The PANAS-X: Manual for the Positive and Negative Affect Schedule - Expanded Form

Watson, David; Clark, Lee Anna

Watson, D., & Clark, L. A. (1994). The PANAS-X: Manual for the Positive and Negative Affect Schedule - Expanded Form. https://doi.org/10.17077/48vt-m4t2

updated 8/99

THE PANAS-X

Manual for the

Positive and Negative Affect Schedule - Expanded Form

The University of Iowa

Table of Contents

List of Tables

I. The Hierarchical Structure of Self-Rated Affect

In recent research, two broad, general factors--typically labeled Positive Affect (PA) and Negative Affect (NA)--have emerged reliably as the dominant dimensions of emotional experience. These factors have been identified in both intra- and interindividual analyses, and they emerge consistently across diverse descriptor sets, time frames, response formats, languages, and cultures (Almagor & Ben-Porath, 1989; Mayer & Gaschke, 1988; Meyer & Shack, 1989; Tellegen, 1985; Watson, 1988b; Watson, Clark, & Tellegen, 1984; Watson & Tellegen, 1985; Zevon & Tellegen, 1982). To measure these factors, Watson, Clark, and Tellegen (1988) developed the Positive and Negative Affect Schedule (PANAS), which consists of two 10-item scales for PA and NA, respectively.

These two general dimensions account for most of the variance in self-rated affect--together they account for roughly one-half to three-quarters of the common variance in mood terms (see Watson, 1988b; Watson & Tellegen, 1985). Nevertheless, specific emotional states can also be identified in the same data. In fact, on the basis of earlier work by Tellegen , Watson and Tellegen (1985) proposed a hierarchical taxonomic scheme in which the two broad, higher order dimensions are each composed of several correlated, yet ultimately distinguishable affective states (see also Watson & Clark, 1989, 1992a). In this model, the higher level reflects the *valence* of the mood descriptors (i.e., whether they represent negative or positive states), whereas the lower level reflects their specific *content* (i.e., the distinctive qualities of the individual affects).

To assess these specific emotional states, we have created a 60-item, expanded version of the PANAS (the PANAS-X). In addition to the two original higher order scales, the PANAS-X measures 11 specific affects: Fear, Sadness, Guilt, Hostility, Shyness, Fatigue, Surprise, Joviality, Self-Assurance, Attentiveness, and Serenity. The PANAS-X thus provides for mood measurement at two different levels.

The PANAS-X is simple and easy to administer. Most subjects complete the entire 60-item schedule in 10 minutes or less. However, investigators facing more severe time constraints can select and assess only those scales that are most relevant to their research. A sample PANAS-X protocol is shown in Table 1. In addition, the terms comprising each of the PANAS-X scales are shown in Table 2.

The PANAS-X was created in three relatively distinct stages. As mentioned, the two higher order scales were developed first. Seven specific affect scales primarily involving

---

Table 1 *Sample* PANAS-X *Protocol Illustrating "Past Few Weeks" Time Instructions*

This scale consists of a number of words and phrases that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent you have felt this way during the past few weeks. Use the following scale to record your answers:

1 very slightly or not at all	2 a little	3 moderately	4 quite a bit	5 extremely
_____ cheerful	_____ sad	_____ active		_____ angry at self
_____ disgusted	_____ calm	_____ guilty		_____ enthusiastic
_____ attentive	_____ afraid	_____ joyful		_____ downhearted
_____ bashful	_____ tired	_____ nervous		_____ sheepish
_____ sluggish	_____ amazed	_____ lonely		_____ distressed
_____ daring	_____ shaky	_____ sleepy		_____ blameworthy
_____ surprised	_____ happy	_____ excited		_____ determined
_____ strong	_____ timid	_____ hostile		_____ frightened
_____ scornful	_____ alone	_____ proud		_____ astonished
_____ relaxed	_____ alert	_____ jittery		_____ interested
_____ irritable	_____ upset	_____ lively		_____ loathing
_____ delighted	_____ angry	_____ ashamed		_____ confident
_____ inspired	_____ bold	_____ at ease		_____ energetic
_____ fearless	_____ blue	_____ scared		_____ concentrating
_____ disgusted with self	_____ shy	_____ drowsy		_____ dissatisfied with self

---

Table 2 *Item Composition of the* PANAS-X *Scales*

General Dimension Scales	
Negative Affect (10)	afraid, scared, nervous, jittery, irritable, hostile, guilty, ashamed, upset, distressed
Positive Affect (10)	active, alert, attentive, determined, enthusiastic, excited, inspired, interested, proud, strong
Basic Negative Emotion Scales	
Fear (6)	afraid, scared, frightened, nervous, jittery, shaky
Hostility (6)	angry, hostile, irritable, scornful, disgusted, loathing
Guilt (6)	guilty, ashamed, blameworthy, angry at self, disgusted with self, dissatisfied with self
Sadness (5)	sad, blue, downhearted, alone, lonely
Basic Positive Emotion Scales	
Joviality (8)	happy, joyful, delighted, cheerful, excited, enthusiastic, lively, energetic
Self-Assurance (6)	proud, strong, confident, bold, daring, fearless
Attentiveness (4)	alert, attentive, concentrating, determined
Other Affective States	
Shyness (4)	shy, bashful, sheepish, timid
Fatigue (4)	sleepy, tired, sluggish, drowsy
Serenity (3)	calm, relaxed, at ease
Surprise (3)	amazed, surprised, astonished

Note. The number of terms comprising each scale is shown in parentheses.

different negative affects were constructed next. Finally, four specific positive affect scales were created. In this manual, we recap briefly the development of the original PANAS scales, and present their basic reliability and validity data (the interested reader is referred to Watson et al., 1988, for more detail). We then describe the development of the specific affect (PANAS-X) scales, and present data supporting their reliability and validity. Finally, we discuss several important issues regarding the instrument as a whole.

II. The Higher Order Scales

A. Construction of the Original Positive and Negative Affect Scales

The goal in developing these scales was to create reliable and valid measures that were also brief and simple to administer. The primary concern was to select descriptors that were relatively pure markers of either Negative Affect or Positive Affect; that is, terms that had a substantial loading on one factor but a near-zero loading on the other. As a starting point, we used the 60 terms included in the factor analyses reported by Zevon and Tellegen (1982). Tellegen constructed this set from an initial pool of 117 affective words and phrases derived from the earlier studies of Izard (1972), Nowlis (1965), Zuckerman and Lubin (1965), and Ekman (1971). A principal components analysis of content sortings of this larger pool of items identified 20 synonym groups, and the final list of 60 terms was constructed by choosing three marker terms from each content group (see Zevon & Tellegen, 1982, Table 1). Thus, these terms provide a comprehensive assessment of the affective lexicon.

From this list of 60 terms we selected those descriptors that had an average loading of .40 or greater on the relevant factor across both the within- and between-subjects analyses reported in Zevon and Tellegen (1982). Twenty Positive Affect and 30 Negative Affect markers met this initial criterion. However, as noted previously, we were also concerned that the terms be relatively pure markers of a factor. We therefore specified that the terms not have a secondary loading of |.25| or greater in either analysis. This reduced the pool of candidate descriptors to 12 for Positive Affect and 25 for Negative Affect.

Preliminary reliability and validity analyses indicated that 10 terms were sufficient for the higher order Positive Affect scale. We therefore dropped two terms that had relatively higher secondary loadings on the Negative Affect factor, yielding the final set of 10 descriptors (shown in Table 2).

The 25 Negative Affect candidate terms included all 3 terms from seven of Tellegen's content categories, plus 2 from each of two others. Because we wanted to tap a broad range of content, we constructed a preliminary 14-item scale that included 2 terms from each of the seven complete triads. We found, however, that the terms from the Contempt and Revulsion content categories did not significantly enhance the reliability and validity of the scale. Moreover, these

terms were less familiar to our subjects (primarily undergraduates) and were occasionally left unanswered. Therefore, the final 10-item version consisted of 2 terms each from the other five triads (displayed in Table 2).

B. Normative and Internal Consistency Data

Between-subjects data. Most of our basic psychometric data were gathered from undergraduates enrolled in various psychology courses at Southern Methodist University (SMU), a private southwestern university. However, we (and others) have also collected data on various student, adult and psychiatric patient samples, and we present these results as well.

We have obtained PANAS-X ratings using eight different temporal instructions. Subjects have rated how they felt: (a) "right now (that is, at the present moment)" (*Moment* instructions); (b) "today" (*Today*); (c) "during the past few days" (*Past Few Days*); (d) "during the past week" (*Past Week*); (e) "during the past few weeks (*Past Few Weeks*); (f) "during the past month" (*Past Month*), (g) "during the past year (*Past Year*); and (h) "in general, that is, on the average" (*General*).

Table 3 *Means and Standard Deviations for the General Positive Affect and Negative Affect Scales as a Function of Rated Time Frame and Subject Population*

		Positive Affect		Negative Affect	
Time Frame/Sample	*N*	*M*	*SD*	*M*	*SD*
Moment					
SMU undergraduates	2,213	29.0	8.0	15.8	5.9
Australian undergraduates[a]	279	25.3	7.2	15.2	5.0
VA substance abusers	158	32.0	8.5	23.7	10.1
Psychiatric inpatients	56	32.6	7.8	21.1	9.0
Today					
SMU undergraduates	1,664	30.1	8.2	17.6	7.0
Past Few Days					
SMU undergraduates	1,577	33.7	7.3	18.5	6.9
Iowa undergraduates	502	31.5	7.4	21.7	7.4
Past Week					
SMU undergraduates	1,521	32.4	7.3	20.4	7.0
Dallas-area adults	328	31.1	7.5	18.0	7.1
Australian adult men[a]	114	32.6	5.9	16.3	4.7
Australian adult women[a]	115	30.7	7.1	15.8	4.9
Past Few Weeks					
SMU undergraduates	2,076	32.6	7.1	20.2	7.2
SMU employees	164	33.1	6.8	17.9	6.4
Past Month					
SMU undergraduates	1,006	34.5	7.2	20.2	7.3
Past Year					
SMU undergraduates	964	35.9	6.4	22.8	6.6
General					
SMU undergraduates	3,622	35.7	6.2	19.5	6.0
SMU employees	202	35.1	7.4	18.9	6.4
Detroit-area adults[b]	815	36.0	6.0	18.2	6.3
Australian adult men[a]	114	33.5	5.9	14.2	4.1
Australian adult women[a]	115	33.9	5.1	15.5	5.3
Psychiatric inpatients	117	32.4	8.1	25.5	10.0
Mixed clinical sample	107	30.2	6.6	26.3	9.0

Note. These data include those reported in Watson et al. (1988), Table 1.

[a]Unpublished data reported by Ross Wilkinson, The Australian National University, April, 1993.

[b]These data are reported in Quinn (1989).

Table 3 presents basic descriptive statistics for the general Positive Affect and Negative Affect scales for these different time instructions in various samples. We have not found any large or consistent gender differences on these scales (the issue of gender differences will be discussed in more detail subsequently), so our own data are collapsed across gender. Nevertheless, we recommend testing for gender differences in any new (especially nonstudent) sample.

