

Cloninger's Tridimensional Personality Questionnaire: Psychometric Properties and Construct Validity in Taiwanese Adults

Wei J. Chen, Hsing-Me Chen, Chwen-Cheng Chen, Chiao-Chicy Chen, Wu-Yang Yu,
and Andrew T.A. Cheng

Despite the wide use of the Tridimensional Personality Questionnaire (TPQ) in Western populations as a useful tool integrating both genetic and environmental influences on personality, some of its constructs remain questionable. In this study, we examined the psychometric properties of the TPQ and its relationship with aggression in Taiwanese adults. The subjects were 201 Taiwanese adults of wide ranges in age and educational level. Subjects were assessed using a Chinese version of the TPQ and the Brown-Goodwin Aggression Inventory. The internal consistency of the Chinese version of the TPQ scales is found to be mostly in the acceptable range except for the reward dependence (RD) scale and its subscales. The results

of factor analysis of the 12 TPQ subscales partially support the four-factor model rather than the original three-factor model. The construct validity of the novelty seeking (NS) and harm avoidance (HA) dimensions is supported by the findings that the NS is negatively correlated with age, the NS1 subscale is slightly negatively correlated with all the HA subscales, and the NS is positively correlated with aggression. We conclude that both the HA and NS scales of the TPQ are cross-culturally robust, while the RD scale needs to be refined.

Copyright 2002, Elsevier Science (USA). All rights reserved.

A LARGE NUMBER of studies have consistently revealed that there are three to five discrete quantifiable and heritable traits that comprise human personality and predispose individuals to behave in certain consistent ways.¹⁻⁶ Recently, there has been increased interest in applying molecular genetic techniques to decipher the genes responsible for the variations in personality trait.⁷ In this regard, the biosocial theory of personality put forth by Cloninger^{1,8} is heuristically useful because it attempts to integrate information from family studies, longitudinal development research, psychometric descriptions of personality structure, and neuropharmacological and neuroanatomical studies of conditioning and learning in man and animals.

Cloninger^{1,8} proposed that there were three genetically independent dimensions of personality: novelty seeking (NS), harm avoidance (HA), and

reward dependence (RD). NS is characterized by an inherited tendency toward the activation of behavior such as exploratory activity in response to novelty, impulsive decision making, extravagance in the approach to cues of reward, and quick loss of temper. HA is defined as an inherited tendency toward the inhibition or cessation of behaviors, such as pessimistic worry in anticipation of future problems, passive avoidant behaviors such as fear of uncertainty and shyness of strangers, and rapid fatigability. RD is assumed to be an inherited tendency toward the maintenance of ongoing behaviors, and is exhibited as sentimentality, persistence, social attachment, and dependence on the approval of others. Individual differences along these three dimensions reflect variations in the monoaminergic brain systems of behavioral activation (dopamine), behavioral inhibition (serotonin), and behavioral maintenance (norepinephrine). However, in response to novel, aversive, and appetitive stimuli, the three dimensions interact with one another. HA is said to modulate both NS and RD. Thus, on the phenotypic levels, HA is negatively correlated with NS and is positively correlated with RD. On the basis of Cloninger's theory, many predictions pertaining to human behaviors can be made and hence provide possible tests of fundamental assumptions of the theory.⁹ One particular example that has been examined in this study is that aggression will be determined by high NS, low HA, and low RD at all ages.

On the basis of the three-dimensional model,

From the Institute of Epidemiology, College of Public Health, National Taiwan University, Taipei; TakMing College, Taipei; Department of Psychiatry, College of Medicine, National Cheng-Kung University, Tainan; Taipei City Psychiatric Center, Taipei; Kai-Suan Psychiatric Hospital, Kaohsiung; Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.

Supported by a three-year grant from the National Science Council, Taiwan (NSC88-2318-B-001-004-M51).

Address reprint requests to Andrew T.A. Cheng, M.D., Ph.D., Division of Epidemiology and Public Health, Institute of Biomedical Sciences, Academia Sinica, Taipei 108, Taiwan.

