 5.34**** | 34.44**** | 7.64**** | 37.6* |

more angry), but improved much more. Alcoholics rated themselves higher on depression and tension than major depressives. Twelve significant changes were detected by both instruments. The POMS reached statistical significance in detecting decreasing anger in bipolars and decreasing depression in major depressives, confirming nonsignificant trends on the VAS. The VAS detected increasing vigor in major depressives ($p < 0.05$) confirming the nonsignificant POMS trend. The POMS found significant decreasing anger in major depressives, which the VAS did not confirm.

As table II indicates, both instruments provided substantial evidence that VA patients were a different population from the University patients. This confirmed a clinical impression based on assessment of premorbid adjustment, work history, length of stay, and treatment response. More VA patients were alcoholics (all 7) or schizoaffective (4 of 5) and fewer minor depressives (1 of 8). In VA patients, both POMS and VAS found higher mean scores on every negative scale (except fatigue on VAS day 1: VA 54.7; UN 60.1) and lower mean scores on the vigor score, on both day 1 and day 2. Both scales showed little improvement in group means among VA patients from day 1 to day 2. In contrast, both scales showed significant improvement in group means among UN patients. Overall the POMS was superior to the VAS in providing statistically significant results. POMS change scores among UN patients were all significant at the $p < 0.0001$ level. The VAS confirmed all 6 of these changes but because of greater variance only reached $p < 0.0001$ twice, $p < 0.001$ once, $p < 0.05$ twice, while failing to reach significance on the anger scale. However, among VA patients the VAS did demonstrate a significant decrease in depression ($p < 0.05$) which the POMS confirmed with only $p = 0.08$.

Table III compares correlations between each day and between each scale. There were significant relationships between subjects' score on day 1 and day 2 with both scales (except fatigue on VAS). However, POMS

**Table III.** Correlations between days, tests, change scores

| | | Tension | Anger | Vigor | Depression | Fatigue | Confusion |
|---|---|---|---|---|---|---|---|
| Day 1 vs. day 2 | VAS | 0.54 | 0.48 | 0.47 | 0.54 | 0.10 (0.46) | 0.36 (0.004) |
| | POMS | 0.78 | 0.77 | 0.61 | 0.78 | 0.59 | 0.72 |
| VAS vs. POMS | Day 1 | 0.73 | 0.51 | 0.45 (0.0002) | 0.61 | 0.60 | 0.47 |
| | Day 2 | 0.69 | 0.63 | 0.67 | 0.71 | 0.64 | 0.54 |
| ΔVAS vs. ΔPOMS | | 0.53 | 0.26 (0.04) | 0.37 (0.003) | 0.56 | 0.60 | 0.26 (0.04) |

All significant at p = 0.0001, exceptions with p values in parentheses.

scores were considerably less variable with its 6 r values averaging 0.71 versus 0.42 for the VAS. POMS scores and VAS scores were significantly correlated (r = 0.47 on Confusion to r = 0.73 on Tension). Consistency between the two increased on the 2nd day (average r value 0.65 vs. 0.56 on day 1).

Overall, based on correlation calculations, patients were more consistent on tension, depression, and fatigue subscales. Confusion and vigor subscales showed less consistency on both instruments. The POMS anger scale was more helpful than the VAS in detecting changes and correlations.

The total distress scores on both POMS and VAS correlated moderately well with the objective rating. However, the POMS showed a greater correlation (r = 0.63, p = 0.0001) than the VAS (r = 0.52, p = 0.0001).

## Discussion

Both the POMS and VAS subscales demonstrated concurrent validity in appropriately detecting diagnostic differences, clinical setting differences, and clinical change. The degree of agreement between the two also suggests construct validity. Comparatively, the POMS was more reliable, achieving lower variability and greater degrees of statistical significance in the change scores analysis (table II) and objective rating correlation. Despite the opportunity the VAS provides subjects for making graduated distinctions, psychiatric inpatients sometimes provide impulsive or extreme responses, even though instructed of the significance of such markings. Test-retest agreement on the VAS, as measured by r value, was lower and appeared less optimal than the POMS, generating variance rather than sensitivity. This study provides new evidence that the POMS is valid in inpatients and is appropriately sensitive to intermediate and clinically significant mood change (measuring neither transient states nor long-term traits, as questioned by previous reviewers [Peterson and Headon, 1984]. Considering the brevity and ad hoc nature of our VAS instrument, its performance might be viewed favorably. Other construct labels might correlate better with POMS results, though not decreasing variability. It is not clear that either of these instruments would be sensitive enough to use in making individual diagnostic formulations or treatment decisions.

