

A Visual Analogue Scale Technique to Measure Global Vigor and Affect

Timothy H. Monk

Received November 30, 1987; revised version received April 6, 1988; accepted May 21, 1988.

Abstract. This article describes an easily administered Visual Analogue Scale (VAS) technique that can be used to detect changes in mood and subjective activation. The method yields two summary measures: Global Vigor (GV) and Global Affect (GA), each ranging in value from 0 to 100. The instrument was administered about six times per day in 38 healthy control and 6 depressed patients participating in temporal isolation studies. This yielded a total of 5,734 control subject sessions and 575 patient sessions. For both groups, frequency distributions of GV and GA were shown to be approximately Gaussian, and evidence was obtained suggesting that the instrument was being completed properly. On average, depressives were about 0.6 standard deviations lower than controls in GV and about 1.5 standard deviations lower in GA, confirming the validity of the scales. Measures of GV in controls were shown to be sensitive to both jet lag and diurnal variation, thus confirming the reliability and validity of GV in these situations. In two studies involving more than 50 days of voluntary seclusion, measures of GA were found to show an almost monotonic decline, tracking the decline in mood and confirming the reliability and validity of GA in that situation.

Key Words. Mood, affect, activation, visual analogue scale.

Very often in psychiatric research one needs to know whether a particular drug or manipulation has upset subjects or made them feel drowsy. There is a need to assess affective state (feelings, mood) and level of vigor (alertness, vigilance) using a quick, easily administered technique. Particularly in within-subject repeated measures designs, the instrument should not be unduly burdensome to subjects, since otherwise their feelings will be influenced by the instrument itself, making for facetious or stereotyped responses. This article describes a rapidly administered Visual Analogue Scale (VAS) technique that yields global measures of vigor and affect, and seeks to demonstrate its validity, reliability, and sensitivity in both healthy controls (5,734 subject sessions) and unipolar depressives (575 patient sessions).

There are a number of different techniques currently available for measuring vigor and affective state. Probably the most widely used is the Profile of Mood States (POMS) (McNair et al., 1971). The POMS requires the subject to rate a level of agree-

Timothy H. Monk, Ph.D., is Associate Professor of Psychiatry and Director, Human Chronobiology Research Program, Sleep Evaluation Center, Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine. (Reprint requests to Dr. T.H.Monk, Sleep Evaluation Center, WPIC, 3811 O'Hara St., Pittsburgh, PA 15213, USA.)

ment/disagreement to a list of 65 adjectives or phrases. Advantages of the POMS are the several different dimensions or factors that result. Disadvantages are the test's longwindedness and occasionally archaic language (e.g., "peevied," "full of pep"). To do the test properly requires time and effort; for many subjects, being repeatedly asked to rate 65 items can become tiresome and irritating, violating the premise of the instrument, which requires an evaluation of feelings *during the past week*. Similar arguments apply to Thayer's (1978) AD-ACL, which uses an equivalent technique.

At the other extreme, there are very rapid techniques to assess vigor, notably the Stanford Sleepiness Scale (Hoddes et al., 1973), in which subjects merely select the single sentence (from a list of 7) that most accurately describes their feelings. Here the problems become coarseness of resolution (most ratings from healthy normal subjects are either 1, 2, or 3), and response stereotyping, by which subjects "decide they are a 2" and respond as that all the way through the study (Weitzman, personal communication).

The present technique sought to avoid the excessive longwindedness of the mood-adjective ratings, while still yielding a fine resolution of response. Its aim was to provide an easily obtained indication of mood and activation, lacking the comprehensiveness of the POMS and Thayer AD-ACL, but still able to detect the effects of various manipulations in both normal healthy subjects and psychiatric patients. The instrument was devised for general use in a temporal isolation facility where subjects are tested approximately six times per day for several weeks at a time. Protocols required that it be difficult for subjects to remember their previous response, thus avoiding the problem of "one step back" comparisons (e.g., "at breakfast I was a 6, so now I must be a 7").