Inspecting Table 3, one sees that subjects report more Positive Affect than Negative Affect, regardless of the rated time frame or the population studied. Moreover, mean scores on both scales tend to increase as the rated time frame lengthens. This pattern is not surprising: As the rated time period increases, the probability that a subject will have experienced a significant amount of a given affect also increases. Finally, it is interesting to compare the mean scores of the college student, adult, and psychiatric patient samples. Table 3 indicates that the mean Positive Affect scores of college students and adults are generally quite similar. The findings for Negative Affect are more variable, but the bulk of the data suggest that adults report slightly lower levels of Negative Affect than college students. Overall, these data suggest that normative data collected on college student samples can be applied to community-dwelling adult samples with some confidence, particularly with regard to Positive Affect.

Table 4 *Internal Consistency Reliabilities (Coefficient Alpha) and Intercorrelations of the General Positive Affect and Negative Affect Scales*

Time Frame/Sample	*N*	Coefficient α		Scale Intercorrelation
		Positive Affect	Negative Affect	
Moment				
SMU undergraduates	2,213	.88	.85	-.06
VA substance abusers	158	.87	.91	-.21
Psychiatric inpatients	56	.83	.88	-.23
Today				
SMU undergraduates	1,664	.89	.87	-.05
Past Few Days				
SMU undergraduates	1,577	.88	.87	-.17
Iowa undergraduates	502	.88	.87	-.31
Past Week				
SMU undergraduates	1,521	.88	.85	-.14
Dallas-area adults	328	.90	.90	-.38
Australian adults[a]	229	.86	.79	-.07
Past Few Weeks				
SMU undergraduates	2,076	.87	.87	-.13
SMU employees	164	.86	.87	-.09
Past Month				
SMU undergraduates	1,006	.89	.89	-.15
Past Year				
SMU undergraduates	964	.87	.84	-.23
General				
SMU undergraduates	3,622	.87	.85	-.13
SMU employees	201	.90	.87	-.31
Detroit-area adults[b]	815	.84	.88	-.28
Australian adults[a]	229	.85	.87	.01
Psychiatric inpatients	117	.87	.93	-.33
Mixed clinical sample	107	.83	.89	-.32

Note. These data include those reported in Watson et al. (1988), Table 2.

[a]Unpublished data reported by Ross Wilkinson, The Australian National University, April, 1993.

[b]These data are reported in Quinn (1989).

The psychiatric patients, however, clearly represent a distinct population. Most notably, consistent with previous research (e.g., Watson & Clark, 1984; Watson, Clark, & Carey, 1988), psychiatric patients report significantly higher levels of Negative Affect than do nonpatients. The data for Positive Affect are less clear, but one can anticipate that many patients will obtain unusually low scores on this scale.

Internal consistency reliabilities (Cronbach's coefficient alpha) for the two higher order scales, and also their intercorrelations, are presented in Table 4. The alpha reliabilities for both scales are high, generally ranging from .83 to .90 for Positive Affect, and from .85 to .90 for Negative Affect. It is noteworthy that the reliabilities of the scales essentially are unaffected by the time instructions that are used or by the type of subject population (student, adult, or patient) that is assessed.

Table 4 also shows that the correlation between the Positive Affect and Negative Affect scales is generally low, typically ranging from -.05 to -.35. These discriminant values indicate quasi-independence, an attractive feature for many research and data-analytic purposes, and they are somewhat lower than those of many other brief measures of Negative Affect and Positive Affect (see Watson, 1988b). Interestingly, the scale intercorrelation is not systematically influenced by the rated time frame. Thus, the higher order scales maintain their quasi-independence regardless of whether state (shorter term) or trait (longer term) affect is assessed.

Aggregated within-subject data. The data we have presented thus far have been derived from single, between-subjects assessments of large subject samples. Some investigators, however, may wish to use the higher order PANAS-X scales in designs that necessitate repeated within-subject assessments. Accordingly, we present basic descriptive statistics from three groups of subjects who completed the higher order PANAS-X scales (using Today instructions) on a daily basis over a period of several weeks. All subjects completed a minimum of 30 daily mood assessments; all of the assessments were completed in the evening, so that the ratings would provide a reasonable estimate of the subjects' moods over the course of the day (for more details regarding this type of design, see Kennedy-Moore, Greenberg, Newman, & Stone, 1992; Watson, 1988a; Watson, Clark, McIntyre, & Hamaker, 1992).

The three groups consisted of: (1) 425 SMU undergraduates, who completed an average of 43.3 mood assessments; (2) 90 adult men (*M* age = 42.7 years), who completed an average of 77.9 mood ratings; and (3) 27 adults (age range = 21 to 76 years) diagnosed with chronic fatigue syndrome (CFS). The CFS patients were drawn from the Minnesota Regional Chronic Fatigue Syndrome Research Program at the Hennepin County Medical Center (see Marshall et al., 1994, for more details); they completed an average of 53.8 mood assessments.

Mean Positive Affect and Negative Affect scores initially were calculated for each subject. Overall sample means and standard deviations were then computed for each scale; these values are shown in Table 5. Looking first at the college student data, it is noteworthy that the means for both Positive Affect and Negative Affect are comparable--but slightly lower--than the corresponding single-assessment between-subjects values (using Today instructions) that are shown in Table 3. Interestingly, the adult men reported somewhat lower levels of both affects than did the college students. The most striking data, however, are those of the CFS patients. Their mean Positive Affect score was substantially lower than that of the adult men, and more than a full standard deviation below that of the college students. In marked contrast, their Negative Affect levels were unremarkable, falling in between those of the college students and adults. Thus, these findings indicate clearly that CFS is associated with a marked reduction in positive emotional experiences (for further discussion of these data, see Marshall et al., 1994).

Table 5 *Descriptive Statistics for the General Positive Affect and Negative Affect Scales Derived from Mean Daily Mood Scores*

Sample	No. of Subjects	Total No. of Observations	Positive Affect		Negative Affect	
			M	*SD*	*M*	*SD*
SMU undergraduates	425	18,420	28.3	5.9	16.4	4.1
Adult men[a]	90	7,013	25.0	7.4	12.7	3.7
CFS patients	27	1,453	20.1	6.1	15.3	5.6

Note. CFS = Chronic Fatigue Syndrome. See text for more details.

[a]These data are reported in Kennedy-Moore et al. (1992), Table I.

Table 6 *Correlations Between the General Positive Affect and Negative Affect Scales and Regression-Based Scores on the First Two Varimax Factors in Six Samples Assessed with Tellegen's Set of 60 Mood Descriptors*

		Positive Affect Scale Correlations		Negative Affect Scale Correlations	
Rated Time Frame	*N*	Factor 1	Factor 2	Factor 1	Factor 2
Moment	660	-.02	.95	.91	-.15
Today	657	-.02	.95	.93	-.11
Past Few Days	1,002	-.15	.92	.93	-.10
Past Few Weeks	586	-.10	.92	.92	-.18
Past Year	649	-.17	.89	.93	-.09
General	663	-.08	.94	.93	-.12

Note. This table is adapted from Watson et al. (1988, Table 4). See text for details.

C. Construct Validity

Factorial validity. An important step in validating these general scales was to demonstrate that they captured the underlying higher order dimensions adequately. We have examined this issue in three series of analyses. First, we subjected ratings on Tellegen's 60 mood descriptors (see Zevon & Tellegen, 1982) to a principal factor analysis (squared multiple correlations as the initial communality estimates) in six large data sets, each of which was based on a different rated time frame. The *N*s were 660 (Moment), 657 (Today), 1,002 (Past Few Days), 586 (Past Few Weeks), 649 (Year), and 663 (General).

Two dominant factors emerged in each solution that jointly accounted for roughly two-thirds of the common variance, ranging from 62.8% in the Moment solution to 68.7% in the General ratings. The first two factors in each solution were then rotated using varimax. Each of the six solutions generated two sets of factor scoring weights that were used to compute regression-based estimates of the underlying Negative Affect and Positive Affect factors. Within each data set, we then correlated these estimated factor scores with the PANAS-X Negative Affect and Positive Affect scales.

The results, shown in Table 6, demonstrate the expected convergent/discriminant pattern: Both scales are very highly correlated with their corresponding regression-based factor scores in each solution, with convergent correlations ranging from .89 to .95. In contrast, the discriminant correlations are quite low, ranging from only -.02 to -.18. Furthermore, Watson et al. (1988) present other data indicating that the PANAS-X scales offer a better convergent/discriminant correlational pattern with the underlying factors than do other commonly used scale pairs.

The second series of analyses involved 10 additional data sets. These analyses were identical to those already discussed, except that the subjects were assessed on the 60 PANAS-X terms, rather than Tellegen's set of mood descriptors. Eight of the data sets were collected from SMU undergraduates; they differed only on the rated time frame. The *N*s were 1,027 (Moment), 1,007 (Today), 289 (Past Few Days), 1,278 (Week), 678 (Past Few Weeks), 1,006 (Past Month), 315 (Past Year), and 1,657 (General). The ninth data set consisted of 502 University of Iowa undergraduates who rated themselves using Past Few Days instructions. The final sample was composed of 328 Dallas-area adults who rated their mood using Past Week instructions.

The results of these analyses are presented in Table 7 and it can be seen that they replicate closely those of the first series. Again, the convergent correlations were quite high, ranging from .90 to .95 for Positive Affect, and from .92 to .95 for Negative Affect. As before, the discriminant correlations generally were quite low, ranging from -.02 to -.28 for Positive Affect, and from .00 to -.16 for Negative Affect. In other words, these data again demonstrate that the general Negative Affect and Positive Affect scales of the PANAS-X are excellent measures of the underlying higher order factors.

The data presented in Tables 6 and 7 are based entirely on between-subjects data. In contrast, the final series of analyses examined whether the general PANAS-X scales are also excellent measures of the two higher order factors that emerge in within-subject data. To investigate this issue we conducted factor analyses in two very large data sets. The first consisted of 226 SMU undergraduates who rated their current, momentary mood on the 60 PANAS-X terms repeatedly over a 1-2 month period (M = 45.0 assessments per subject). The subjects were instructed to complete one mood assessment each day; the times for these ratings varied from day to day according to a pre-arranged, randomized

Table 7 *Correlations Between the General Positive Affect and Negative Affect Scales and Regression-Based Scores on the First Two Varimax Factors in Ten Samples Assessed with the 60* PANAS-X *Mood Descriptors*

		Positive Affect Scale Correlations		Negative Affect Scale Correlations	
Rated Time Frame	*N*	Factor 1	Factor 2	Factor 1	Factor 2
SMU undergraduates					
Moment	1,027	-.02	.94	.94	-.05
Today	1,007	-.05	.95	.94	.00
Past Few Days	289	-.01	.93	.95	-.01
Past Week	1,278	-.09	.94	.94	-.03
Past Few Weeks	678	-.09	.92	.93	-.08
Past Month	1,006	-.12	.94	.95	-.05
Past Year	315	-.17	.90	.92	-.05
General	1,657	-.10	.93	.93	-.01
Iowa undergraduates					
Past Few Days	502	-.19	.93	.94	-.16
Dallas-area adults					
Past Week	328	-.28	.92	.93	-.15

schedule. The second sample consisted of 248 SMU undergraduates who rated their mood on the full PANAS-X on a daily basis (using Today instructions) over a 1-2 month period (M = 45.7 assessments per subject); as in previous analyses of daily affect, all ratings were made in the evening so that they would provide a reasonable estimate of the subjects' moods over the course of the day. All subjects in both data sets completed a minimum of 35 mood assessments.