*Copyright 2002, Elsevier Science (USA). All rights reserved.
0010-440X/02/4302-0005\$35.00/0
doi:10.1053/comp.2002.30797*

Cloninger developed the Tridimensional Personality Questionnaire (TPQ), a self-report inventory designed to measure the three higher-order dimensions and their 12 subscales.⁸ Cloninger has argued that the proposed three-dimensional factor structure of the TPQ is useful in classifying and describing both normal and abnormal variations in personality. The scales were initially generated a priori, and were then tested empirically using factor analysis. The initial assessment of the psychometric properties of the original 80-item TPQ among 101 medical students indicated that the TPQ did have a three-dimensional factor structure conforming to Cloninger's theory with acceptable internal consistency and test-retest reliability.⁸ These psychometric properties were further confirmed in a study among a US national sample of 1,019 adults using an expanded version of the TPQ, containing 100 items with two items dropped from scoring.¹⁰

Many subsequent studies using various versions of the TPQ have been conducted to examine its reliability and validity in a variety of nonclinical populations, including university students,¹¹⁻¹⁶ adult twins and their families,¹⁷ voluntary adults,^{14,18-20} and representative samples from the general population.²¹ The main findings of these studies are that both the HA and NS scales are psychometrically sound and conform to Cloninger's theory. However, the RD scale is psychometrically weaker and may be composed of two factors, i.e., RD2 (persistence) as an independent factor, while the other three RD subscales represent the RD dimension. In general, these studies were limited to Caucasian (US white, English, Czechoslovakian, Yugoslavian, Canadian, German, and Swedish) and African-American populations. The applicability of the TPQ in Asian populations has not been reported except in one study of Japanese university students.¹⁵

In this study, we examined the psychometric properties of a Chinese version of the TPQ in a randomly selected sample of Taiwanese adults of wide ranges in age and educational level. In addition to internal consistency and factor analysis, the construct validity of the Chinese version of the TPQ was assessed through the relationship between its scales and gender, age, educational level, and aggression.

METHOD

Subjects

Participants of this study were mainly recruited from those who attended the Health Screening at National Taiwan University Hospital from October 1998 to March 1999. The sampling scheme was to recruit 20 consecutively admitted persons for each of five 10-year age intervals (20 to 29, 30 to 39, 40 to 49, 50 to 59, and 60 to 69 years) for each sex. At the end of the study, there were 176 health screening attenders participating in the study, with the age intervals of 20 to 29 not meeting the goal of 20 persons. In addition, 19 staff members and students at National Cheng-Kung University Hospital and 14 staff members at the Institute of Biomedical Sciences, Academia Sinica, were recruited for the completion of the TPQ. A participant's TPQ results were deleted from analysis if any item was missing. In total, 201 subjects (100 males and 101 females) with complete data were analyzed.

Measurement Instruments

Chinese version of the TPQ. TPQ-version IV was employed in this study. This version contains 100 "true" or "false" items covering the three scales, NS, HA, and RD, with four subscales each (Table 1). In accordance with Cloninger et al.,¹⁰ items 61 and 71 were dropped from scoring.

The translation of the TPQ into the Chinese version used in this study was carried out in two stages. In the first stage, the TPQ was translated, revised after pretest in 15 college students, and then applied in a convenient sample of 360 adults in 1990.²² Before the start of this study, a further revision was undertaken by a focus group comprised of five researchers (W.J.C., C.-C.C., C.-C.C., W.Y.Y., and A.T.A.C.). Three of them had obtained their doctoral degrees in the United Kingdom or United States and had many years' experience in translating English questionnaires into Chinese for use in Taiwan. Any suggestions for psycholinguistic modification of individual items from the majority of group members were taken. A field trial was then conducted by three research assistants among 10 normal subjects under the observation of the investigators, by either self-administration or interview. This was to ensure that they understood all of the TPQ items and were able to read and interpret them clearly to the interviewees. If any difficulty of comprehension about a specific item was raised by the interviewees, further revisions were made by the focus group and tested again among interviewees. In total, the original Chinese translates of 22 items were modified further to make them more accurate in reflecting the English meaning and more comprehensive to interviewees.