The method chosen was one based on Visual Analogue Scales (VAS). VAS techniques have a long history of use in the evaluation of mood and activation (Freyd, 1923; Bond and Lader, 1974). Essentially they involve presenting subjects with a question, below which appears a line (usually 10-cm long) on which they place a mark. Labels at either end of the line indicate opposite extremes of the mood, and the mark represents the subject's feelings at that time, between the two extremes. Studies by Bond and Lader (1974) have shown the technique to yield data that are approximately Gaussian in their distribution, and thus amenable to parametric tests of significance.

Subjects and patients like the technique because it is brief and does not "push one into yes or no boxes." It has been used both in psychiatric settings (Folstein and Luria, 1973), as well as in studies of continuous operations (Froberg, 1977), shift work (Folkard et al., 1978), intractable pain (Folkard et al., 1976), circadian rhythms in subjective activation (Weitzman et al., 1982; Monk et al., 1983), sleep loss (Carskadon and Dement, 1985), and fatigue levels in chemical plant process controllers (Monk and Embrey, 1981) and commercial airline pilots (Graeber et al., 1986).

Methods

Rationale. The present method is based on eight unipolar VAS ratings, four primarily concerned with subjective activation or vigor (alertness, sleepiness, motivation loss [effort],

and weariness) and four concerned more with feelings or “affective state” (happiness, sadness, calmness, and tension). Each group of four is then summed algebraically to give single global values of vigor (GV) and affective state (GA). The eight unipolar scales all have “very little” at the left end of the 10-cm line and “very much” at the right end. Unipolar scales are preferable to bipolar ones for two reasons. First, they provide a useful check of whether the subject is answering in a sensible way. Thus, alertness should show a negative correlation with sleepiness, happiness with sadness, and calmness with tension. Second, they overcome the problem that some subjects have in conceptualizing gradations of particular moods. Thus, several subjects would always rate themselves as 100% alert whenever awake, but would admit to gradations of *sleepiness* over the waking day, allowing a diurnal variation in subjective activation to be detected using a combination of the two.

The Instrument. A pencil-and-paper version of the instrument is illustrated in Fig. 1. In the actual version, the horizontal lines are exactly 10 cm long. The data to be reported here come from a computerized version of the instrument which is described in the following paragraph. The subject or patient is seated in front of a computer screen with a push-button in each hand. Each of the eight scale questions is presented one at a time on the screen with a 10-cm line and a cursor, which is initially placed off the left end of the line. The cursor has to be moved to complete the trial. Pressing the left-hand button causes the cursor to move left; pressing the right-hand button causes it to move right. The subject is allowed as many movements of the cursor as he likes. When the cursor is stationary on the line for 3 sec, a value corresponding to the distance of the cursor from the left end of the line (1 to 36 arbitrary units) is recorded on the computer, a tone sounds, and the next scale is presented. At the completion of all eight scales, the subject is presented with a further scale measuring well-being (not discussed here) and then a message thanking him for participating. No numerical values are presented to the subject, removing any temptation for him to remember how he responded on a previous trial session. The eight scales are always given exactly in the order they are presented in Fig. 1, thus

Fig. 1. Pencil-and-paper version of Global Vigor and Affect (GVA) instrument

Name _____ Day _____ Date ____/____/____ Time _____

How alert do you feel?
 very little _____ very much

How sad do you feel?
 very little _____ very much

How tense do you feel?
 very little _____ very much

How much of an effort is it to do anything?
 very little _____ very much

How happy do you feel?
 very little _____ very much

How weary do you feel?
 very little _____ very much

How calm do you feel?
 very little _____ very much

How sleepy do you feel?
 very little _____ very much

At actual size the horizontal lines are exactly 10-cm long.

avoiding problems of adjacent opposites (e.g., “happy” immediately following “sad”) which randomized presentation might incur.

Other computer techniques—one involving a light pen, the other involving keyboard input on a personal computer—have also been developed by the author and found to be equally effective.

