

# Cognitive-Behavioral Assessment of Depression: Clinical Validation of the Automatic Thoughts Questionnaire

Thomas H. Harrell and Nancy B. Ryon  
Florida Institute of Technology

Hollon and Kendall's (1980) Automatic Thoughts Questionnaire (ATQ-30) was designed to identify and measure the frequency of occurrence of automatic negative thoughts associated with depression. This 30-item self-statement inventory was constructed and cross-validated using male and female undergraduates as subjects. The present study examined the applicability of the ATQ-30 to clinical populations. Compared to nondepressed mental-health-center clients and nondepressed medical patients, significantly higher ATQ-30 scores were found for mental-health-center clients who met the criteria for inclusion in a depressed group. The ATQ-30 correlated significantly with therapist ratings of depression, the Minnesota Multiphasic Personality Inventory Depression scale, and the Beck Depression Inventory. Split-half and coefficient alpha reliability coefficients were also highly significant, as were item-to-total correlations. The ATQ-30 appears to be a valid and reliable measure of depression-related cognitions in clinical as well as nonclinical populations.

The study of cognitive factors in depression has led to the belief that the etiology and maintenance of this clinical syndrome are, at least in part, a consequence of maladaptive thinking styles and negative self-statements (Beck, 1976). Therapeutic procedures designed to alter negative beliefs have proved particularly effective in both symptom reduction and prevention of relapse (Rush, Beck, Kovacs, & Hollon, 1977; Kovacs, Rush, Beck, & Hollon, 1981). In order to understand their role and confirm the effects of treatment, cognitions must be assessed in a reliable and valid way (Kendall & Korgeski, 1979).

Studies of perceptual and recall factors in depression have revealed possible perceptual deficits associated with depression (Krantz & Hammen, 1979) and better recall of negative material by depressed subjects (Nelson & Craighead, 1977). Investigations of thinking styles and personal constructs have focused on either specification of thinking errors, identification of automatic thoughts or irrational beliefs, or assessment of the nature of response biases (Shaw & Dobson, 1981). Questionnaires

have recently been developed to examine depressive distortion (Krantz & Hammen, 1979; Lefebvre, 1981), to determine the relationship between attitudes and the tendency to become depressed (Weissman & Beck, Note 1), and to assess automatic negative thoughts associated with depression (Hollon & Kendall, 1980). However, as the small number of measures reported in the literature indicates, cognitive-behavioral assessment of depression is not well advanced, and evidence for the validity and reliability of measures is sparse.

Hollon and Kendall (1980) developed the Automatic Thoughts Questionnaire (ATQ-30) to identify and measure the frequency of occurrence of automatic negative thoughts associated with depression. This 30-item self-statement inventory, constructed and cross-validated using male and female undergraduates as subjects, significantly discriminated subclinically depressed from nondepressed criterion groups. The purpose of the present investigation was to investigate the reliability and validity of the ATQ-30 in a clinical population.

## Method

### *Subjects*

Participants were 114 male and female clients of a local mental health center and office patients of two physicians in private practice. Subjects ranged in age from 18 to 55.

---

The authors wish to thank the anonymous reviewers for their helpful comments on an earlier version of this article.

Requests for reprints should be addressed to Thomas H. Harrell, School of Psychology, Florida Institute of Technology, Melbourne, Florida 32901.

Sixty-one members of this sample met criteria for inclusion into one of three groups; a clinically depressed group (DEP), a nondepressed psychopathology group (NDEP-P), and a nondepressed medical group (NDEP-M).

Mental-health-center clients scoring 17 and above on the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), 70T and above on the Minnesota Multiphasic Personality Inventory (MMPI) Depression scale (*D*), and who were rated three or higher on a depression rating scale by their therapists were included in the DEP group. The NDEP-P criterion group, also composed of mental-health-center clients, and the NDEP-M group, consisting of patients seeking treatment for medical problems, were formed on the basis of scores of 9 or lower on the BDI, 59T and below on the MMPI-*D*, and depression rating scale values of 2 or less.

### Instruments

The MMPI-*D* scale has been widely used as a screening instrument for selecting depressed and nondepressed samples for research purposes (e.g., Hollon & Kendall, 1980; Lewinsohn, Mischel, Caplin, & Barton, 1980). The scale consists of 60 true-false items and is considered to be a sensitive index of current mood regardless of general adjustment. Higher scores reflect greater severity of depression and elevation tends to increase with age.

The BDI is a 21-item scale assessing affective, behavioral, cognitive, motivational, and physical aspects of depression. Each item consists of four related statements, scored on a range from 0 to 3. Total scores range from 0 to 63. The BDI has demonstrated good concurrent validity (Beck et al., 1961).