Chinese version of the Brown-Goodwin Aggression Inventory. We applied the Brown-Goodwin Aggression Inventory²³ to assess the construct validity of the TPQ. High scores on the Brown-Goodwin Aggression Inventory has been found to be correlated with cerebrospinal fluid concentration of monoamine metabolites.^{23,24} A revised version by Dr Lawrence Greenhill at Columbia University (personal communication) was used in this study. It contains 11 items enquiring about aggression events during childhood, adolescence, and adulthood. There were four possible answers to each question: never, rarely (one to three times), occasionally (four to 10 times), and often (> 10 times). For the calculation of the final score, "never" was coded

Table 1. TPQ Scale and Subscale Means (SD) and Internal Consistency (Cronbach's α)

Scale	Features of High Scorers	No. of Items	Men (n = 101)		Women (n = 100)		Total (n = 201)	
			Mean (SD)	α	Mean (SD)	α	Mean (SD)	α
Novelty Seeking								
NS1	Exploratory excitability	9	3.9 (1.7)	0.43	3.9 (1.8)	0.52	3.9 (1.8)	0.47
NS2	Impulsiveness	8	2.0 (1.8)	0.65	2.0 (1.8)	0.63	2.0 (1.8)	0.64
NS3	Extravagance	7	2.8 (1.6)	0.67	3.1 (1.3)	0.42	3.0 (1.5)	0.51
NS4	Disorderliness	10	4.5 (1.7)	0.27	4.3 (1.8)	0.36	4.4 (1.8)	0.32
NS		34	13.1 (4.7)	0.72	13.3 (4.3)	0.67	13.2 (4.5)	0.70
Harm Avoidance								
HA1	Anticipatory worry	10	3.1 (2.5)	0.74	3.5 (2.2)	0.64	3.3 (2.3)	0.70
HA2	Fear of uncertainty	7	4.2 (1.8)	0.68	4.4 (1.8)	0.63	4.3 (1.8)	0.66
HA3	Shyness with strangers	7	2.6 (2.2)	0.79	2.9 (2.1)	0.72	2.7 (2.2)	0.76
HA4	Fatigability and asthenia	10	3.3 (2.4)	0.71	3.6 (2.4)	0.68	3.4 (2.4)	0.70
HA		34	13.2 (7.1)	0.89	14.3 (6.5)	0.85	13.8 (6.9)	0.87
Reward Dependence								
RD1	Sentimentality	5	4.4 (0.9)	0.37	4.4 (0.9)	0.48	4.4 (0.9)	0.42
RD2	Persistence	9	4.3 (1.4)	0.21	4.9 (1.5)	0.19	4.6 (1.5)	0.23
RD3	Attachment	11	6.9 (2.0)	0.56	6.8 (2.2)	0.62	6.8 (2.1)	0.59
RD4	Dependence	5	2.1 (1.3)	0.40	2.3 (1.2)	0.30	2.2 (1.3)	0.34
RD		30	17.7 (3.2)	0.50	18.4 (3.5)	0.55	18.1 (3.4)	0.52
RD†	RD1 + RD3 + RD4	21	13.4 (2.8)	0.54	13.5 (2.9)	0.56	13.5 (2.8)	0.55

as 0, "rarely" as 1, "sometimes" as 2, and "frequently" as 3. If an item was not applicable to a participant (e.g., questions about aggression events at school during adulthood were not applicable for someone who had an educational level of 6 years only), the participants' final score was weighted by the full score of the sum of the remaining items. A Chinese version of this inventory was developed following the same procedure of two-stage translation, focus group, and field trial. Its psychometric properties, norm data, and construct validity in Taiwanese adults will be detailed in a separate report.

Procedures

Potential participants were approached between 6:30 PM and 8:30 PM immediately after their orientation session at the Health Screening Ward. In a typical day, a team of three research assistants administered both the TPQ and the Brown-Goodwin Aggression Inventory to approximately 10 to 15 subjects. Because the Brown-Goodwin Aggression Inventory must be conducted by interview, while the TPQ can be either self-administered or interviewed, we adopted two sequences for administering the two instruments to optimize the efficiency of the data collection. For literate participants, research assistants first interviewed them using the Brown-Goodwin Aggression Inventory, and then left the TPQ for them to fill out. If a participant was illiterate, a research assistant interviewed him or her using the TPQ first and then the Brown-Goodwin Aggression Inventory. The rationale for the latter sequence was that we learned from pilot study that we would encounter less resistance in answering the Brown-Goodwin Aggression Inventory from illiterate subjects when the TPQ was administered first.