The mood ratings in each sample were standardized on a within-subject basis; that is, each subject's responses were converted to standard scores (M = 0, SD = 1). This procedure eliminates between-subjects variability, so that subsequent analyses reflect only within-subject variation. The data in each sample were then subjected to a separate principal factor analysis (squared multiple correlations in the diagonal); note that this is equivalent to factoring each subject's data separately and then averaging the resulting solutions in each sample. The first two factors in each solution were then rotated using varimax; as before, regression-based estimates of these factors were computed. Finally, within each data set, we correlated these factor scores with the PANAS-X Negative Affect and Positive Affect scales. The results (shown in Table 8) again strongly demonstrate the factorial validity of these scales. Replicating the between-subjects results, the convergent correlations for the Positive Affect scale were .93 and .90 in the Moment and Today data, respectively; the corresponding values for the Negative Affect scale were .89 and .89, respectively. Moreover, the discriminant coefficients were low, ranging from -.16 to -.23. Thus, we see clear evidence that the general PANAS-X scales are excellent measures of the major dimensions underlying intraindividual mood experience.

Convergence with peer ratings. The data we have presented thus far consist entirely of subjects' self-reports. In order to establish construct validity, however, it is important to consider other types of evidence as well. Accordingly, we have conducted two studies in which self-ratings on the two higher order scales (as well as the specific affect scales, to be discussed later) were correlated with corresponding judgments made by well-acquainted peers. All affect ratings in both studies were based on trait (i.e., General) time instructions. The first study (discussed in more detail by Watson & Clark, 1991) involved dormitory residents at SMU. To participate in the study, subjects were required to sign up in five-person groups, with the additional proviso that all group members know each other reasonably well. Each subject rated all five group members, thereby generating one set of self-ratings and four sets of peer-ratings. The results reported here were computed from the data of 89 subjects, each of whom was rated by at least three well-acquainted peers; the peer judgments for each subject were averaged to yield a single overall peer rating score on each scale for that subject.

The second study was an examination of currently dating couples at various universities in the Dallas-Fort Worth metropolitan area. To be included in the study, the couple had to have been dating each other for at least one month (M time of dating = 21.8 months). All subjects rated both themselves and their partners on the full PANAS-X.

Before presenting these results, we should note that one can expect only moderate self-peer convergence on these scales. Several studies have found that more externally visible behavioral dispositions show better self-peer convergence than do more internal, subjective traits (Albright, Kenny, &

Table 8 *Correlations Between the General Positive Affect and Negative Affect Scales and Regression-Based Scores on the First Two Varimax Factors Emerging in Within-Subject Data*

Time Instructions	No. of Subjects	Total No. of Observations	Positive Affect Scale Correlations		Negative Affect Scale Correlations	
			Factor 1	Factor 2	Factor 1	Factor 2
Moment	226	10,169	.93	-.20	-.16	.89
Today 248	11,322	.90	-.23	-.18	.89	

Malloy, 1988; Funder & Colvin, 1988; Funder & Dobroth, 1987; Kenrick & Stringfield, 1980; Norman & Goldberg, 1966; Watson, 1989; Watson & Clark, 1991). Because affective experience is strongly subjective, one cannot expect the convergent correlations to be very high. Furthermore, earlier research has also found that self-peer convergence increases as more peer judges are used (McCrae & Costa, 1987; Watson, 1989; Watson & Clark, 1991). Accordingly, one would predict that convergence would be better in the dormitory study (in which each target was rated by 3-4 well-acquainted peers) than in the dating study (in which each target was rated by only a single peer).

The results of both studies are presented in Table 9. These data clearly support the convergent and discriminant validity of the general Positive Affect and Negative Affect scales. The convergent correlations all are statistically significant; those for Positive Affect are especially good, with values of .48 and .35 in the dormitory and dating studies, respectively. The convergent correlations for Negative Affect (.36 and .21, respectively) are lower, but still respectable given the limitations discussed earlier. Finally, it is noteworthy that the discriminant coefficients are invariably low and non-significant.

In summary, the two general PANAS-X scales provide reliable, valid, and largely independent measures of the higher order Positive Affect and Negative Affect dimensions, regardless of the subject population or time frame used. For more information regarding the reliability and validity of these scales, see Watson (1988b) and Watson et al. (1988).

III. The Specific Affect Scales

A. Construction of the Fear, Sadness, Guilt, Hostility, Shyness, Fatigue, and Surprise Scales

In constructing the lower order scales, we continued to be concerned with developing reliable and valid measures that could be used with different time instructions. Therefore, the initial scale construction was based on the six large data sets (with *N*s ranging from 586 to 1,002) that were used in the first series of analyses on factorial validity (see Table 6 and the accompanying discussion). As noted previously, each of these data sets included Tellegen's list of 60 mood terms (see Zevon & Tellegen, 1982), differing only in the time frames used to generate the mood ratings (Moment, Today, Past Few Days, Past Few Weeks, Past Year, General).

Each of these data sets was factored separately using principal factor analysis (squared multiple correlations in the diagonal) and varimax rotation. To determine the final number of factors in each data set, a range of solutions--starting at two factors--was examined, until a solution was reached that contained an uninterpretable factor (i.e., one with fewer than three marker terms). In these analyses, we defined a marker as a variable that loaded |.40| or higher on the factor and had its highest loading on the factor. These criteria yielded the following numbers of factors: Moment (7), Today (8), Past Few Days (8), Past Few Weeks (8), Past Year (10), and General (6). In each case, this represents the *maximum* number of interpretable factors that could be identified in the data.

Table 9 *Convergence Between Self- and Peer-ratings on the Two Higher Order* PANAS-X *Scales*

Peer Rated	Self Rated PA	Self Rated NA
Dormitory Study (N = 89)		
Positive Affect	.48*	-.05
Negative Affect	-.13	.36*
Dating Study (N = 137)		
Positive Affect	.35*	-.09
Negative Affect	-.03	.21*

*$p < .05$, two-tailed.

Table 10 *Varimax-Rotated Factor Loadings of the 60 Tellegen Mood Terms in the Past Few Weeks Solution (N=586)*

	Factor							
Descriptor	1	2	3	4	5	6	7	8
enthusiastic	.74							
happy	.70							
joyful	.70							
interested	.69							
determined	.66							
delighted	.66							
confident	.65							
excited	.65							
friendly	.63							
warmhearted	.62							
alert	.61							
proud	.59							
strong	.59							
active	.58							
sociable	.57							
inspired	.56							
content	.56							
at ease	.55		-.40					
attentive	.49							
healthy	.46							
calm	.45		-.39					
** hostile		.69						
** scornful		.68						
disdainful		.63						
** loathing		.62						
** angry		.61						
contemptuous		.59						
** disgusted		.56			.35			
** irritable		.55						
revulsion		.53						
tormented		.39	.36					
** scared			.76					
** afraid			.71					
** frightened			.70					
** shaky			.57					
** nervous			.56					
** jittery			.52					
distressed		.35	.42	.33	.33			

(table continues)

Table 10 (cont.)

	Factor							
Descriptor	1	2	3	4	5	6	7	8
** alone				.75				
** lonely				.71				
** sad				.62				
** blue		.31		.61				
** downhearted				.55				
rejected				.54				
upset		.39	.32	.49				
** dissatisfied with self					.74			
** angry at self					.73			
** disgusted with self					.68			
** guilty					.54			
** blameworthy					.51			
** ashamed					.51			
** sleepy						.76		
** tired						.74		
** sluggish						.57		
** astonished							.62	
** amazed							.62	
** surprised							.60	
** bashful								.67
** shy								.66
** sheepish								.44

Note. Loadings below |.30| are omitted.

**Item was included as a marker in the preliminary version of the scale assessing this factor.

Overall, these multifactorial solutions were remarkably clear and clean. Most terms defined one and only one factor, and there were few significant secondary loadings. Moreover, the factors that emerged were both easily interpretable and broadly consistent with existing theoretical models of emotion (e.g., Ekman, 1971; Izard, 1972). A representative factor loading matrix (the Past Few Weeks solution) is presented in Table 10. Eight factors emerged clearly in at least four of the six solutions. Seven of these dimensions--Fear, Sadness, Guilt, Hostility, Shyness, Fatigue, and Surprise-- represented specific emotional states and therefore were used as the basis for scale development. The terms that most strongly and consistently defined these factors across all of the solutions were selected as the component scale items, and are indicated in Table 10. The eighth replicable factor was the higher order Positive Affect dimension; we describe our further efforts to identify specific positive emotional factors in the next section.

Initial internal consistency reliabilities (Cronbach's coefficient alpha) were computed for all seven scales in the six scale-development data sets. Consistent with their factor-analytic derivation, the scales were all reasonably homogeneous. Moreover, the reliability of the scales did not vary systematically with the rated time frame. Not surprisingly, however, the shorter (3-item) scales were somewhat less reliable than those with five or six items (Fear, Hostility, Guilt, and Sadness). We therefore decided that they could be improved by adding a fourth term. Using data from the samples to be described in the next section, we tested an additional marker for Shyness (*timid*) and Fatigue (*drowsy*), and found that they significantly improved the psychometric properties of their respective scales. These terms were therefore added to the Shyness and Fatigue scales, bringing them to their final 4-item form. Although the psychometric properties of Surprise also might have been improved through the addition of a fourth term, no further revisions were made on this scale because we were unable to find a commonly used term that did not greatly affect its valence or correlate strongly with other lower order scales. We will present psychometric and normative information on these finalized scales after we have described the development of the remaining specific affect scales.

B. Construction of the Joviality, Self-Assurance, Attentiveness, and Serenity Scales

The one unexpected aspect of these multifactorial solutions was the failure of Positive Affect to split into more specific positive emotional states. Separate positive affect factors emerged only in the Past Year solution. In these data, Positive Affect split into Joy/Sociability and Interest/Energy. To test whether these specific positive emotional factors should be measured separately, we constructed scales using the marker variables from the two factors in the Past Year solution. Across the six data sets, these two scales correlated from .66 to .77 with one another (median r = .73). Because of this strong level of intercorrelation, we did not pursue these two scales further. More generally, positive mood descriptors appear to be even more highly interrelated than are negative terms, and we therefore believe that positive emotional states may prove ultimately to be less differentiable in self-report.

Nevertheless, we continued to seek an alternative method for developing scales to assess specific positive affects. To achieve a desirable level of differentiation, we needed a broader sample of positive mood descriptors. We therefore created an expanded pool of 36 positive mood terms. This list included 20 of the previous set of 21 terms (*healthy* was dropped because preliminary analyses showed that it consistently split across several factors), plus additional terms assessing venturesomeness, energy, cheerfulness, sociability, mental alertness, and serenity.
This expanded set of positive mood descriptors was subjected to principal components analyses (with varimax rotation) in three new student samples. First, 607 subjects rated themselves using Past Week instructions; 347 of these subjects were then retested (using the same instructions) 2 months later. Finally, 327 individuals rated themselves using General instructions. Each of these solutions indicated the presence of a large general factor, which accounted for 30.3% to 37.5% of the total variance. Nevertheless, four interpretable factors could be identified in each solution. Terms that did not load clearly or consistently on one of these factors were eliminated gradually from subsequent analyses, leading to a reduced set of 21 clear marker terms. Analyses on this reduced set of descriptors yielded a clean simple-structure rotation in each solution. A representative solution (the initial Past Week sample) is shown in Table 11.