The Formulas. Each of the eight scales is assigned a value between 0 and 100, corresponding to the distance in mm of the subject’s mark from the left end of the line, or to an equivalent appropriate transformation of the computer value. This, of course, represents a spurious level of accuracy, but is easier to measure than a coarser metric (e.g., cm) which would require rounding decisions. The formulas used are then simply:

$$GV = [(alert) + 300 - (sleepy) - (effort) - (weary)] / 4$$

and

$$GA = [(happy) + (calm) + 200 - (sad) - (tense)] / 4$$

Each formula yields a value between 0 and 100 which is rounded to the nearest whole number.

Control Subject Validation Group. The control subject validation group consisted of healthy paid volunteers who had passed rigorous mental and physical health screening procedures (Monk et al., 1985). Thirty of the subjects were male, five female; ages ranged from 23 to 81 years (mean = 48). These volunteers took part in various forms of temporal isolation experiment. All experiments were conducted in special apartments (Monk et al., 1985), which allowed all external time cues (e.g., windows, clocks, phones, TV) to be eliminated. Subjects lived alone in these apartments for between 15 and 99 calendar days (mean = 23.6), during which they were unaware as to what time of day it was. Social contact was maintained with specially trained technicians who took blood samples if required, brought in meals, and prepared subjects for polysomnography at bedtime. In some protocols, the subject was told throughout when to go to bed, get up, and take meals (usually according to a routine involving phase shifts and/or abnormal day lengths). In others, subjects were allowed at some time to make all their own decisions about these events (“free-running”). However, all protocols started with a baseline portion in which subjects were required to live on a 24-hour routine tailored to be equivalent to their typical routine at home, as determined by a sleep diary kept for the 2 weeks immediately preceding the experiment. The duration of each baseline segment ranged from 4 to 10 days (mean = 5.5) and involved between 21 and 81 test sessions (mean = 41). On average, there were 151 test sessions per whole study, yielding a grand total of 5,734 subject sessions in the 38 studies.

To avoid providing temporal information to the subjects, mood test sessions were tied to particular activities: waketime, bedtime, before and after exercise, after urination, and every time solid food (meal or snack) was requested or scheduled. No test was given within 60 min of another. This yielded about six tests per “day” spread throughout the subject’s waking day. Tests were not given during the subject’s “night.”

Patient Group. The patient group comprised six depressives who all met Research Diagnostic Criteria for recurrent unipolar primary major endogenous depressive disorder (Spitzer et al., 1978) and had scores > 20 on the Hamilton Rating Scale for Depression (Hamilton, 1967). All were outpatients who were medication-free at the time of study. Five of the six patients were female; ages ranged from 23 to 57 years (mean = 43). Experimental procedures for the patient group were exactly the same as those for the control subjects. In each case, the protocol was a “free-running” one (lasting about 18 days), starting as usual with 5 days of entrainment to the patient’s habitual home routine. In some cases, the last few days of the study also included an enforced “afternoon” nap. On average, there were 27 test sessions per baseline segment, 96 test sessions per whole study (range: 56-125).

Subject and Patient Training. Subjects and patients were carefully trained in completing the mood scales using both pencil-and-paper and computerized versions of the instrument. It was explained that the VAS technique avoided the need to put answers into “yes or no boxes” and that they should think of the lines as representing *gradations* of the relevant mood. They were invited to regard the ends of the line as representing *their own* range of feelings, and to avoid comparing themselves to other people. Thus, for example, gradations of sadness would emerge even for naturally happy people whose sadness was “nowhere near as bad as other people’s” (and conversely for the depressives). In particular, subjects and patients were specifically asked to *avoid* simply checking one or other end of the line, but were urged to give quick responses, not pausing to agonize over positioning their marks.