Statistical Analysis

The means, standard deviation, skewness, and kurtosis of each scale and subscale were examined, and Pearson product

moment correlations between scales and subscales were computed. The means of each scale and subscale in males were compared to that in females using Student's *t* test (two-tailed). The effects of age, educational level, and sex on the TPQ scales were assessed using multiple linear regressions.

Both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were employed to test the degree to which the three-factor or the four-factor model describes the TPQ factor structure in our sample. Since the individual items of the TPQ are ordinal and skewed, appropriate methods of both the measure of association and fit function in conducting factor analysis on the basis of individual items require much larger sample size than the current one.²⁵ In a preliminary factor analysis using individual items in each scale, we extracted four factors for the NS (75.3% of the variance), four factors for the HA (81.1% of the variance), and three factors for the RD (65.8% of the variance) using the criteria of eigenvalue greater than 1. However, the pattern of factor loading was not very coherent with the original subscales of the TPQ due to the relatively small sample size of this study. Thus, the following factor analyses were done using the 12 subscales rather than individual items.

In EFA, principal component analysis with promax rotation was performed on the 12 subscales of the TPQ. This procedure increases the stability of the measures by examining the inter-correlations of the subscales rather than individual items, but may hide problems with items that deviate considerably from others on the same subscale. Nevertheless, given the large number of items on the TPQ, this approach appeared most feasible.¹² In CFA, the generalized least squares method was used for estimation with the factor loadings from the exploratory factor analysis as the initial values. Traditionally, the χ^2 test is applied to assess the goodness of fit for a model. However, the χ^2 test is very sensitive to sample size (may reject a

model with good fit if sample size is large) and departures from normality.²⁶ In this study, three other indexes, including goodness-of-fit index (GFI), adjusted GFI (AGFI), and the root of the mean square residual (RMR), were chosen to assess the fit of the model. The GFI reflects the proportion of variance and covariance accounted for by the model in the observed covariance matrix between the TPQ subscales. AGFI adjusts GFI for the number of estimated parameters in the model. Meanwhile, RMR measures the average of the residual correlation between the predicted correlations and the observed ones. Values of GFI greater than 0.9, AGFI greater than 0.8, and RMR less than 0.1 indicate a good fit.²⁷ Both the three-factor model and the four-factor model were examined.

RESULTS

Descriptive statistics

The mean age (\pm SD) of subjects was 44.4 (\pm 13.9) years, ranging from 20 to 69 years, and the mean years of education was 13.6 (\pm 3.9) years, ranging from 0 to 19 years. Young subjects had higher educational levels than older subjects (correlation between age and education, -0.32), and males (14.4 ± 3.2 years) had higher educational levels than females (12.0 ± 4.5 years) ($P < .05$).

The skewness and kurtosis for the three higher-order scales fell in the range of $(-1,1)$, with the exception of RD in males (kurtosis = 1.26). Most of the subscales also had skewness and kurtosis falling in the range $(-1,1)$. The only exception was those of the subscale RD1, in which the distribution of scores was negatively skewed (-1.84 to -1.19) for males, females, and all subjects together, and kurtotic (2.62 to 4.41) for females and all subjects together. Thus, the scales and subscales of the TPQ in this study were approximately normally distributed. The means and standard deviations of the NS, HA, and RD and their subscales were given in Table 1. Comparing the scores of males to that of females, the only significant difference was that females had a higher mean RD2 score than males ($P = .002$). Females also tended to have higher mean scores in the HA and RD

scales than males, although the differences did not reach statistical significance.

The internal consistency (Cronbach's α) of the scales and subscales are also listed in Table 1. The α values of both the NS and HA were greater than 0.60, while that of the RD was not (range, 0.50 to 0.55). The α values of all the subscales of the HA were also greater than 0.60. With regard to the NS, two subscales (NS1 and NS4) had a preponderance of α values below 0.6, ranging from 0.27 to 0.57. For the RD scale, the α values for three subscales (RD1, RD2, and RD4) were low for all of the groups, especially the RD2 subscale (0.19 to 0.23). Notably, the α value of the RD scale was lower than that of one of its subscales (RD3). If RD2 was not counted as part of RD scale, the α values of the of RD \dagger only improved very slightly over that of RD.