Four specific positive affect scales--Joviality, Self-Assurance, Attentiveness, and Serenity--were constructed from the clearest and most consistent markers of each of these factors. The descriptors comprising these PANAS-X scales are presented in Table 2.

C. Normative and Internal Consistency Data

Basic between-subjects data. Table 12 reports basic descriptive statistics and internal consistency reliabilities (Cronbach's coefficient a) for the lower order PANAS-X scales in 11 data sets reflecting eight different time instructions. Looking first at the descriptive statistics, it should be noted that these data show a pattern similar to that observed with the higher order scales (see Table 3); that is, mean scale scores show a general tendency to increase as the rated time frame lengthens.

Table 12 also clearly demonstrates that the alpha reliabilities of the longer (i.e., 5-8 items) PANAS-X scales are high. Joviality is both the longest and the most reliable of the lower order scales, with a median internal consistency estimate of .93 (range = .88 to .94) across the 11 samples. Furthermore, Fear (median a = .87), Sadness (median a = .87), Guilt (median a = .88), Hostility (median a = .85), Fatigue (median a = .88), Self-Assurance (median a = .83) and Shyness (median a = .83) also consistently show good reliabilities. In contrast, three of the shorter (i.e., 3-4 item) scales--Attentiveness, Serenity, and Surprise--yielded slightly lower reliability estimates; across the 11 samples,

Table 11 *Varimax-Rotated Factor Loadings of the Positive Mood Terms Defining the Joviality, Self-Assurance, Attentiveness, and Serenity Factors (Past Week Instructions, N = 607)*

	Factor			
Descriptor	1	2	3	4
happy	.79			
cheerful	.78			
joyful	.78			
excited	.77			
enthusiastic	.73			
lively	.70	.30		
energetic	.64		.31	
delighted	.63			
bold		.71		
fearless		.69		
strong		.68		
proud		.64		
confident		.60		.34
daring	.42	.57		
concentrating			.79	
attentive			.77	
determined			.66	
alert			.61	
calm				.78
relaxed				.76
at ease	.37			.62

Note. Loadings below |.30| are omitted.

Table 12 *Descriptive Statistics and Internal Consistency Reliabilities for the 11 Lower Order* PANAS-X *Scales*

Scale	*M*	*SD*	Coefficient α
Moment (1,027 SMU undergraduates)			
Fear	9.9	4.5	.87
Sadness	9.4	4.4	.86
Guilt	8.7	4.0	.86
Hostility	9.3	3.9	.82
Shyness	6.7	2.9	.80
Fatigue	11.1	4.2	.88
Joviality	21.7	7.5	.93
Self-Assurance	16.5	5.1	.83
Attentiveness	12.5	3.1	.72
Serenity	9.8	2.8	.74
Surprise	4.8	2.4	.80
Today (1,007 SMU undergraduates)			
Fear	11.0	4.9	.87
Sadness	9.8	4.7	.87
Guilt	10.1	5.2	.89
Hostility	10.7	4.9	.86
Shyness	7.5	3.4	.83
Fatigue	11.5	4.4	.89
Joviality	23.9	7.8	.94
Self-Assurance	16.0	5.2	.84
Attentiveness	13.1	3.3	.79
Serenity	9.4	2.7	.76
Surprise	5.9	2.7	.75
Past Few Days (289 SMU undergraduates)			
Fear	11.2	5.2	.90
Sadness	9.9	4.8	.89
Guilt	10.4	5.2	.90
Hostility	11.7	5.5	.89
Shyness	7.0	3.0	.85
Fatigue	11.5	4.1	.89
Joviality	26.5	7.0	.93
Self-Assurance	17.3	5.1	.84
Attentiveness	13.5	3.1	.79
Serenity	9.1	2.6	.79
Surprise	6.7	2.9	.79
Past Few Days (502 Iowa undergraduates)			
Fear	12.9	5.0	.86
Sadness	11.5	4.6	.86
Guilt	11.8	5.3	.87
Hostility	12.4	4.7	.83
Shyness	8.2	3.2	.80
Fatigue	12.6	3.8	.86
Joviality	25.7	6.9	.93
Self-Assurance	17.2	5.0	.83
Attentiveness	13.0	3.1	.75
Serenity	9.2	2.8	.79
Surprise	6.8	2.7	.78
Past Week (1,278 SMU undergraduates)			
Fear	12.2	5.2	.88
Sadness	11.2	4.9	.89
Guilt	10.8	4.9	.87
Hostility	11.6	4.6	.85
Shyness	8.1	3.4	.83
Fatigue	11.7	3.8	.87
Joviality	26.5	6.9	.93
Self-Assurance	17.1	4.8	.83
Attentiveness	13.1	3.1	.79
Serenity	8.9	2.5	.74
Surprise	6.6	2.7	.80
Past Week (328 Dallas-area adults)			
Fear	9.6	4.1	.88
Sadness	9.5	4.4	.88
Guilt	10.1	4.9	.91
Hostility	11.1	4.6	.88
Shyness	6.3	2.7	.83
Fatigue	9.0	3.8	.89
Joviality	25.0	6.8	.93
Self-Assurance	16.5	4.6	.81
Attentiveness	12.7	3.0	.78
Serenity	8.7	2.4	.75
Surprise	5.4	2.3	.75
Past Few Weeks (678 SMU undergraduates)			
Fear	12.3	4.9	.86
Sadness	11.7	4.8	.87
Guilt	12.0	5.2	.86
Hostility	12.9	5.0	.85
Shyness	7.7	3.1	.81
Fatigue	12.7	3.9	.88
Joviality	26.8	6.6	.93
Self-Assurance	17.7	4.7	.81
Attentiveness	13.5	2.9	.75
Serenity	8.9	2.6	.79
Surprise	6.8	2.8	.80
Past Month (1,006 SMU undergraduates)			
Fear	11.3	4.6	.86
Sadness	10.6	4.6	.88
Guilt	11.0	5.2	.89
Hostility	12.3	4.8	.86
Shyness	7.7	3.2	.82
Fatigue	11.3	3.7	.87
Joviality	27.7	6.4	.92
Self-Assurance	18.5	4.8	.83
Attentiveness	13.8	3.0	.78
Serenity	9.4	2.6	.79
Surprise	6.9	2.7	.77

Table 12 *(cont.)*

Scale	*M*	*SD*	Coefficient α
Past Year (315 SMU undergraduates)			
Fear	13.7	4.6	.84
Sadness	12.9	4.7	.86
Guilt	13.1	5.5	.88
Hostility	14.3	4.7	.83
Shyness	8.7	3.3	.84
Fatigue	12.2	3.6	.84
Joviality	28.0	5.8	.91
Self-Assurance	19.0	4.4	.80
Attentiveness	14.0	2.8	.78
Serenity	9.2	2.5	.74
Surprise	7.5	2.3	.74
General (1,657 SMU undergraduates)			
Fear	11.3	3.8	.83
Sadness	10.1	3.7	.83
Guilt	10.8	4.3	.85
Hostility	11.5	4.0	.83
Shyness	8.3	3.3	.83
Fatigue	10.3	3.4	.86
Joviality	28.1	6.0	.93
Self-Assurance	19.1	4.3	.81
Attentiveness	14.2	2.6	.76
Serenity	9.8	2.3	.73
Surprise	6.7	2.3	.76
General (107 mixed inpatients/outpatients)			
Fear	15.1	9.0	.92
Sadness	14.8	5.3	.88
Guilt	17.7	6.6	.90
Hostility	14.4	4.8	.79
Shyness	9.9	4.3	.86
Fatigue	10.5	4.2	.89
Joviality	21.3	6.1	.88
Self-Assurance	15.8	4.9	.80
Attentiveness	13.1	3.0	.70
Serenity	7.9	2.8	.83
Surprise	6.4	2.5	.72

these scales had median internal consistency values of .78, .76, and .77, respectively. Nevertheless, these median values reflect mean inter-item correlations of .45 or greater, indicating that the scales are appropriately homogeneous; thus, their reliability estimates simply reflect the fact that they have relatively few items. As mentioned earlier with regard to Surprise, the data suggest that these scales could be improved through the inclusion of additional marker terms. Unfortunately, the English language contains few suitable terms in these content domains and we have not been able to identify additional markers for these scales.

Analyses of gender. An issue of general interest concerns the possibility of systematic gender differences in affective experience. We examined this issue in the 10 large data sets (with *N*s ranging from 289 to 1,657) that were used in the second series of analyses on factorial validity (see Table 7 and the accompanying discussion). As was discussed earlier, subjects in these samples were assessed on the full set of 60 PANAS-X terms. Accordingly, we were able to conduct *t*-tests comparing men and women on all 13 PANAS-X scales in each sample. Note that all of these data sets contained at least 100 subjects of each gender (the gender breakdowns of these samples are reported in Table 13), and most of them contained many more than that. Given the large size of these samples, these *t*-tests will identify significant effects even when the absolute difference between groups is quite small.

Generally speaking, these analyses revealed few consistent gender-related differences in affective experience. Five scales--general Negative Affect, general Positive Affect, Attentiveness, Surprise and Sadness--showed virtually no significant gender-related effects. General Positive Affect and Attentiveness both yielded significant but inconsistent group differences in two samples: In each case men scored higher in one sample and women scored higher in the other. Sadness (women scored higher in the Past Week/Adult sample) and Surprise (men scored higher in the Past Month sample) each produced one significant group difference. Finally, the general Negative Affect scale showed no significant effects whatsoever.

Five additional scales displayed modest gender-related differences. Specifically, women scored significantly higher on Joviality in four samples (Today, Past Week/SMU, Past Month, General), on Fatigue in three samples (Moment, Past Month, Past Week/Adult), and on Fear in two samples (Past Few Days/Iowa, Past Year). Conversely, in four samples men scored significantly higher on Guilt (Today, Past Few Days/SMU, Past Week/SMU, General) and Shyness (Past Few Days/SMU, Past Few Weeks, Past Month, General).

The three remaining scales showed more consistent gender effects. Specifically, men reported significantly higher levels of Self-Assurance in all 10 samples; in addition, they scored significantly higher on Serenity in nine samples (the sole exception being the Today data set) and on Hostility in seven samples (Moment, Today, Past Few Days/SMU, Past Week/SMU, Past Few Weeks, Past Month, General). Accordingly, Table 13 reports gender-specific normative data on these three scales in all 10 data sets. Note that even though these scales yielded consistent gender differences, the absolute size of the difference generally is quite small. That is, across the 10 data sets the mean scores for men are approximately two points higher on Self-Assurance, one point higher on Hostility, and a half-point higher on Serenity. Thus, it appears that the affective experiences of women and men generally are quite similar, so that the overall normative statistics presented in Tables 3 and 12 can be used with some confidence.