The instructor then went through each of the eight terms, giving synonyms and anecdotal examples such as *alert* = wide awake, vigilant, full of vigor; *sad* = sorrowful, blue; *tense* = uptight, nervous, jittery; *effort* = how hard it is to get yourself going; *happy* = cheerful, contented; *weary* = tired and fed-up, beat; *calm* = relaxed, laid back; *sleepy* = ready for bed. Subjects and patients were given a booklet of 20 practice sessions to be completed at home during the week preceding the study at a rate of about 4 per day. Typically, very few problems emerged; any that did occur were quickly detected and rectified during the practice sessions.

Aims of the Present Analysis. The present analysis had five aims: (1) to test whether negative correlations did indeed occur in the control group between alertness and sleepiness, happiness and sadness, and calmness and tension, as would be predicted if the VAS items were correctly completed; (2) to chart the frequency distributions of the global measures of vigor (GV) and affect (GA) in both (a) transverse and (b) longitudinal control studies; (3) to test the sensitivity of GV to (a) the effect of a 6-hour phase advance (jet lag) and (b) diurnal variation (changes from one time of day to another); (4) to test the sensitivity of GA to changes in control subjects’ mood that result from 8 or more weeks of voluntary seclusion; (5) to verify the usefulness of the instrument in a sample of depressives and to determine how they differ from controls in GV and GA measures.

Results

Aim 1. Two forms of correlational analysis were used. In the first, each control subject’s data were considered separately, and a rank order correlation (Spearman’s rho) was calculated *within that subject* for each of the three pairs of items (*alert* vs. *sleepy*, *happy* vs. *sad*, *calm* vs. *tense*). Ninety-three percent of the 114 correlations were negative. All of the *alert* vs. *sleepy* correlations achieved statistical significance (median rho = -0.71), as did 68% of the *happy* vs. *sad* correlations (median rho = -0.34) and 76% of the *calm* vs. *tense* correlations (median rho = -0.34). A second correlational analysis was concerned with *between*-subjects comparisons. The average value of each of the six items (for the experiment as a whole) was first calculated for each subject, yielding 38 averages for each item. A rank-order correlation coefficient was then calculated for each of the three pairs of items. All were statistically significant ($p < 0.001$) and negative in direction (*alert* vs. *sleepy*: rho = -0.69; *happy* vs. *sad*: rho = -0.59; *calm* vs. *tense*: rho = -0.48). There thus appeared to be little doubt that sensible ratings were being given by the control subjects, whether evaluated on a within- or between-subjects basis.

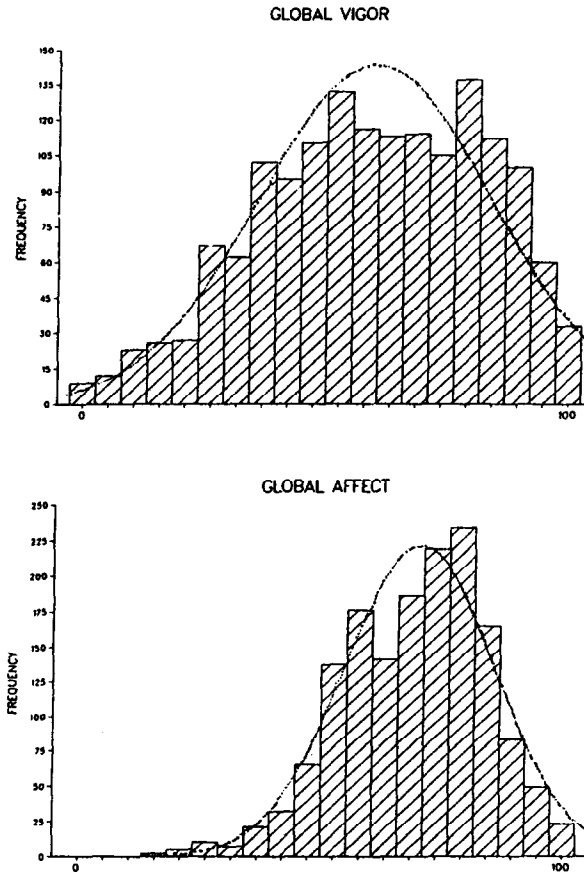
Aim 2. The transverse analysis used only data from the baseline section of each control study. This yielded a grand total of 1,555 test sessions from the 38 studies. Global measures of vigor (GV) and activation (GA) were calculated for each test session, and frequency distributions plotted (Fig. 2). Although not perfectly

Gaussian in their distribution, it is clear that both measures showed a sufficiently normal distribution for parametric tests of significance to be appropriate. The distribution of GV had a mean of 59.0 and SD of 21.7; that of GA, a mean of 68.3 and SD of 14.0 (standard Normal curves with these parameters have been superimposed on the histograms).