To assess the effect of demographic features on the TPQ scores, each higher-order dimension score was regressed on subjects' age, education, and sex (Table 2). NS scores were negatively associated with age, and HA scores were negatively associated with both age and educational levels. The differences between sexes on the scores of the three scales were not significant after adjusting for age and educational level.

Internal Structure

Because males and females did not differ in the scoring of the TPQ scales in this study, and the number of subjects of each sex was not sufficient for factor analysis, all subsequent analyses were done with all subjects grouped together. The correlations among the three scales of the TPQ were very low, none of them being significant (Table 3). In subscale-scale correlations, the highest positive values were within the proper scale. The intersubscale correlations within the equivalent scale were positive and moderately high, except for the RD

Table 2. Linear Regression Coefficients of the TPQ Scales on Age, Education Level, and Sex (N = 201)

	NS	HA	RD	RD*	PS
	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)
Age	-0.12 (0.02)†	-0.13 (0.04)†	0.01 (0.02)	-0.01 (0.02)	0.02 (0.01)†
Education	-0.07 (0.09)	-0.41 (0.13)†	0.05 (0.07)	0.01 (0.06)	0.04 (0.07)
Sex	-0.03 (0.65)	-0.06 (0.97)	-0.68 (0.50)	0.06 (0.42)	-0.75 (0.22)†

Abbreviation: PS, persistence.

*RD does not include the RD2 subscale.

†Significant at $P < .05$.

Table 3. TPQ Interscale Correlation Matrix (N = 201)

Scale	NS	HA	RD	RD*	NS1	NS2	NS3	NS4	HA1	HA2	HA3	HA4	RD1	RD2	RD3RD4	
NS	—															
HA	-.04	—														
RD	-.03	-.11	—													
RD†	.05	-.11	.90†	—												
NS1	.61†	-.33†	.08	.13	—											
NS2	.66†	.13	-.14	-.11	.09	—										
NS3	.65†	.07	.08	.15†	.23†	.29†	—									
NS4	.74†	.04	-.08	.00	.28†	.36†	.32†	—								
HA1	.04	.82†	-.11	-.10	-.25†	.17†	.08	.10	—							
HA2	-.25†	.73†	-.08	-.06	-.40†	-.09	-.09	-.06	.46†	—						
HA3	.05	.80†	-.14	-.13	-.22†	.17†	.10	.09	.56†	.53†	—					
HA4	-.01	.78†	-.02	-.05	-.20†	.12	.10	-.02	.52†	.42†	.45†	—				
RD1	-.14	-.22†	.50†	.48†	-.03	-.23†	-.09	-.01	-.23†	-.04	-.20†	-.21†	—			
RD2	-.17†	-.04	.56†	.13	-.06	-.11	-.12	-.17†	-.05	-.06	-.06	.04	.21†	—		
RD3	.10	-.19†	.74†	.83†	.19†	-.10	.17†	.03	-.16†	-.17†	-.21†	-.08	.19†	.10	—	
RD4	.04	.24†	.41†	.49†	-.02	.08	.12	-.06	.21†	.18†	.20†	.17†	.06	-.01	.03	—

*RD does not include the RD2 subscale.

†Significant at $P < .05$.

subscales. The intersubscale correlations across the scales were either less positive or negative. The only highly negative correlations could be found between the NS1 subscale and all the HA subscales, and between the RD1 subscale and the HA1, HA3, and HA4 subscales. However, the correlations between the RD4 subscale and all of the HA subscales were modestly positive.

Factor Analysis

The standardized factor loadings following promax rotation are given in Table 4. Scree plots indicated that the three-factor solutions were appropriate. The proportion of the sample variance

explained by the three factors was 51.5%. The first factor loaded positively on all HA subscales and on the RD4 subscale, and negatively on the NS1 subscale. The second factor loaded positively on all NS subscales, and the third factor loaded positively on all RD subscales. On the basis of communality (h^2), the RD2 and RD4 subscales were not sufficiently saturated by any of the three factors.