We found evidence suggesting possible limitations with either instrument. Subjects with certain diagnoses may have distorted their reports. Personality-disordered patients rated themselves very low on anger, and alcoholics were highest on depression. However, 3 of our PD patients were diagnosed as primarily passive-aggressive and none were felt to be borderline, antisocial or histrionic. Our small sample size and lack of detail limit generalizations. We also believe the high depression scores of alcoholics and minor depressives may be valid. Single cross-sectional examinations often disclose convincingly intense verbal, facial and bodily expressions of sadness that are found not pervasive when longer sampling periods in other circumstances are undertaken.

It is unclear if POMS stability would remain optimal using shorter intervals (within a single day). Our evidence does show, however, in fairly ill inpatients, that subjects can comply with the demands of the POMS at a 3- to 5-day interval. More frequent testing might be less reliable due to decreased subject effort and POMS performance might deteriorate below that of the VAS. However, the POMS has been shown to be sensitive to changes after drugs at intervals of 2–4 h [Chait et al., 1985], while there is evidence that VAS performance shows marked regression to the mean after multiple retests [greater than 10 trials in normals, McClellan, 1987]. Our data shows greater agreement between the POMS and VAS on day 2 (table III), perhaps demonstrating improved reliability because of greater subject concentration and practice, but also potentially because of regression to the mean on both scales. Single tests have difficulty in distinguishing between chronic traits and temporary states, which practical and valid multiple tests may alleviate if these centripetal tendencies can be minimized [Boyle, 1985].

More generally, these results suggest that one cannot assume a simple relationship between a certain mood and a certain syndrome. Moods and emotions are rarely 'pure'. Polivy's [1981] work suggests that moods involve unique mixtures of pure anxiety, depression, anger, etc. Distinct psychiatric syndromes with specific disordered emotional states seen as the core symptoms often clinically involve complex mixtures of 'pure sadness, anger, anxiety, etc'. A considerable number of patients exhibit the coexistence of anxiety, depression, and anger in varying degrees. This is consistent with evidence for familial and neurobiological overlap between anxiety and depressive disorders [Mountjoy and Roth, 1982].

Our evidence showing differences among subscales suggests subjects may have a clearer and more consistent idea of how to assess and quantitate some moods better than others. However, selection factors and social desirability influences may account for the differences.

Despite the fact that the VAS has been gaining greater favor among psychiatrists [Kendall et al., 1984; Eastwood et al., 1984; Davis et al., 1985], our results may suggest some caution. More extensive sampling is achieved by the POMS, which poses similar questions a number of times from slightly different semantic perspectives (Are you miserable?... dependent?... discouraged? etc.). In clinical (and practical) usage, moods refer to complex families of related sensations, feelings, and thoughts, Individuals and cultures vary as to which sensations, feelings and thoughts they include within a mood boundary, how central or peripheral they conceptualize each of these phenomena, and

how much each actually affects behavior [Lutz, 1985; Shields, 1984]. No single label will likely sharply delimit moods. The present subacute study does not directly address the utility of the VAS in response to acute, drug-induced perturbations. Unless the VAS label can crystallize the varied sensations, feelings and thoughts after a drug, the VAS as compared to multi-item mood scales like the POMS or drug-sensitive scales like the ARCI [Haertzen, 1984] would seem mildly inferior. This study suggests that individual VAS results will not be equivalent to those after the POMS, and that group scores on the VAS are less reliable indicators of change. Despite these reservations, continued efforts at standardizing VAS and understanding the appropriate contexts for their use should be undertaken.