Since this analysis included more influence from subjects with a larger number of sessions, a secondary analysis was performed using only the average baseline GV and GA scores from each subject. The distributions of these 38 averages could then be plotted. Again, acceptably Gaussian distributions emerged which were very similar to those obtained from the whole data set, with only the expected reduction in standard deviation (GV: mean = 58.6, SD = 11.8; GA: mean = 68.3, SD = 10.8).

The longitudinal analysis used the data from four young male subjects (aged 19-27 years) who had each contributed at least 393 consecutive test sessions, from studies

Fig. 2. Frequency distributions of Global Vigor (GV) and Global Affect (GA) scores



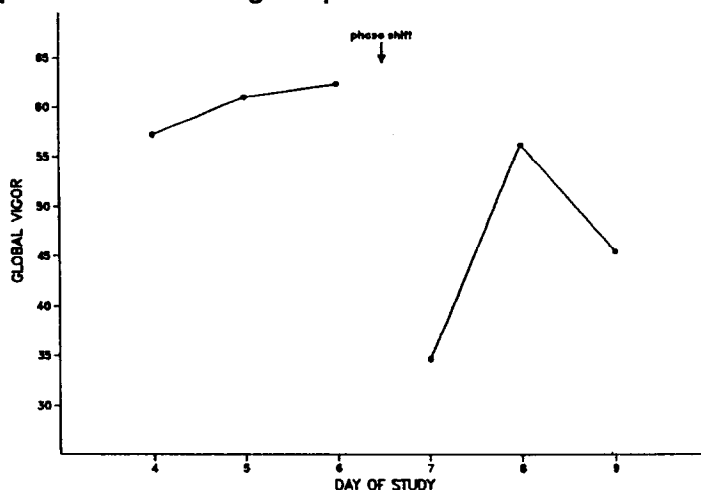
The scores are from 38 baseline studies, which together yielded a total of 1,555 subject sessions. Normal distributions with appropriate mean and variance have been superimposed on the histograms (see text).

lasting at least 62 days. For each study, frequency distributions were calculated for GV and GA. Again, both measures showed acceptably Gaussian distributions (GV means: 55.3, 61.7, 64.5, 61.2; GV SDs: 17.3, 12.9, 17.5, 9.6; GA means: 56.4, 57.2, 55.0, 54.5; GA SDs: 15.1, 10.1, 16.3, 7.0).

Aim 3a. The effects of jet lag were tested using data from eight middle-aged men (mean age = 47 years), who each experienced the same standard 15-day protocol. This protocol kept the subjects on their habitual routine for the first 5 days of the study. On night 6, there was an abrupt 6-hour phase advance in routine, truncating the sleep episode to about 1 or 2 hours. The phase change was unheralded, and only four of the subjects realized what had happened.

The present analysis was based on daily mean scores of GV from the last 3 days before the phase change (days 4, 5, and 6), comparing them to the first 3 days after it (days 7, 8, and 9). The analysis of variance (ANOVA) thus had “days” (3 levels) and “condition” (*before vs. after* phase shift) as within-subject factors. Results are illustrated in Fig. 3. Statistically significant deleterious effects of the phase shift were observed ($F = 80.62$; $df = 1, 7$; $p < 0.001$) in a zig-zag adjustment pattern characteristic of phase advances (Monk, 1987a). This finding would thus appear to confirm the sensitivity of GV to jet-lag effects, and to confirm its validity in this situation. (The equivalent effect in GA failed to achieve statistical significance [$F = 3.36$; $df = 1, 7$; $p = 0.11$].)