Given that Cloninger has suggested that the RD2 (persistence) subscale of the TPQ may be a factor itself, four-factor solutions were also attempted. The proportion of the sample variance explained by the four factors improved to 59.3%. The pattern of loadings of the first three factors were very

Table 4. Factor Loadings and Communality of TPQ Subscales (promax rotation, N = 201)

Subscale	Three-Factor Solution				Four-Factor Solution				
	I	II	III	h^2	I	II	III	IV	h^2
NS1	-0.42	0.55		0.50	-0.44	0.54			0.50
NS2		0.61		0.49		0.65			0.55
NS3		0.73		0.57		0.72			0.57
NS4		0.69		0.49		0.68			0.49
HA1	0.78			0.65	0.79				0.65
HA2	0.75			0.62	0.74				0.67
HA3	0.77			0.65	0.78				0.65
HA4	0.75			0.54	0.77				0.61
RD1			0.64	0.47			0.62		0.51
RD2			0.50	0.30				0.92	0.86
RD3			0.66	0.52			0.60		0.52
RD4	0.49		0.45	0.37	0.41		0.60		0.54
% of variance	24.03	16.64	10.79		24.03	16.64	10.79	7.88	

NOTE. Only loadings with absolute values >0.30 are shown.

Table 5. Factor Loadings and Squared Multiple Correlations From Confirmatory Factor Analysis of TPQ Subscales (N = 201)

Subscale	Three-Factor Model				Four-Factor Model				
	I	II	III	SMC	I	II	III	IV	SMC
NS1		0.11		0.01		0.29			0.05
NS2		0.53		0.29		0.50			0.25
NS3		0.42		0.18		0.44			0.20
NS4		0.73		0.53		0.78			0.61
HA1	0.69			0.47	0.68				0.46
HA2	0.53			0.28	0.53				0.28
HA3	0.70			0.49	0.69				0.47
HA4	0.58			0.34	0.59				0.34
RD1			0.86	0.74			0.86		0.73
RD2			0.29	0.08				0.98	0.96
RD3			0.23	0.05			0.23		0.05
RD4	0.49		0.11	0.01	0.41		0.11		0.01

Abbreviation: SMC, squared multiple correlation.

similar to that of three-factor model, with the exception that the third factor no longer loaded on the RD2 subscale, and the fourth factor loaded solely on the RD2 subscale. The four-factor solution also improved the communality of both RD2 and RD4.

The results of CFA for the three-factor model showed that $\chi^2 = 112.17$, $df = 51$, $P < .001$, GFI = 0.912, AGFI = 0.858, and RMR = 0.160. Although the values of GFI and AGFI met the criteria of a good fit, the value of RMR suggested that some subscales might be poorly fit. The factor loadings for the 12 subscales ranged from 0.11 to 0.86 (Table 5), and with the exception of NS1 and RD4, all of the factor loadings were significant at $P < .05$. When the CFA was performed on the basis of a four-factor solution, the results showed that $\chi^2 = 105.21$, $df = 48$, $P < .001$, GFI = 0.912, AGFI = 0.858, and RMR = 0.156. These results were very similar to those of the three-factor model, with the value of RMR still not meeting the criteria of less than 0.10. The main improvements in factor loadings for the 12 subscales were in NS1 (from 0.11 to 0.29) and RD2 (from 0.29 to 0.98). However, the loading of RD4 remained low and the only one that was not significant at $P < .05$.

Correlation With Aggression

Among the three TPQ scales, only the NS scale scores correlated consistently with aggression scores in every period (childhood, adolescence, and adulthood) (Table 6). Scores on the other scales did not correlate with aggression at any period.

DISCUSSION

Psychometric Properties

According to Cloninger’s theory,¹ genetic variation in each dimension follows a normal distribution, with most people having intermediate values. However, the observed behaviors are influenced by other effects, such as environment, and therefore may not be normally distributed. In our sample, only the RD scale and the RD1 subscale were negatively skewed. This result was similar to that of the US normative data,¹⁰ in which the distribution of RD1 scores was negatively skewed (−1.70 to −1.46) and kurtotic (1.83 to 3.23) in white females and both black females and males. One likely explanation is that the RD1 (sentimentality) subscale is very short (five items) and had very

Table 6. Correlation Between Scores on the TPQ Scales and Those on the Brown-Goodwin Aggression Inventory (N = 201)

Aggression in	NS	HA	RD	RD*	PS
Childhood	0.21†	0.11	0.01	0.01	0.00
Adolescence	0.32†	0.10	0.09	0.15	−0.08
Adulthood	0.19†	0.09	−0.04	0.01	−0.11
Lifetime	0.29†	0.13	0.01	0.06	−0.09

*RD does not include the RD2 subscale.