The 30-item ATQ-30, devised by Hollon and Kendall (1980), measures the frequency of occurrence of automatic negative thoughts (negative self-statements) associated with depression. It has been found to significantly discriminate subclinically depressed from nondepressed criterion groups of male and female college undergraduates. Subjects are instructed to read each thought and indicate, on a scale of 1 to 5, how frequently, if at all, that thought occurred to them over the last week. Scores range from 30 to 150. In the Hollon and Kendall (1980) sample, the mean ATQ-30 score for depressed student subjects was 79.64 ( $SD = 22.29$ ), whereas the mean ATQ-30 score for nondepressed subjects was 48.57 ( $SD = 10.89$ ).

A 5-point Likert-type scale permitting the therapist or physician to indicate a rating of the overall severity of depression was devised. Guidelines for completion of the Likert scale were based on the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III; American Psychiatric Association, 1980) diagnostic criteria for Major Depressive Episode (without psychotic features) and Dysthymic Disorder (Depressive Neurosis) and the DSM-III Draft (APA, 1978) criteria for levels of severity of a Depressive Episode. The scale was devised to obtain a numerical rating on the continuum of severity of depression, which is not provided in DSM-III. DSM-III diagnostic criteria were used to determine the presence or absence of a Major Depressive Episode or Dysthymic Disorder. The DSM-III draft criteria for levels of severity of the disorder were used to indicate a rating of overall severity (no depression, mild or nonclinical, moderate, marked, severe). This resulted in a single quantitative measure of severity of depression as evaluated by the interviewer.

### Procedure

Subjects were invited to participate in an investigation of the utility of psychological tests. Volunteers received a test packet containing a consent form and the three self-report questionnaires in one of three counterbalanced sequences. Severity of depression was rated on the rating scale by a therapist at the mental health center or by one of the participating physicians following an extensive intake interview. Subjects requesting additional information regarding the purpose of the study were given a one-page statement by the therapist or physician after completing the questionnaires. Data collection continued until a minimum of 20 subjects for each of the three groups met the criteria for inclusion. All subjects who initially agreed to participate completed the test packet.

### Results

#### Criterion Group Characteristics

Twenty-one subjects (7 males and 14 females) met the criteria for inclusion in the depressed group. The nondepressed psychopathology and nondepressed medical criterion groups consisted of 20 subjects each (NDEP-P, 9 males and 11 females; NDEP-M, 11 males and 9 females).

As the proportion of females to males was higher in the depressed group than in the two nondepressed groups, chi-square analyses were conducted. Comparison of the male/female ratio across the three groups was not significant,  $\chi^2(2) = 1.96$ . A second chi-square for the comparison of the male/female ratio in the depressed group to the male/female ratio in the nondepressed groups combined was also nonsignificant,  $\chi^2(1) = 1.55$ . The analyses indicate that for groups of this size the 2 to 1 ratio of females to male subjects in the depressed sample did not constitute a significant gender difference. Additionally, since 18% to 23% of adult females and only 8% to 11% of adult males are estimated to have at some time had a major depressive episode (APA, 1980), the ratio of female to male depressed subjects in this investigation appears to reflect the base rate for the general population.

The depressed and nondepressed psychopathology groups did not differ significantly in age (DEP,  $M = 31.71$ ,  $SD = 9.98$ , range = 19 to 53 years; NDEP-P,  $M = 32.6$ ,  $SD = 6.62$ , range = 20 to 47 years). The nondepressed medical patient sample was significantly older ( $M = 39.15$ ,  $SD = 10.51$ , range = 20 to 51 years);  $F(2, 58) = 3.93$ ,  $p < .05$ .

Ten of the DEP subjects were receiving tri-

cyclic antidepressants, and 5 of the NDEP-P and 4 of the NDEP-M subjects were receiving benzodiazepines. Two NDEP-P subjects were receiving butyrophene medications. However, there were no significant differences between subjects receiving medication and non-medicated subjects on the criterion measures.

Depressed subjects had significantly higher mean scores than nondepressed subjects on both the BDI,  $F(2, 58) = 152.70, p < .001$ , and the MMPI-D,  $F(2, 58) = 136.77, p < .001$ . The mean therapist rating was also significantly higher for the depressed criterion group,  $F(2, 58) = 182.30, p < .001$ . Scheffé and Duncan a posteriori contrast procedures revealed no significant differences between the two nondepressed groups on these measures. As these criterion checks indicate, compared to both nondepressed criterion groups, the depressed sample showed evidence psychometrically and by interview of moderate to severe clinical depression.

#### Criterion Group Differences on the ATQ-30

One-way analysis of variance (ANOVA) indicated significantly higher total scores on the ATQ-30 for the depressed subjects  $F(2, 58) = 78.20, p < .001$ . The mean for this group was 88.90 ( $SD = 21.15$ ), whereas the mean for the nondepressed medical patients averaged 38.35 ( $SD = 8.17$ ). The NDEP-P and NDEP-M groups did not display significantly different total ATQ-30 scores. These two groups formed a homogeneous subset when the criterion groups were analyzed using Scheffé and Duncan procedures. Table 1 presents the means and standard deviations of the inventory scores and therapist ratings for each of the three criterion groups.