Table 13 *Gender-Specific Normative Statistics for the Self-Assurance, Hostility, and Serenity Scales*

	Men		Women	
Time Frame/Sample	*M*	*SD*	*M*	*SD*
	Self-Assurance			
Moment[ab]	17.8	5.0	15.6	5.0
Today[ab]	16.9	5.3	15.3	5.0
Past Few Days[ab]	18.8	4.7	16.4	5.2
Past Few Days[ac]	17.7	4.6	16.6	5.3
Past Week[ab]	18.0	4.8	16.5	4.7
Past Week[ad]	17.6	4.4	15.5	4.7
Past Few Weeks[ab]	19.1	4.4	16.8	4.6
Past Month[ab]	19.6	4.8	17.7	4.6
Past Year[ab]	20.3	4.2	18.0	4.2
General[ab]	19.8	4.1	18.6	4.4
	Hostility			
Moment[ab]	10.0	4.1	8.8	3.6
Today[ab]	11.4	5.1	10.2	4.7
Past Few Days[ab]	12.7	5.2	11.0	5.4
Past Few Days[c]	12.6	4.5	12.2	4.9
Past Week[ab]	12.4	4.8	11.0	4.4
Past Week[d]	10.9	4.4	11.3	4.8
Past Few Weeks[ab]	13.7	4.7	12.4	5.2
Past Month[ab]	12.7	4.8	11.9	4.6
Past Year[ab]	15.0	5.3	14.0	4.3
General[ab]	12.4	4.3	10.8	3.6
	Serenity			
Moment[ab]	10.0	2.7	9.6	2.8
Today[b]	9.4	2.7	9.3	2.7
Past Few Days[ab]	9.6	2.5	8.9	2.6
Past Few Days[ac]	9.5	2.7	8.7	2.9
Past Week[ab]	9.1	2.5	8.7	2.5
Past Week[ad]	9.1	2.4	8.5	2.5
Past Few Weeks[ab]	9.3	2.6	8.7	2.6
Past Month[ab]	9.8	2.6	9.2	2.6
Past Year[ab]	9.6	2.7	9.0	2.4
General[ab]	10.0	2.3	9.6	2.2

Note. *Ns* by gender: Moment (437 M, 585 W), Today (420 M, 583 W), Past Few Days/SMU (102 M, 182 W), Past Few Days/Iowa (281 M, 221 W), Past Week/SMU (502 M, 769 W), Past Week/Adult (142 M, 186 W), Past Few Weeks (259 M, 409 W), Past Month (391 M, 602 W), Past Year (128 M, 183 W), General (660 M, 989 W).

[a]Means across gender differed ($p < .05$, two-tailed). [b]SMU undergraduates. [c]Iowa undergraduates. [d]Dallas-area adults.

Within-subject data. As with the higher order scales, some investigators may wish to use the specific affect scales in designs that necessitate repeated within-subject assessments. Accordingly, Table 14 presents basic descriptive statistics from a sample of 262 SMU undergraduates who completed the full PANAS-X (using Today instructions) on a daily basis over a period of several weeks. Paralleling the data reported in Table 5, all subjects completed a minimum of 30 daily mood ratings ($M = 45.0$ observations per subject); all of the assessments were completed in the evening, so that the ratings would provide a reasonable estimate of the subjects' moods over the course of the day. As with the Table 5 data, mean scale scores initially were calculated for each subject; overall sample means and standard deviations were then computed for each scale. It is noteworthy that--paralleling the pattern observed with the higher order scales--the means for all 11 scales are lower than the corresponding single-assessment between-subjects values (using Today instructions) that are shown in Table 3. These results strongly suggest that subjects tend to give slightly lower mood ratings with repeated measurement.

D. Construct Validity

Comparison analyses with the POMS scales. To demonstrate the convergent and discriminant validity of a subset of the PANAS-X scales, we created an affect questionnaire that included the descriptors for Fear, Hostility, Sadness, Fatigue, general Positive Affect, and the six scales from the Profile of Mood States (POMS; McNair, Lorr, & Droppleman, 1971). This questionnaire was administered to 563 SMU undergraduates using Past Few Weeks instructions.

These five PANAS-X scales measure the same basic affects as corresponding scales from the POMS. Specifically, Fear can be identified with POMS Tension-Anxiety, Hostility with POMS Anger-Hostility, Sadness with POMS Depression-Dejection, Fatigue with POMS Fatigue, and general Positive Affect with POMS Vigor (which, despite its name, includes a broad sample of positive mood terms). Intercorrelations among these conceptually overlapping scales are presented in Table 15. The table shows that each of the PANAS-X scales is strongly related to its POMS counterpart, with convergent correlations ranging from .85 to .91. These very high coefficients partly reflect item overlap, as the corresponding scales have one to three items in common. Nevertheless, the important point is that the PANAS-X scales assess affective states that are broadly similar to those measured in existing multi-affect inventories such as the POMS.

Table 15 also demonstrates that the PANAS-X scales offer an important advantage over their POMS counterparts: They tend to be less highly correlated with one another, and thus show better discriminant validity. The mean correlation among the PANAS-X Fear, Hostility, Sadness and Fatigue scales was .45, which was significantly lower than the mean correlation (.60) among the corresponding POMS

Table 14 *Descriptive Statistics for the 11 Lower Order Scales Derived from Mean Daily Mood Scores*

Scale	*M*	*SD*
Fear	9.7	2.7
Sadness	8.2	2.4
Guilt	9.0	2.8
Hostility	9.4	2.3
Shyness	5.5	1.5
Fatigue	8.6	2.2
Joviality	22.7	4.8
Self-Assurance	14.7	3.5
Attentiveness	11.3	2.3
Serenity	8.3	1.8
Surprise	5.5	1.8

Note. $N = 262$. Total number of observations = 11,783. See text for details.

scales ($p < .01$, two-tailed; Fisher's r to z transformation was used in the computation of these and subsequent mean correlations). Follow-up comparisons indicated that five of the six individual correlations were also significantly lower ($p < .01$, two-tailed) in the PANAS-X scales; specifically, only the Fear-Hostility correlation did not differ significantly between the two instruments. Thus, the PANAS-X scales generally provide a less redundant, more differentiated assessment of affect.

Convergence with peer ratings. To document the construct validity of the lower order PANAS-X scales further, we again will consider evidence from two studies in which self-ratings on these scales were correlated with corresponding judgments made by well-acquainted peers. First, in the dormitory study that was discussed previously (see Table 9 and the accompanying discussion), self- and aggregated peer-ratings (averaged across the responses from three or four well-acquainted peers) were available on seven of the lower order scales: Fear, Hostility, Guilt, Sadness, Shyness, Fatigue, and Surprise (note, however, that preliminary 3-item versions of Shyness and Fatigue were used). The heteromethod correlations from this sample are presented in Table 16. These data produced strong self-peer convergence, and generally support the convergent and discriminant validity of the PANAS-X scales. The single exception is Surprise, which produced a nonsignificant level of self-peer agreement ($r = .14$) in these ratings. The six remaining scales, however, yielded convergent correlations ranging from .27 to .52, with a mean value of .38. In contrast, the discriminant coefficients were generally low and nonsignificant. Using the criterion that the convergent correlation should be higher than any of the other values in its row or column of the heteromethod block (Campbell & Fiske, 1959), five of the scales (Fear, Hostility, Sadness, Shyness, and Fatigue) showed acceptable discriminant validity. (For a more detailed discussion of these data, see Watson & Clark, 1991).

Table 15 *Correlations among the* PANAS-X *Scales and Corresponding Scales from the Profile of Mood States (*POMS*; McNair, Lorr & Droppleman, 1971)*

	PANAS-X Scales					POMS Scales			
Instrument/Scale	1	2	3	4	5	6	7	8	9
PANAS-X Scales									
1. Fear	---								
2. Hostility	.58	---							
3. Sadness	.61	.49	---						
4. Fatigue[a]	.40	.31	.27	---					
5. Positive Affect	.02	-.06	-.25	-.07	---				
POMS Scales									
6. Tension-Anxiety	**.85**	.62	.57	.48	.02	--			
7. Anger-Hostility	.59	**.91**	.51	.35	.00	.63	--		
8. Depression-Dejection	.74	.66	**.85**	.34	-.24	.69	.66	--	
9. Fatigue	.53	.46	.40	**.89**	-.07	.61	.48	.47	--
10. Vigor	-.03	-.09	-.28	-.08	**.86**	-.05	-.03	-.25	-.07

Note. $N = 563$. Mood ratings based on Past Few Weeks instructions. Convergent correlations are shown in boldface. All correlations greater than $|.10|$ are significant at $p < .01$, two-tailed.

[a]Data are based on a preliminary, 3-item version of this scale.

Table 16 *Self- versus Peer-Ratings on Seven Lower Order* PANAS-X *Scales (Dormitory Study)*

	Self Rated						
Peer Rated	Fear	Hostility	Guilt	Sadness	Shyness	Fatigue	Surprise
Fear	**.40***	-.07	.16	.25*	.25*	.00	.10
Hostility	.08	**.31***	.10	.22*	-.08	-.04	-.12
Guilt	.34*	.04	**.34***	.39*	.20	.09	.16
Sadness	.31*	-.03	.30*	**.52***	.27*	-.01	.13
Shyness[a]	.23*	-.18	.09	.03	**.42***	.01	-.02
Fatigue[a]	-.02	.12	-.10	-.02	.08	**.27***	-.04
Surprise	.03	-.19	-.02	.02	.02	-.05	**.14**

Note. $N = 89$. Convergent correlations are shown in boldface. Correlations of .30 or greater are underlined.
[a]Data are based on preliminary, 3-item versions of these scales. *$p < .05$, two-tailed.

Second, self- and peer-ratings were available for all 11 lower order scales from the dating study described earlier. The heteromethod correlations from this sample are shown in Table 17. Similar to the results presented in Table 16, these data yielded strong self-peer convergence and broadly support the convergent and discriminant validity of the lower order PANAS-X scales. Eight of the 11 scales (Joviality, Self-Assurance, Attentiveness, Serenity, Sadness, Hostility, Shyness and Fatigue) had convergent correlations of .30 or greater; in fact, the mean coefficient across these scales was .36. Moreover, using the same criterion described in connection with Table 16, all eight of these scales displayed acceptable discriminant validity. Thus, the data from this study clearly support the construct validity of these eight scales. A ninth scale (Fear) also had a significant convergent correlation (.21), but failed to demonstrate acceptable discriminant validity. Finally, the two remaining lower order scales (Surprise and Guilt) did not yield significant evidence of either convergent or discriminant validity. Taken together with the results of the dormitory study, these data seriously challenge the validity of trait ratings on Surprise.