†Significant at $P < .05$.

high endorsement rates ($> 72\%$ for the Taiwanese data and $> 75\%$ for the US normative data).

Our results showed that the HA scale was the most consistent, followed by the NS, and the RD. Many previous studies have also reported that the internal consistency is less satisfactory for the RD scale than the other two.^{10,12,14,17,18,20,21,27} The only exception is a study of Japanese university students, in which the internal consistency of the RD scale was higher than that of the NS scale.¹⁵ However, the number of items and the scoring method of that study were based on an earlier version of the TPQ and could not be compared directly with the other studies. The low internal consistency of the three RD subscales may be in part attributable to the fact that two of the three subscales (RD1 and RD4) have only five items each.¹⁰ However, the α of the RD2 subscale was still low in our sample, even though it contains nine items. This indicates that the composition of the RD scales is probably questionable.

Structural Validity

An EFA of the TPQ subscales restricted to a three-factor solution conforms to Cloninger's hypothesis of three independent dimensions except for the deviations on some RD subscales. Factor I was defined mainly by the four HA subscales, and factor II specifically loaded the four NS subscales. Meanwhile, factor III was defined by the RD1, RD2, and RD3 subscales. However, the RD4 subscale could not be assigned to factor III because it also showed almost equivalent correlation with factor I. Furthermore, both the RD2 and RD4 subscales were insufficiently saturated by any of the three factors. Thus, factor III is problematic and represented RD dimension only to some extent.

The EFA results of a four-factor solution of the TPQ subscales support the view that the RD2 subscale is an independent scale itself. In the four-factor-solution, the loadings of factors I and II did not change much compared to those of the three-factor solution. Meanwhile, the RD4 subscale could be assigned to factor III, because it had a relatively higher loading on factor III than on factor I, and the new factor IV is exclusively defined by a high loading of the RD2 subscale. Furthermore, the communality of both RD2 and RD4 improved substantially as compared to those in the three-factor solution. However, the results of CFA indicated that the fit of the three-factor model and

that of the four-factor model are equivalent and both only partially meet the criteria of a good fit (their values of GFI and AGFI indicate a good fit, while that of RMR does not). Thus, from the psychometric point of view, the TPQ subscales can be composed of four factors rather than the original proposition of three factors. This corroborates the findings of Svrakic et al.,¹⁶ Weyers et al.,²⁰ Krebs et al.,¹⁹ and Brandstrom et al.²¹ However, several other factor-analytic studies of the TPQ failed to extract the RD2 subscale as the fourth factor, despite their findings that the homogeneity of the RD dimension is smaller compared to the NS and HA dimensions.^{10,14} This inconsistency is complicated by the fact that the internal consistency of the RD subscales tend to be less than ideal and hence make the interpretation of the results more difficult. It is worthwhile to note that Giancola et al.¹³ applied a new version of the TPQ (160 items) that was designed to improve the psychometric properties of the scales but still failed to support the existence of the fourth factor.

Despite the controversy over the existence of the fourth factor, the results of factor analyses over a wide range of study populations, including the present one, shared two strikingly similar features: the NS1 subscale tends to have a modest negative loading on the factor representing the HA dimension, while the RD4 subscale tends to have a modest positive loading on the HA factor.^{10,14,16,18-21} This may reflect Cloninger's original proposition that HA inhibit NS, especially those of excitable behaviors (NS1), and enhance RD, especially those of dependence (RD4).

Construct Validity

In addition to the assessment in the structural validity of the TPQ in Taiwanese adults, several predictions made on the basis of Cloninger's bio-social theory of personality were tested in this study to examine the construct validity of the TPQ. First, it is predicted that females have lower NS scores but higher HA and RD scores than males, which has been corroborated by Cloninger et al.¹⁰ Many other studies have found gender differences in HA and RD but not in NS.^{13,14,18,19,27} One study even reported that females had higher NS scores than males.²⁰ On the basis of these previous studies, it seems that the gender difference in HA and RD is more culturally independent, while NS is not.