In order to examine the influence of gender on these findings, a post hoc two-way ANOVA was performed on ATQ-30 scores, using group and sex as factors. A significant main effect was again obtained for group,  $F(2, 55) = 73.39, p < .001$ , whereas the main effect for sex,  $F(1, 55) = .42$ , and the interaction,  $F(2, 55) = .83$ , failed to reach significance.

A direct discriminate analysis using all ATQ-30 items was performed. Wilks's lambdas, evaluated using the  $F$ -test approximation, were significant at the .001 level for all items. Depressed subjects demonstrated significantly higher mean scores on all 30 items than sub-

Table 1  
Means and Standard Deviations of Inventory Scores and Therapist Ratings for Depressed and Nondepressed Criterion Groups

Measure	DEP ( <i>n</i> = 21)		NDEP-P ( <i>n</i> = 20)		NDEP-M ( <i>n</i> = 20)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ATQ-30	88.90	21.15	42.35	9.96	38.35	8.17
BDI	25.00	6.63	4.25	2.73	4.00	2.40
MMPI-D	85.19	10.99	49.05	5.18	49.70	6.58
Rating	3.24	.43	1.25	.44	1.10	.31

Note. DEP = depressed; NDEP-P = nondepressed psychopathology; NDEP-M = nondepressed medical; ATQ-30 = the Automatic Thoughts Questionnaire; BDI = the Beck Depression Inventory; MMPI-D = the Minnesota Multiphasic Personality Inventory, Depression scale.

jects in either of the two nondepressed groups. Classification subanalysis used the linear combination of all 30 items. This analysis yielded correct classification of 100.0% of the DEP subjects, 84.2% of the NDEP-P subjects, and 95.0% of the NDEP-M group. No nondepressed subjects were incorrectly classified as depressed. The overall percentage of correctly classified cases was 93.1%.

#### Interrelationships Among Measures

Intercorrelations among the ATQ-30, the BDI, the MMPI-D, and the therapist ratings were calculated for the full criterion sample (all three criterion groups combined). The obtained correlation coefficients were significant at the .001 level.

However, these coefficients may reflect constraints on test scores due to the criteria used for inclusion into the criterion groups, resulting in artificial inflation of the correlations. In order to examine the covariational relationship among the measures in subjects demonstrating the full range of scores, intercorrelations were computed for the total sample (including those participants who did not qualify for inclusion into one of the three criterion groups). These correlations were also significant at the .001 level. The obtained correlation coefficients for the criterion sample and the total sample are displayed in Table 2.

#### Reliability

Split-half and coefficient alpha reliability coefficients were computed to estimate the re-

Table 2  
*Intercorrelations of Inventory Scores and  
 Therapist Ratings for the Combined Criterion  
 Groups and the Total Sample*

Measure	BDI	MMPI-D	Rating
ATQ-30	.87 (.79)	.85 (.69)	.79 (.58)
BDI		.89 (.71)	.86 (.63)
MMPI-D			.86 (.60)

Note.  $n = 61$  for combined criterion groups sample;  $n = 114$  for total sample,  $r$  value shown in parentheses; All correlations are significant at  $p < .001$ . BDI = the Beck Depression Inventory; MMPI-D = the Minnesota Multiphasic Personality Inventory, Depression scale; ATQ-30 = the Automatic Thoughts Questionnaire.

liability of the ATQ-30 with a clinical population. Nine cases with missing data were deleted for these analyses (two from the DEP group, one from the NDEP-P group, and six from the remaining participants). The split-half reliability coefficient, calculated on odd versus even items, for the three criterion groups combined and for the total sample was .96 ( $p < .001$ ) and .96 ( $p < .001$ ), respectively. Coefficient alpha for both of these samples was .98 ( $p < .001$ ).

Reliability measures for the depressed cases revealed a coefficient alpha of .94 ( $p < .001$ ) and a split-half coefficient of .91 ( $p < .001$ ). For the nondepressed medical patients, the reliability coefficients were as follows: coefficient alpha, .91 ( $p < .001$ ) and split-half, .87 ( $p < .001$ ). For the nondepressed psychopathology cases these values were as follows: coefficient alpha, .89 ( $p < .001$ ) and split-half, .59 ( $p < .01$ ).

All ATQ-30 item-to-total correlations for the criterion groups sample and for the total sample were significant at the .001 level. These correlations ranged from .56 to .91.

## Discussion

The findings of this investigation of the ATQ-30 support its applicability to clinical populations. A sample of clinically depressed, treatment-seeking clients was found to differ significantly in reported frequency of automatic negative thoughts from two nondepressed samples.