Table 17 *Self- versus Peer-Ratings on the 11 Lower Order* PANAS-X *Scales (Dating Study)*

	Self Rated										
Peer Rated	Jov	Assur	Atten	Seren	Fear	Sad	Guilt	Host	Shy	Fat	Surp
Joviality	**.36***	.28*	.15	.13	-.02	-.26*	.00	-.16	-.29*	-.03	.20*
Self-Assurance	.17*	**.44***	.25*	.09	-.06	-.17*	.02	-.11	-.27*	-.03	.09
Attentiveness	.17*	.24*	**.34***	.02	-.02	-.13	-.05	-.10	-.25*	-.10	.04
Serenity	.09	.17*	.16	**.37***	-.04	-.14	-.08	-.15	.10	.04	.07
Fear	.00	-.17*	-.06	-.25*	**.21***	.15	.10	.20*	.15	.14	.06
Sadness	-.02	-.19*	.07	-.23*	.13	**.30***	.13	.20*	.19*	-.03	-.03
Guilt	-.07	-.09	.00	-.20*	.04	.14	**.15**	.19*	.10	.07	-.07
Hostility	-.11	-.02	.02	-.22*	-.05	.11	.04	**.32***	.00	.01	-.09
Shyness	-.06	-.20*	-.09	-.04	.13	.03	.05	.06	**.37***	.09	.08
Fatigue	.01	-.01	.08	-.06	.05	.02	-.01	.10	.27*	**.37***	-.02
Surprise	.16	.11	.16	-.04	.01	-.14	-.08	-.10	-.02	.02	**.14**

Note. $N = 137$. Convergent *r*s in boldface; *r*s of $\geq .30$ underlined. Jov = Joviality; Assur = Self-Assurance; Atten = Attentiveness; Seren = Serenity; Sad = Sadness; Host = Hostility; Shy = Shyness; Fat = Fatigue; Surp = Surprise.
*$p < .05$, two-tailed.

IV. General Issues Concerning the PANAS-X Scales

A. Testing the Hierarchical Arrangement of the PANAS-X *Scales*

As was discussed previously, the PANAS-X was designed explicitly to reflect the hierarchical structure of self-rated affect. That is, self-rated affect is characterized by two broad higher order dimensions (Negative Affect and Positive Affect), each of which is composed of several correlated, yet ultimately distinguishable specific affect states. The PANAS-X includes scales assessing both of these structural levels--that is, the two higher order dimensions and 11 specific lower order states.

How accurately does the PANAS-X capture this hierarchical structure? To answer this question, we subjected the 11 lower order PANAS-X scales to second-order principal components analyses in two large samples of SMU undergraduates, who rated themselves using Moment (N = 1,027) and General (N = 1,657) instructions. Two large second-order factors emerged in each solution: These two dimensions jointly accounted for 52.8% and 54.9% of the total variance in the Moment and General data, respectively.

Two components were therefore extracted in each solution and were rotated using varimax; the rotated loadings are presented in Table 18. In each solution, the first factor clearly can be identified as general Negative Affect. Fear, Sadness, Guilt, Hostility and Shyness are all very strong markers of the high end of this second-order factor, with loadings ranging from .59 to .83 across the two solutions. Fatigue and Surprise also have moderate positive loadings (ranging from .33 to .53) on this dimension. Finally, Serenity serves to define the low end of the factor, with loadings of -.50 and -.26 in the Moment and General solutions, respectively.

The second higher order dimension clearly corresponds to the general Positive Affect dimension that consistently emerges in self-report studies. It is most strongly and clearly defined on its high end by Joviality, Self-Assurance, and Attentiveness, which have loadings ranging from .70 to .86 across the two solutions. Surprise (loadings of .51 and .62 in the Moment and General solutions, respectively) and Serenity (.26 and .46, respectively) also are reasonably good markers of high Positive Affect. Finally, Fatigue is the best marker of low Positive Affect, with loadings of -.39 and -.25 in the Moment and General solutions, respectively.

How closely do these higher order dimensions correspond to the general PANAS-X scales? To examine this issue, we computed regression-based factor scores for the second-order dimensions in each solution. Within each data set, we then correlated these computed factor scores with the general Negative Affect and Positive Affect scales. These results are displayed in Table 19, and they again confirm the factorial validity of the higher order scales. Specifically, scores on Factor 1 correlated .92 and .91 with general Negative Affect in the Moment and General solutions, respectively; conversely, scores on Factor 2 correlated .94 and .91 with general Positive Affect in the Moment and General solutions, respectively. Note also that the discriminant correlations were invariably low, ranging only from -.01 to -.11.

Table 18 *Varimax-Rotated Loadings of the 11 Lower Order PANAS-X Scales with Two Different Time Instructions*

	Factor 1		Factor 2	
Scale	Moment	General	Moment	General
Fear	**.79**	**.83**	.05	.07
Sadness	**.77**	**.79**	-.27	-.21
Hostility	**.78**	**.77**	-.07	-.01
Guilt	**.69**	**.78**	-.09	-.15
Shyness	**.59**	**.62**	.04	-.09
Fatigue	**.33**	**.53**	**-.39**	-.25
Joviality	-.14	-.13	**.86**	**.85**
Self-Assurance	-.04	-.11	**.80**	**.78**
Attentiveness	-.06	-.13	**.76**	**.70**
Surprise	**.49**	**.42**	**.51**	**.62**
Serenity	**-.50**	-.26	.26	**.46**

Note. *N*s = 1,027 (Moment) and 1,657 (General). Loadings of |.30| or greater are shown in boldface.

These data clearly demonstrate the hierarchical arrangement of the PANAS-X scales. Furthermore, on the basis of these and other analyses, we have grouped the 11 lower order scales into three broad subcategories (see Table 2). First, Fear, Sadness, Guilt, and Hostility scales are classified as *Basic Negative Emotion Scales*. As the data in Table 18 illustrate, these scales are consistently and substantially intercorrelated, and therefore, are strong and clear markers of the higher order Negative Affect dimension (see also Watson & Clark, 1989, 1992a). Second, Joviality, Self-Assurance and Attentiveness are classified as *Basic Positive Emotion Scales*. As can be seen in Table 18, these scales are highly correlated with one another, and so are strong and consistent markers of the second-order Positive Affect dimension.

Finally, Shyness, Fatigue, Surprise and Serenity are grouped as *Other Affective States* because they do not strongly or consistently define either of the second-order factors. Shyness tends to load moderately to strongly on Negative Affect (see Table 18), but its loading generally is somewhat lower than those of Fear, Sadness, Guilt, and Hostility. That is, Shyness appears to be less strongly saturated with general Negative Affect variance than are these other scales (see also Watson & Clark, 1989). In contrast, Surprise typically has moderate positive loadings on both higher order factors. Fatigue and Serenity also tend to load

significantly on both general factors: Fatigue is a marker of high Negative, low Positive Affect, whereas Serenity is a marker of low Negative, high Positive Affect. However, these loadings vary considerably across different samples and time frames.

Table 19 *Correlations between the General Positive Affect and Negative Affect Scales and Regression-Based Scores on the First Two Varimax Factors Defined by the 11 Lower Order* PANAS-X *Scales*

	Correlations with	
Scale	Factor 1	Factor 2
Moment Instructions (N = 1,027)		
Positive Affect	-.03	.94
Negative Affect	.92	-.05
General Instructions (N=1,657)		
Positive Affect	-.11	.91
Negative Affect	.91	-.01

B. Using the PANAS-X *Scales as State Measures*

We have presented various types of evidence that establish the convergent and discriminant validity of the PANAS-X scales. In these final sections we examine more specific validation issues, namely, using the PANAS-X scales as measures of state and trait affect, respectively.

Relations with other measures of mood and symptomatology. Many researchers will be interested in using the PANAS-X scales as measures of *state* affect--that is, to assess relatively short-term fluctuations in mood. What evidence suggests that the PANAS-X can be used validly in this way? First, the PANAS-X scales are strongly correlated with other existing measures of shorter term affect. We have already seen, for example, that five of the PANAS-X scales are very highly correlated (convergent correlations ranged from .85 to .91) with corresponding measures from the POMS, which typically uses short-term instructions (in this case, Past Few Weeks). Moreover, the PANAS-X scales showed better discriminant validity--that is, they were less highly intercorrelated than were their POMS counterparts (see Table 15).

Watson and Clark (1992a, Study 1) report additional convergent and discriminant validity data for the Fear, Sadness, and Hostility scales in a sample of 195 SMU undergraduates (see their Tables 1 and 2). For instance, scores on the PANAS-X Sadness scale (assessed using Past Few Weeks time instructions) correlated .59 with the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), .69 with the Depression scale from the Hopkins Symptom Checklist (HSCL; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974). and .75 with the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977). Similarly, the PANAS-X Fear scale correlated .74 with the HSCL Anxiety scale and .56 with the STAI State Anxiety Scale (Spielberger, Gorsuch, & Lushene, 1970). Finally, the PANAS-X Hostility scale correlated .55 with the Hostility scale from the SCL-90 (Derogatis, 1977; Derogatis & Cleary, 1977) and .45 with the state form of the State-Trait Anger Scale (STAS; Spielberger, Jacobs, Russell, & Crane, 1983). Furthermore, it is important to note that all three PANAS-X scales showed evidence of significant discriminant validity (see Watson & Clark, 1992a, Table 2).

Similarly, Watson et al. (1988) present correlations between the two higher order PANAS-X scales and other commonly used measures of state affect and psychiatric symptomatology (see Watson et al., 1988, Table 7). For example, in a sample of 880 SMU undergraduates, the BDI correlated .56 and -.35 with general Negative Affect and Positive Affect, respectively (the PANAS-X ratings were based on Past Few Days instructions). Similarly, in another sample of 208 SMU students, the BDI correlated .58 and -.36 with Negative Affect and Positive Affect, respectively (Past Few Weeks instructions). In this latter sample (N = 203), the STAI State Anxiety Scale also correlated .51 and -.35 with Negative Affect and Positive Affect, respectively. Thus, both of the general PANAS-X scales are significantly related to state measures of depression and anxiety. Finally, in a sample of 398 SMU undergraduates, general Negative Affect (using Past Few Weeks instructions) correlated .74 with the total score on the HSCL, a widely used measure of clinical symptomatology.

Studies of intraindividual mood fluctuation. A second line of evidence indicates that, when used with short-term time instructions (i.e., Moment or Today), the PANAS-X scales are sensitive to changing internal or external circumstances. We have used the general PANAS-X scales in four within-subjects investigations that illustrate their utility in studying qualitatively distinctive intraindividual mood fluctuations. In the first study, 80 SMU students completed a mood questionnaire consisting of the two higher order scales each evening for 5-7 weeks, using Today instructions (Watson, 1988a). At each assessment the subjects also estimated their daily social activity (number of hours spent with friends that day), rated the level of stress they had experienced, and noted whether or not they had exercised during the day. A total of 3,554 measurements were collected (M = 44.4 per subject). As hypothesized (see Clark & Watson, 1988, 1990b; Watson, 1988a; Watson & Tellegen, 1985), within-subject variations in perceived stress were strongly correlated with fluctuations in Negative Affect, but not in Positive Affect. Also, as expected, social activity and physical exercise were more highly related to Positive Affect than to Negative Affect.

The second study was primarily concerned with diurnal

variation in mood (Clark, Watson, & Leeka, 1989). The subjects were 196 SMU students who completed the two general PANAS-X scales approximately seven times a day for 1 week. Subjects also rated their current stress, and noted whether or not they had engaged in a number of activities (including social interaction and exercise) within the past hour. A total of 8,700 assessments were collected (M = 44.4 per subject). As hypothesized, Positive Affect showed significant diurnal variation: It rose sharply from early morning until noon, remained relatively constant until 9 p.m., and then fell rapidly. Negative Affect, however, did not exhibit a systematic diurnal pattern in these data. Furthermore, replicating the results of the earlier study, perceived stress was again strongly correlated with within-subject fluctuations in Negative Affect but not Positive Affect. And, as before, social interaction and exercise were more strongly related to Positive Affect than to Negative Affect (Clark & Watson, 1990b).