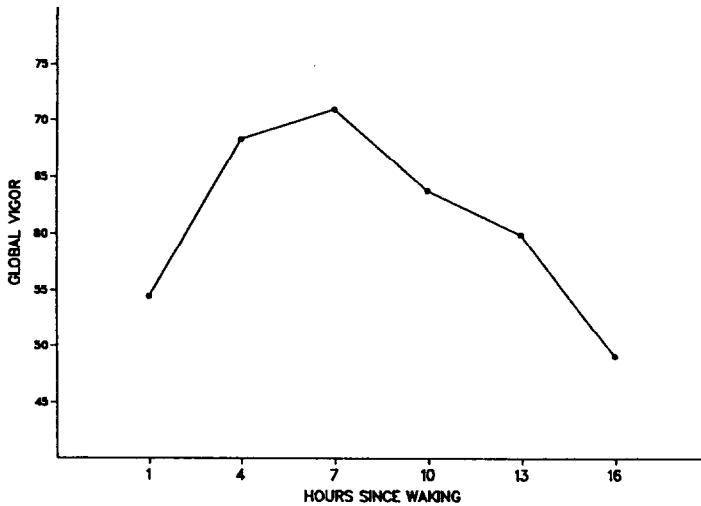
Fig. 3. Mean Global Vigor (GV) scores from 3 days before and 3 days after a 6-hour phase shift involving sleep reduction



The average from 8 middle-aged men is plotted.

Aim 3b. Diurnal variation in GV was examined using baseline data from the 38 control studies. Data were cast into six time-of-day “bins.” The first bin contained data collected within the first hour after waking. Subsequent bins had a time interval width of 3 hours. Each study thus yielded one composite time-of-day function. The average function from all 38 studies is illustrated in Fig. 4. One-way ANOVA

Fig. 4. Average time-of-day function in Global Vigor (GV) from 38 baseline studies

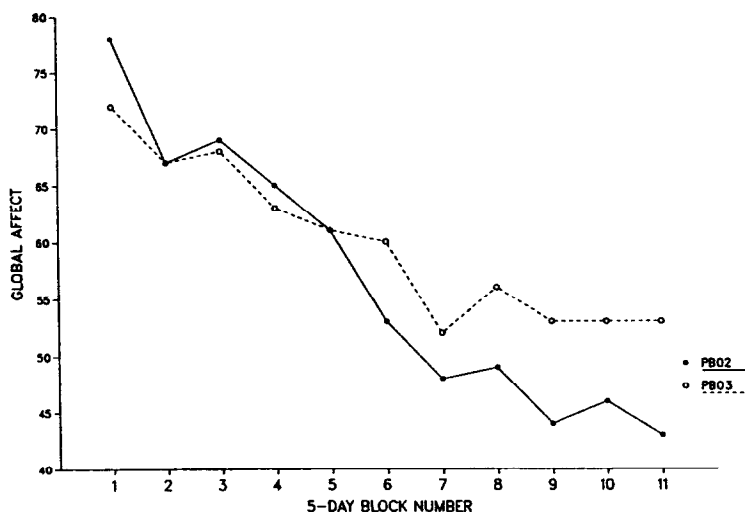


confirmed an extremely robust time-of-day effect in GV ($F = 27.29$; $df = 5, 185$; $p < 0.001$) with a mid-day peak, replicating the time-of-day function in subjective alertness observed in previous studies (Thayer, 1978; Folkard, 1982; Monk et al., 1983; Monk, 1987b). In addition to confirming the sensitivity of GV to diurnal variation, this also confirmed the test-retest reliability of the instrument, since several days at a time were being averaged to yield a given time-of-day score.