Based on these findings, we feel encouraged to pursue the use of the POMS at multiple intervals. It may usefully serve as a supplement to standard Depression Self-Inventories such as the Carroll or Zung Scales for documenting stability and interactions of mood states. It might also help validate diagnostic subclasses such as hysteroid dysphoria, subaffective dysthymia, demoralization syndrome, and quantitate differences between personality disorder types, or endogenous versus nonendogenous major depressives. Additionally, the relationships between mood states, biological markers and treatment outcome might be profitable addressed with this instrument.

**Acknowledgement**

The authors would like to acknowledge the assistance of Virginia Lynn Gift and Debbie Howard for typographical services.

**References**

Aitken, R.C.B.: Measurement of feelings using visual analogue scales. Proc. R. Soc. Med. *62:* 989–993 (1969).

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders III – Revised. p. 222 (APA, Washington 1987).

Boyle, G.J.: Self-report measures of depression: some psychometric considerations. Br. J. clin. Psychol. *24:* 45–59 (1985).

Chait, L.D.; Uhlenhuth, E.H.; Johanson, C.E.: The discriminative stimulus and subjective effects of *d*-amphetamine in humans. Psychopharmacology *86:* 307–312 (1985).

Davis, R.; Freeman, R.; Solyom, L.: Mood and food: an analysis of bulimic episodes. J. psychiat. Res. *19:* 331–335 (1985).

Eastwood, M.R.; Whitton, J.L.; Kramer, P.M.: A brief instrument for longitudinal monitoring of mood states. Psychiat. Res. *11:* 119–125 (1984).

Fahndrich, E.: Biological predictors of antidepressant drug therapy. Psychiat. Dev. *2:* 151–171 (1987).

Folstein, M.F.; Luria, R.E.: Reliability, validity, and clinical application of the visual analogue mood scale. Psychol. Med. *3:* 479–486 (1973).

Haertzen, C.A.: An overview of Addiction Research Center Inventory Scales (ARCI): an appendix and manual of scales (US Government Printing Office, Washington 1974).

Kendall, R.E.; MacKenzie, W.E.; West, C.; McQuire, R.J.; Cox, J.L.: Day-to-day mood change after childbirth: further data. Br. J. Psychiat. *145:* 620–625 (1984).

Lemus, C.Z.; Asnis, G.M.; Halbreich, V.; et al.: Clinical variables and hypothalamic-pituitary-adrenal function in depression: the importance of mood reactivity. J. affect. Disorders *12:* 219–221 (1987).

Luria, R.E.: The validity and reliability of the visual analogue mood scale. J. psychiat. Res. *12:* 51–57 (1975).

Lutz, C.: Depression and the translation of emotional worlds; in Kleinman, Good, Culture and depression, pp. 63–100 (University of California, Berkeley 1985).

McClellan, G.R.: The effects of practice on measures of performance. Human Psychopharm. *2:* 109–118 (1987).

McNair, D.M.; Lorr, M.; Droppleman, L.F.: Profile of mood states manual (EDITS, San Diego 1981).

Mountjoy, C.Q.; Roth, M.: Studies in the relationship between depressive disorders and anxiety states. J. affect. Disorders *4:* 127–147 (1982).

Nelson, J.C.; Mazure, C.; Quinlan, D.M.; Gatlow, P.I.: Drug-responsive symptoms in melancholia. Archs gen. Psychiat. *41:* 663–668 (1984).

Peterson, R.A.; Headon, S.W.: Profile of mood states; in Kerper, Sweetland, Test critiques, vol. 1 (Test Corp of America, Kansas City 1984).

Polivy, J.: On the induction of emotion in the laboratory: discrete moods or multiple affect states? J. Pers. soc. Psychol. *41:* 803–817 (1981).

Shields, S.A.: Distinguishing between emotion and nonemotion: judgements about experience. Motiv Emot. *4:* 355–369 (1984).

Karley Little, MD
Department of Psychiatry
Annex 2, Room 206
Lexington, KY 40536–0084 (USA)