In the third study (McIntyre, Watson, & Cunningham, 1990), 18 students completed a mood questionnaire (including the two general scales) at the beginning of a 1-week period in order to establish baseline levels on the two higher order dimensions. Three other questionnaires were completed within the week by each subject: Once after social interaction, once after physical exercise, and once prior to a stressful examination. Consistent with the results of the other studies, within-subjects analyses of variance revealed that Positive Affect was increased significantly by social interaction and exercise, but was not affected by test stress; conversely, Negative Affect was increased significantly by the stressful examination, but was not influenced by social activity or exercise.

In the fourth study (Watson et al., 1992, Study 2), 127 SMU undergraduates completed the two higher order PANAS-X scales each evening for 5-7 weeks using Today instructions (M = 42.7 assessments per subject); a subset of these subjects (N = 96) also completed the three lower order positive emotion scales (i.e., Joviality, Self-Assurance, and Attentiveness). At each assessment, all subjects also rated the amount of time they spent in each of 21 social/interpersonal activities during that day. Ratings on these activities were summed to create three subscales (Social Entertainment, Active Participation, Social Responsibilities) as well as a measure of Overall Social Activity. Replicating and extending previous research in this area, within-subject correlational analyses indicated that all four positive affect scales were significantly related to three of the social activity measures (Social Entertainment, Active Participation, Overall Social Activity); in contrast, general Negative Affect was completely unrelated to social activity.

Finally, data from a sample of 308 undergraduates who were retested after 2 months on the PANAS-X using Past Week instructions are shown in the first column of Table 20. These correlations range from .23 (for Surprise) to .49 (for Shyness), indicating a moderate level of stability. These results suggest that subjects have a characteristic range of affect within which short-term fluctuations occur. In summary, the PANAS-X scales are strongly correlated with commonly used measures of state affect and current psychiatric symptomatology, and are sensitive to changing endogenous and exogenous conditions. These data indicate that the PANAS-X scales can be used validly to assess short-term, state affect.

Table 20 *Test-Retest Reliabilities of the* PANAS-X *Scales (2-Month Retest Interval)*

	Past Week Instructions	General Instructions	
Scale	(N = 308)	(N = 502)	(N = 399)
Higher Order Scales			
Positive Affect	.43	.70	.64
Negative Affect	.41	.71	.59
Lower Order Scales			
Fear	.35	.62	.57
Hostility	.39	.65	.58
Guilt	.36	.68	.65
Sadness	.35	.62	.60
Shyness	.49	.70[a]	.64
Fatigue	.30	.57[a]	.53
Surprise	.23	.56	.52
Joviality	.43	–[b]	.64
Self-Assurance	.47	–[b]	.68
Attentiveness	.42	–[b]	.55
Serenity	.32	–[b]	.51

Note. All correlations significant at $p < .01$, two-tailed. [a]Data are based on preliminary, 3-item versions of these scales. [b]Not assessed in this sample.

C. Using the PANAS-X Scales as Measures of Trait Affect

Test-retest stability data. What evidence supports the validity of the PANAS-X scales as trait measures, that is, as measures of long-term, individual differences in affectivity? In this regard, we have already presented data indicating that--with the exception of Surprise--trait (i.e., General) versions of the PANAS-X scales are significantly correlated with corresponding judgments made by well-acquainted peers. Moreover, most of these scales also displayed acceptable discriminant validity. Thus, the PANAS-X scales generally appear to possess adequate external validity.

Test-retest reliability also is an important consideration in establishing the construct validity of trait measures. Table 20 presents 2-month stability coefficients--based on General instructions--for nine of the PANAS-X scales in one sample of undergraduates (N = 502) and all of the scales in another (N = 399). Scores on all of these scales were quite stable,

with retest coefficients ranging from .51 (for Serenity) to .71 (for general Negative Affect in the larger sample). As expected, the stability coefficients based on General instructions are consistently higher than those based on Past Week instructions, which further validates the use of different time instructions with the PANAS-X scales.

We have also completed one longer-term analysis of temporal stability on the two higher order scales. A sample of 239 SMU undergraduates initially completed the general Negative Affect and Positive Affect scales (using General instructions) from January, 1985 to September, 1988; these subjects were reassessed on these scales (again using General instructions) during the spring and summer of 1993. Overall, the retest interval ranged from 56 to 99 months (*M* retest interval = 72.4 months). It also should be noted that all subjects graduated from the university during this intervening time period. Thus, all of the respondents experienced a major life change over the course of the study.

Stability correlations from this study are presented in Table 21. The most important aspect of these data is that both of the higher order scales demonstrated significant, moderate stability over this extended time interval, with retest correlations of .43 and .39 for Negative Affect and Positive Affect, respectively. Note also that the discriminant correlations were substantially lower, again demonstrating the discriminant validity of these scales (for more details regarding this study, see Watson & Walker, 1996).

Table 21 *Long-term Retest Stabilities of the Higher Order PANAS-X scales (Mean Retest Interval = 72.4 Months)*

	Time 1 Score	
Time 2 Score	Positive Affect	Negative Affect
Positive Affect	.42	-.24
Negative Affect	-.18	.43

Note. *N* =237. All scales were assessed using General Time Instructions. All correlations significant at $p < .01$, two-tailed.

Convergence with aggregated state measures. Another important approach to validating the trait versions of the PANAS-X scales is to demonstrate that they converge with aggregated measures of shorter term affect. We have conducted two studies of this type. First, 80 undergraduates enrolled in an undergraduate personality course at the University of Texas at Dallas completed the higher order Positive Affect and Negative Affect scales (using General, trait instructions) during the first week of class. Subjects then were reassessed on these scales (using Past Week instructions) once a week over the next 13 weeks. All subjects completed a minimum of 7 weekly mood ratings (*M* = 12.2 assessments per subject). Mean weekly Positive Affect and Negative Affect scores then were created by averaging each subject's responses over the entire weekly rating period.

Correlations between these average weekly scores and the trait versions of the scales are reported in the upper half of Table 22. Again, both of the higher order scales exhibited an impressive convergent/discriminant pattern. That is, the convergent correlations for both scales are high (*r* s = .66 and .48 for Positive Affect and Negative Affect, respectively), whereas the discriminant correlations are low and non-significant.

In the second study, 410 SMU undergraduates completed a trait form of the two higher order PANAS-X scales, and also rated themselves on these scales (using Today instructions) on a daily basis over a period of several weeks. As in previous analyses of this type, all subjects completed a minimum of 30 daily mood assessments (*M* = 43.3 assessments per subject); all of the ratings were completed in the evening, so that they would provide a reasonable estimate of the subjects' moods over the course of the day. Mean daily Positive Affect and Negative Affect scores then were created by averaging each subject's responses over the entire daily rating period.

Correlations between the trait scales and average daily mood scores are displayed in the lower half of Table 22, and it can be seen that they replicate closely the results of the weekly study. Again, the convergent correlations for both scales are high (*r*s = .64 and .53 for Positive Affect and Negative Affect, respectively), whereas the discriminant correlations are low and non-significant.

Table 22 *Correlations Between General Trait Ratings on the Higher Order PANAS-X Scales and Aggregated State Ratings*

Aggregated	General Ratings	
State Ratings	Positive Affect	Negative Affect
Mean Weekly Ratings (N = 80)		
Positive Affect	.66*	-.10
Negative Affect	-.02	.48*
Mean Daily Ratings (N = 410)		
Positive Affect	.64*	-.03
Negative Affect	-.06	.53*

*$p < .05$, two-tailed.

A subset of these subjects (*N* = 211) completed the full, 60-item PANAS-X, so that comparable data also were available

for the 11 lower order scales. It is noteworthy that six scales had convergent correlations of .50 or greater: Joviality (.59), Guilt (.59), Attentiveness (.58), Sadness (.54), Self-Assurance (.53), and Shyness (.51). Four additional scales had convergent coefficients in the .40 to .50 range: Serenity (.47), Fear (.46), Hostility (.46), and Fatigue (.41). It is noteworthy that Surprise produced the lowest convergent correlation (.36), again challenging the validity of the trait form of this scale.

Table 23 *Varimax-Rotated Factor Loadings of Personality, Emotionality, and Higher Order* PANAS-X *Scales (General Instructions) in a Sample of SMU Undergraduates*

	Factor		
Scale	1	2	3
GTS Negative Temperament	.93		
EPQ Neuroticism	.90		
Goldberg Neuroticism	.84		
PANAS-X Negative Affect	.83		
GTS Positive Temperament		.87	
Goldberg Extraversion		.86	
EPQ Extraversion		.83	
PANAS-X Positive Affect		.79	
GTS Disinhibition			.87
EPQ Psychoticism			.74
Goldberg Conscientiousness			-.83

Note. N = 231. Loadings below |.30| are omitted. GTS = General Temperament Survey; EPQ = Eysenck Personality Questionnaire; PANAS-X = Positive and Negative Affect Schedule (Expanded Form). Adapted from Watson & Clark, 1997, Table 4).

Relations with measures of personality and emotionality. We have also collected extensive data relating trait versions of the higher order scales to various measures of personality and emotionality. These data consistently yield a clear convergent/discriminant pattern: Trait Negative Affect is substantially correlated with measures of Neuroticism or Negative Emotionality, but is generally unrelated to measures of Extraversion or Positive Emotionality. Conversely, trait Positive Affect is strongly related to Extraversion and Positive Emotionality, but not to Neuroticism or Negative Emotionality (Watson & Clark, 1984, 1992b, 1997). Table 23 presents illustrative data from a sample of 231 SMU undergraduates. These subjects were assessed on 11 measures: [a] trait versions of the two higher order PANAS-X scales; [b] the Neuroticism, Extraversion, and Psychoticism scales from the Eysenck Personality Questionnaire (EPQ; Eysenck & Eysenck, 1975); [c] the Neuroticism, Extraversion, and Conscientiousness scales developed by Goldberg (1983); and [d] the Negative Temperament, Positive Temperament and Disinhibition scales from the General Temperament Survey (GTS; Clark & Watson, 1990a). It should be noted that the GTS Negative and Positive Temperament scales are true-false measures designed to assess trait Negative and Positive Affectivity, respectively.

Scores on these 11 personality and affectivity measures were subjected to a principal components analysis. An inspection of the unrotated eigenvalues revealed three strong factors with eigenvalues greater than 1.00; together, these factors accounted for 74.5% of the total variance. These three factors were rotated using varimax and the resulting loadings are shown in Table 23. The first dimension clearly can be identified as Neuroticism/Negative Emotionality, and it is noteworthy that the PANAS-X Negative Affect scale loads as strongly on this factor (.83) as do the other measures. Similarly, the second dimension can be interpreted as Extraversion/Positive Emotionality; the PANAS-X Positive Affect scale loads .79 on this factor. Finally, the third factor can be identified as Disinhibition versus Conscientiousness.