Aim 4. The analysis of mood changes associated with several weeks of voluntary seclusion used data from the longer-term studies. Five-day blocks of data were taken, smoothing out the finer-grained irregularities. In some studies, the pattern of GA changes over the course of the experiment was rather erratic, either due to the particular manipulations of the protocol, or to other changes (letters from home, lost laundry) that are hard to quantify. There were, however, two subjects who showed a gradual deterioration in their mood which was clearly mapped by ever-lowering GA values (Fig. 5). When the GA value was correlated with "5-day block No.," significant negative correlations emerged for both subjects ($\rho = -0.973, -0.890$, $p < 0.001$ both cases). On a priori grounds, one could assert (using the prison system as an example) that seclusion for $N + 1$ weeks is worse than seclusion for N weeks. Thus, at least on a 5-day block basis, seclusion can be considered as likely to upset one in a monotonically increasing manner. This result would therefore seem to confirm both the sensitivity of GA to the effects of seclusion and the validity of GA as a measure in this situation.

Aim 5. As found in earlier VAS studies (Folstein and Luria, 1973), patients had few problems in completing the eight VAS scales, and liked the technique. As in Aim 1, the correlational analysis was primarily concerned with relationships between the items *within* a given patient. Thus, a separate rank order correlation was calculated for each patient. In all six patients, there was a significant ($p < 0.02$, all cases)

Fig. 5. Deterioration in Global Affect (GA) score over 55 consecutive days of voluntary incarceration, for subjects PB02 (solid line) and PB03 (broken line)



Each point represents the mean of 5 days (about 30 test sessions).

negative correlation between *alertness* and *sleepiness* ($\rho = -0.81, -0.38, -0.61, -0.89, -0.90, -0.55$), *happiness* and *sadness* ($\rho = -0.74, -0.36, -0.44, -0.21, -0.58, -0.84$), and *calmness* and *tension* ($\rho = -0.83, -0.57, -0.31, -0.57, -0.55, -0.75$). Thus, it would seem that these patients were indeed completing the GVA scales sensibly.

Frequency distributions for GV and GA across all 575 patient-sessions were acceptably Gaussian, with SDs that were very similar to those of the control group (GV: patient = 21.2, control = 21.7; GA: patient = 16.1, control = 14.0), but mean levels that were substantially lower (GV: patient = 47.2, control = 59.0; GA: patient = 45.6, control = 68.3). Thus, on average, patient values were about 0.6 SD lower than control values in GV and 1.5 SD lower in GA, reflecting the decreased vigor and impaired affect characteristic of the depressed and confirming the reliability of the scales. Time-of-day functions in GV and GA for the patient sample failed to achieve statistical significance.

Discussion

As stated at the beginning of this article, it is not the aim of this test to assess all of the various facets and attributes of mood and subjective activation. Rather, its aim is merely to provide global indications of mood and activation (and thus changes therein) which other, more precise instruments might later explore in detail. This analysis has deliberately avoided applying any form of clustering or factor analysis to the present data. Such techniques might be appropriate for more sophisticated instruments, but are clearly unwarranted for the present, rather simplistic technique.

As with most VAS techniques, the GVA is primarily intended as a tool for detecting changes within an individual (from one time and/or condition to another) rather than between different individuals. It thus came as a pleasant surprise that

simple differences between depressives and controls did emerge for both GV and GA, representing a sensitivity which may be very useful. Investigators should be very careful, however, in concluding that there are no differences in vigor or affect levels between groups simply because they have not emerged in GV or GA levels. More valid would be experimental designs in which the study looked at "interaction" effects (e.g., between time of day and depressive/control, or between drug/placebo and depressive/control) because these are inherently "within-subject" comparisons.

Despite these constraints, the GVA instrument would appear to be valid and reliable, as well as sensitive enough to detect various effects. On average, the eight VAS scales take less than a minute to complete, and subjects and patients thus tolerate it very well, even in studies involving literally hundreds of trials. As other VAS proponents have testified, subjects and patients do like the brevity and simplicity of this technique, accruing benefits to the researcher both in compliance rates and in the accuracy of the subsequent data.