Defining depression in a reliable and valid way is a challenge for both research and clinical

practice. Rather than rely on either clinical observation or personality inventories alone, this investigation combined these criteria for the purpose of forming the depressed and nondepressed groups. The results of the ANOVAs indicated the depressed group obtained significantly higher scores than the nondepressed groups on the BDI, the MMPI-D, and therapist ratings.

The subjects forming the nondepressed groups in this study were either clients of the mental health center or medical patients. These subjects were, therefore, presumably under at least some degree of life stress and/or were diagnosed as having a syndrome psychopathology. Nevertheless, the ATQ-30 significantly discriminated the group defined as depressed from the nondepressed samples. This result suggests that in addition to providing quantitative measurement of reported dysfunctional automatic thoughts, the ATQ-30 also demonstrates considerable specificity for depressive cognitions, as score elevations were not evidenced by nondepressed individuals with diagnosed syndrome psychopathologies or by medical outpatients.

The ATQ-30 demonstrated significant correlation with the BDI, with which it shares a common theoretical base. The ATQ-30 also correlated well with the MMPI-D scale. These findings indicate that the ATQ-30 displays adequate concurrent validity.

The present study was further designed to investigate the reliability of the ATQ-30 with a clinical population. All reliability coefficients were significant at or beyond the .05 level of significance. The results corroborate the Hollon and Kendall (1980) finding that the inventory possesses sufficient reliability to justify its use as a general measure of depression-related automatic negative thoughts.

Like the BDI, the ATQ-30 is a brief, easily administered inventory for the quantitative assessment of the severity of depression. The use of both instruments would enable the diagnostician-therapist to check the consistency of his or her data, while eliciting a broad range of qualitative information for planning treatment. The mean scores of 88.90 for the clinically depressed subjects and 79.64 for the subclinically depressed college students, compared to the mean scores of 38.35, 42.35, and 48.57 for the two nondepressed samples of

this study and the Hollon and Kendall sample, provide general guidelines for interpreting the questionnaire for clinical purposes.

The ATQ-30 identifies cognitions associated with clinical depression. A principal task for the cognitive-behavioral therapist is to reduce depression by systematically altering these negative beliefs. Further study of the ATQ-30 is needed to determine if the measure is sensitive to change as a result of therapeutic interventions, and, therefore, if it is useful as an objective measure for judging improvement resulting from cognitive-behavioral therapy or other forms of treatment.

#### Reference Note

Weissman, A. N., & Beck, A. T. *Development and validation of the dysfunctional attitude scale: A preliminary investigation*. Paper presented at the meeting of the American Educational Research Association, Toronto, 1978.

#### References

- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, D.C.: Author, 1980.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders* (draft; 3rd ed.). Washington, D.C.: Author, 1978.
- Beck, A. T. *Cognitive therapy and the emotional disorders*. New York: International Universities Press, 1976.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. An inventory for measuring depression. *Archives of General Psychiatry*, 1961, 4, 561-571.
- Hollon, S. D., & Kendall, P. C. Cognitive self-statements in depression: Development of an Automatic Thoughts Questionnaire. *Cognitive Therapy and Research*, 1980, 4, 383-395.
- Kendall, P. C., & Korgeski, G. P. Assessment and cognitive-behavioral interventions. *Cognitive Therapy and Research*, 1979, 3, 1-21.
- Kovacs, M., Rush, A. J., Beck, A. T., & Hollon, S. D. Depressed outpatients treated with cognitive therapy or pharmacotherapy. *Archives of General Psychiatry*, 1981, 38, 33-39.
- Krantz, S., & Hammen, C. L. Assessment of cognitive bias in depression. *Journal of Abnormal Psychology*, 1979, 88, 611-619.
- Lefebvre, M. F. Cognitive distortion and cognitive errors in depressed psychiatric and low back pain patients. *Journal of Consulting and Clinical Psychology*, 1981, 49, 517-525.
- Lewinsohn, P. M., Mischel, W., Chaplin, W., & Barton, R. Social competence and depression: The role of illusory self-perception. *Journal of Abnormal Psychology*, 1980, 89, 203-212.
- Nelson, R. E., & Craighead, W. E. Selective recall of positive and negative feedback, self-controlled behaviors, and depression. *Journal of Abnormal Psychology*, 1977, 86, 379-388.
- Rush, A. J., Beck, A. T., Kovacs, M., & Hollon, S. D. Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients. *Cognitive Therapy and Research*, 1977, 1, 17-37.
- Shaw, B. F., & Dobson, K. S. Cognitive assessment of depression. In T. V. Merluzzi, C. R. Glass, & M. Genest (Eds.), *Cognitive assessment*. New York: Guilford Press, 1981.

Received February 7, 1983  
Revision received April 21, 1983 ■