Similarly, we have also related trait scores on the lower order scales to measures of personality and emotionality. These analyses have yielded results that are broadly similar to those observed at the higher order level. That is, the specific negative affects are moderately to strongly correlated with measures of Neuroticism/Negative Emotionality, whereas the individual positive affects are significantly related to Extraversion/Positive Emotionality. However, some interesting additional findings have been noted as well. Table 24 presents illustrative data from a sample of 325 SMU undergraduates. These subjects were assessed on 20 personality and affectivity scales: [a] trait scores on the basic negative emotion (Fear, Sadness, Guilt, Hostility) and positive emotion (Joviality, Self-Assurance, Attentiveness) scales from the PANAS-X; [b] the three GTS scales; [c] the Neuroticism, Extraversion, Conscientiousness, Agreeableness, and Openness scales from the NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992); and [d] factor scales of the same "Big Five" personality traits constructed from an expanded, 80-item version of Goldberg's (1983) scales (see Watson & Clark, 1992b, for a discussion of the development of these factor scales).

An inspection of the unrotated eigenvalues indicated the presence of five strong factors, which jointly accounted for 74.9% of the total variance. The varimax-rotated loadings from this solution are shown in Table 24. Corroborating the results obtained with the general Negative Affect scale, all four negative emotions loaded strongly on the Neuroticism/Negative Emotionality factor. Fear and Guilt were essentially pure markers of this factor, whereas Sadness also had a modest secondary loading on the Extraversion dimension. Hostility, however, split between the Neuroticism and Agreeableness factors, and actually had its highest loading (-.64) on the latter.

Table 24 *Varimax-Rotated Factor Loadings of Personality, Emotionality and Lower-Order* PANAS-X *Scales (General Instructions) in Sample of SMU Undergraduates*

	Factor				
Scale	1	2	3	4	5
NEO-FFI Neuroticism	**.82**	-.21	-.16	.04	-.03
GTS Negative Temperament	**.80**	-.07	.01	-.17	.00
Goldberg Neuroticism	**.80**	-.15	-.13	.00	.02
PANAS-X Fear	**.79**	-.04	.05	-.04	-.10
PANAS-X Guilt	**.76**	-.08	-.17	-.13	-.03
PANAS-X Sadness	**.75**	-.30	.00	-.09	.11
NEO-FFI Extraversion	-.07	**.88**	-.06	.22	.07
Goldberg Extraversion	-.20	**.87**	-.10	.05	.15
PANAS-X Joviality	-.13	**.84**	.13	.27	.00
GTS Positive Temperament	-.19	**.78**	.28	.02	.10
PANAS-X Self-Assurance	-.35	**.59**	.25	-.36	.16
Goldberg Conscientiousness	-.06	.06	**.89**	.04	.02
NEO-FFI Conscientiousness	-.11	.10	**.87**	.02	-.11
PANAS-X Attentiveness	-.14	.25	**.80**	.00	.07
GTS Disinhibition	-.03	.32	**-.73**	-.38	.09
NEO-FFI Agreeableness	-.12	.13	.07	**.89**	.10
Goldberg Agreeableness	-.05	.25	.13	**.84**	.16
PANAS-X Hostility	**.52**	-.02	-.07	**-.63**	-.06
NEO-FFI Openness	.07	.00	-.03	.16	**.88**
Goldberg Openness	-.10	.30	-.04	.08	**.82**

Note. *N*= 325. Loadings above |.40| are shown in boldface. NEO-FFI = NEO Five-Factor Inventory. Adapted from Watson & Clark (1992b, Table 5).

Looking at the positive affects, Table 24 demonstrates that Joviality and Self-Assurance both loaded strongly and primarily on Extraversion/Positive Emotionality. In addition, Self-Assurance also had moderate negative loadings on both Neuroticism and Agreeableness. Finally, Attentiveness was primarily a marker of Conscientiousness, but also had a modest loading on Extraversion (.25). Thus, the data reveal evidence of relations between personality and affectivity that are both general (i.e., between the negative affects and Neuroticism, and between the positive affects and Extraversion) and specific (e.g., between Hostility and Agreeableness, and between Attentiveness and Conscientiousness).

Summary of trait validity evidence. To summarize the extensive data presented here and elsewhere: trait scores on the PANAS-X scales (a) are stable over time, (b) show significant convergent and discriminant validity when correlated with peer-judgments, (c) are highly correlated with corresponding measures of aggregated state affect, and (d) are strongly and systematically related to measures of personality and emotionality. These data clearly demonstrate that --with the possible exception of Surprise--the PANAS-X scales can be used validly to assess long-term individual differences in affect.

V. References

Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology, 55,* 387-395.

Almagor, M., & Ben-Porath, Y. (1989). The two-factor model of self-reported mood: A cross-cultural replication. *Journal of Personality Assessment, 53,* 10-21.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561-571.

Clark, L. A., & Watson, D. (1988). Mood and the mundane: Relations between daily life events and self-reported mood. *Journal of Personality and Social Psychology, 54,* 296-308.

Clark, L. A., & Watson, D. (1990a). *The General Temperament Survey*. Unpublished manuscript, Southern Methodist University, Dallas, TX.

Clark, L. A., & Watson, D. (1990b). *Predicting momentary mood from ordinary life events and personality*. Unpublished manuscript, Southern Methodist University, Dallas, TX.

Clark, L. A., Watson, D., & Leeka, J. (1989). Diurnal variation in the positive affects. *Motivation and Emotion, 13,* 205-234.

Costa, P. T., Jr., & McCrae, R. R. (1989). *The NEO-PI/NEO-FFI manual supplement*. Odessa, FL: Psychological Assessment Resources.

Derogatis, L. R. (1977). *The SCL-90 Manual I: Scoring, administration, and procedures for the SCL-90.* Baltimore: Johns Hopkins U School of Medicine.

Derogatis, L. R., & Cleary, P. A. (1977). Confirmation of the dimensional structure of the SCL-90: A study in construct validation. *Journal of Clinical Psychology, 33,* 981-989.

Derogatis, L. R., Lipman, R S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Science, 19,* 1-15.

Ekman, P. (1971). Universals and cultural differences in facial expressions of emotions. In J. K. Cole (Ed.), *Nebraska Symposium on Motivation* (Vol. 19, pp. 207-283). Lincoln: University of Nebraska Press.

Eysenck, H. J., & Eysenck, S. B G. (1975). *Manual of the Eysenck Personality Questionnaire*. San Diego, CA: Educational & Industrial Testing Service.

Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology, 55,* 149-158.

Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52,* 409-418.

Goldberg, L. R. (1983, June). *The magical number five, plus or minus two: Some conjectures on the dimensionality of personality descriptions*. Paper presented at a research seminar, Gerontology Research Center, Baltimore, MD.

Izard, C. E. (1972). *Patterns of emotions: A new analysis of anxiety and depression*. New York: Academic Press.

Kennedy-Moore, E., Greenberg, M., Newman, M., & Stone, A. A. The relationship between daily events and mood: The mood measure may matter. *Motivation and Emotion, 16,* 143-155.

Kenrick, D. T., & Stringfield, D. O. (1980). Personality traits and the eye of the beholder: Crossing some traditional philosophical boundaries in the search for consistency in all of the people. *Psychological Review, 87,* 88-104.

Marshall, P. S., Watson, D., Steinberg, P., Cornblatt, B., Peterson, P. K., Callies, A., & Schenck, C. H. (1994). *An assessment of cognitive function and mood in chronic fatigue syndrome*. Manuscript submitted for publication.

Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology, 55,* 102-111.

McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of a five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52,* 81-90.

McIntyre, C. W., Watson, D., & Cunningham, A. C. (1990). The effects of social interaction, exercise, and test stress on positive and negative affect. *Bulletin of the Psychonomic Society, 28,* 141-143.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1971). *Manual: Profile of Mood States*. San Diego: Educational and Industrial Testing Service.

Meyer, G. J., & Shack, J. R. (1989). Structural convergence of mood and personality: Evidence for new and old directions. *Journal of Personality and Social Psychology, 57,* 691-706.

Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality & Social Psychology, 4,* 681-691.

Nowlis, V. (1965). Research with the Mood Adjective Check List. In S. S. Tomkins & C. E. Izard (Eds.), *Affect, cognition, and personality* (pp. 352-389).

New York: Springer Publishing.

Quinn, B. A. (1989). *Religiousness and psychological well-being: An empirical investigation*. Unpublished doctoral dissertation, Wayne State Univ., Detroit, MI.

Radloff, L. S. (1977). The CES-D Scale: A new self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385-401.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Jacobs, G. A., Russell, S., & Crane, R. S. (1983). Assessment of anger: The State-Trait Anger Scale. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment* (Vol. 2, pp. 52-76). Hillsdale, NJ: Erlbaum.

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma and J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681-706). Hillsdale, NJ: Erlbaum.

Watson, D. (1988a). Intraindividual and interindividual analyses of Positive and Negative Affect: Their relation to health complaints, perceived stress, and daily activities. *Journal of Personality and Social Psychology, 54,* 1020-1030.

Watson, D. (1988b). The vicissitudes of mood measurement: Effects of varying descriptors, time frames, and response formats on measures of Positive and Negative Affect. *Journal of Personality and Social Psychology, 55,* 128-141.

Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology, 57*, 120-128.

Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96,* 465-490.

Watson, D., & Clark, L. A. (1989, August). General affective dimensions versus discrete emotional factors: A hierarchical model. In L. A. Clark & D. Watson (Chairs), *The emotional bases of personality*. Symposium presented at the 97th Annual Convention of the Amer. Psychological Assoc., New Orleans, LA.

Watson, D., & Clark, L. A. (1991). Self- versus peer-ratings of specific emotional traits: Evidence of convergent and discriminant validity. *Journal of Personality and Social Psychology, 60,* 927-940.

Watson, D., & Clark, L. A. (1992a). Affects separable and inseparable: On the hierarchical arrangement of the negative affects. *Journal of Personality and Social Psychology, 62,* 489-505.

Watson, D., & Clark, L. A. (1992b). On traits and temperament: General and specific factors of emotional experience and their relation to the five-factor model. *Journal of Personality, 60,* 441-476.

Watson, D., & Clark, L. A. (1997). Extraversion & its positive emotional core. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 767-793). San Diego, CA: Academic Press.

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology, 97,* 346-353.

Watson, D., Clark, L. A., McIntyre, C. W., & Hamaker, S. (1992). Affect, personality, and social activity. *Journal of Personality and Social Psychology, 63,* 1011-1025.

Watson, D., Clark, L. A., & Tellegen, A. (1984). Cross-cultural convergence in the structure of mood: A Japanese replication and a comparison with U. S. findings. *Journal of Personality and Social Psychology, 47,* 127-144.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of Positive and Negative Affect: The PANAS Scales. *Journal of Personality and Social Psychology, 54,* 1063-1070.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98,* 219-235.

Watson, D., & Walker, L. M. (1996). The long-term stability and predictive validity of trait measures of affect. *Journal of Personality and Social Psychology, 70,* 567-577.

Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology, 43,* 111-122.

Zuckerman, M., & Lubin, B. (1965). *Manual for the Multiple Affect Adjective Check List*. San Diego, CA: Educational and Industrial Testing Service.