Acknowledgment. Grateful thanks are due to my former colleagues at Cornell University Medical College who collaborated in the collection of data, and in whose studies the mood scales were given. Dr. A.J. Spielman collaborated in the initial conceptualization of the eight mood scales. Miss S.M. Peetz assisted greatly in analyzing the data and preparing the figures. Data collection was supported by grants AG-04135, MH-37814, and NCC-2-253.

References

- Bond, A., and Lader, M. The use of analogue scales in rating subjective feelings. *British Journal of Medical Psychology*, 47:211-218, 1974.
- Carskadon, M.A., and Dement, W.C. Sleep loss in elderly volunteers. *Sleep*, 8:207-221, 1985.
- Folkard, S. Circadian rhythms and human memory. In: Brown, F.M., and Graeber, R.C., eds. *Rhythmic Aspects of Behavior*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1982.
- Folkard, S.; Glynn, C.J.; and Lloyd, J.W. Diurnal variation and individual differences in the perception of intractable pain. *Journal of Psychosomatic Research*, 20:289-301, 1976.
- Folkard, S.; Monk, T.H.; and Lobban, M.C. Short and long-term adjustment of circadian rhythms in "permanent" night nurses. *Ergonomics*, 21:785-799, 1978.
- Folstein, M.F., and Luria, R. Reliability, validity, and clinical application of the visual analogue mood scale. *Psychological Medicine*, 3:479-486, 1973.
- Freyd, M. The graphic rating scale. *Journal of Educational Psychology*, 14:83-102, 1923.
- Froberg, J.E. Twenty-four hour patterns in human performance, subjective and physiological variables and differences between morning and evening active subjects. *Biological Psychology*, 5:119-134, 1977.
- Graeber, R.C.; Lauber, J.K.; Connell, L.J.; and Gander, P.H. International aircrew sleep and wakefulness after multiple time zone flights: A cooperative study. *Aviation, Space, and Environmental Medicine*, 57:B3-B9, 1986.
- Hamilton, M. Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology*, 6:278-296, 1967.

Hoddes, E.; Zarcone, V.; Smythe, H.; Phillips, R.; and Dement, W.C. Quantification of sleepiness: A new approach. *Psychophysiology*, 10:431-436, 1973.

McNair, D.M.; Lorr, M.; and Droppleman, L.F. *Profile of Mood States*. San Diego, CA: Educational and Industrial Testing Service, 1971.

Monk, T.H. Coping with the stress of jet-lag. *Work & Stress*, 1:164-166, 1987a.

Monk, T.H. Subjective ratings of sleepiness—The underlying circadian mechanisms. *Sleep*, 10:343-353, 1987b.

Monk, T.H., and Embrey, D.E. A field study of circadian rhythms in actual and interpolated task performance. In: Reinberg, A.; Vieux, N.; and Andlauer, P., eds. *Night and Shift Work: Biological and Social Aspects*. Oxford: Pergamon Press, 1981.

Monk, T.H.; Fookson, J.E.; Kream, J.; Moline, M.L.; Pollak, C.P.; and Weitzman, M.B. Circadian factors during sustained performance: Background and methodology. *Behavior Research Methods, Instruments & Computers*, 17:19-26, 1985.

Monk, T.H.; Leng, V.C.; Folkard, S.; and Weitzman, E.D. Circadian rhythms in subjective alertness and core body temperature. *Chronobiologia*, 10:49-55, 1983.

Spitzer, R.; Endicott, J.; and Robins, E. Research Diagnostic Criteria: Rationale and reliability. *Archives of General Psychiatry*, 34:773-782, 1978.

Thayer, R.E. Toward a psychological theory of multidimensional activation (arousal). *Motivation and Emotion*, 2:1-34, 1978.

Weitzman, E.D.; Czeisler, C.A.; Zimmerman, J.C.; Ronda, J.M.; and Knauer, R.S. Chronobiological disorders: Analytic and therapeutic techniques. In: Guilleminault, C., ed. *Sleeping and Waking Disorders: Indications and Techniques*. Menlo Park, CA: Addison-Wesley, 1982.