In this study, we did not find any gender differences in the higher-order scales. It is of interest, then, to ponder why Taiwanese women do not have higher HA and RD scores than men as the other populations have shown. Comparing the mean of the HA in Taiwanese (13.2 for men and 14.3 for women) with those of the other populations, i.e., the United States (10.6 to 11.6 for men and 12.9 to 13.2 for women),¹³ England (10.7 for men and 15.0 for women),¹⁷ Germany (14.4 to 14.7 for men and 15.8 to 18.0 for women),¹⁹ and Czechoslovakia (14.7 for men and 17.7 for women),¹⁸ the mean score of Taiwanese women is in the lower end while that of men is in the higher end. A similar but less prominent trend is found for the RD scale, i.e., the mean RD score of Taiwanese women is in the lower end while that of men is in the higher end. Two factors may lead to these patterns in Taiwanese. The first is that the prevailing Confucian culture may shape both Taiwanese men and women to be more harm-avoidant and reward-dependent as compared with other populations. The second is that Taiwanese women have been encouraged or even forced to participate in the working force, both in the earlier agricultural period and in the current industrial one. This may render Taiwanese women to be less harm-avoidant and reward-dependent as would be expected in a traditional male-dominant society. The interaction of these two forces then leads to the closing of the gender gap in the scores of HA and RD in Taiwanese.

Second, it is predicted that as people grow older, their scores become lower in NS and RD and higher in HA. In this study, we found that age is negatively associated with NS and HA, but not with RD. The finding that NS lowers with age is consistent with many other studies that have also examined the effect of age on NS.^{10,14,18-21} However, our finding that HA also lowers with age was contrary to the prediction and was not reported in other studies. Meanwhile, two studies have found that RD also declines with age.^{19,20} The different relationship between age and HA as well as RD in Taiwanese (HA lowers with age, but RD does not) may also be accounted for by the prevailing Confucian culture. The Taiwanese tend to show more respect for people as they grow older, and the majority of the elderly are living with and being taken care of by their offspring. Such cultural characteristics may lead to a decline of HA and no

increase of RD in Taiwanese people as they age. Thus, it is fair to conclude that the decline of NS with age is more culturally universal, while the relationship between age and the other two scales is not.

Third, we found that educational level is negatively associated with HA. Only one study in the United States¹⁰ has explicitly examined this issue; education was positively correlated with NS for both white men and women, and weakly negatively correlated with white and black women. No clear trend can be obtained from these two studies.

Fourth, it is believed that HA is weakly but negatively correlated with NS and weakly but positively correlated with RD. Many studies have found a slight negative association between NS and HA in females^{16,27} or in both males and females,^{10,13,14} but small to no correlation between NS and RD. However, the current study did not find such a correlation between HA and NS. Instead, the modestly negative correlation between the NS1 subscale and all of the HA subscales in our sample indicated that in Taiwanese adults only the excitable behaviors (NS1) of NS might be inhibited by harm avoidance, while the other NS subscales are not.

Finally, our result that NS is positively associated with aggression is consistent with the finding that violent crimes in adulthood were associated with high NS in childhood, in which subjects' personality was rated on the basis of existing records rather than by self-report.⁹ While Sigvardsson et al.⁹ also found that low HA and low RD in childhood were associated with adult violence, we did not find HA and RD to be negatively correlated with aggression. Two caveats should be pointed out in comparing these two studies. First, expression of aggression in Chinese society might be different from that in Western society, and second, whether the scale used in this study (i.e., Brown-Goodwin Aggression Inventory) has sufficient construct validity warrants further investigation. On the other hand, the violent crimes in Sigvardsson et al.⁹ were ascertained through a centralized criminal register and hence are likely to be more severe in nature than the aggression measured in our subjects. This may indicate that while high NS is associated with higher aggression, extremely low HA and RD are associated with violent crimes.

CONCLUSION

Our findings indicate that the internal consistency of the Chinese version of the TPQ scales is mostly in the acceptable range except for the RD scale and its subscales. The results of factor analysis of the 12 TPQ subscales partially support the four-factor model rather than the original three-factor model. The findings that NS is negatively correlated with age, that the NS1 subscale is slightly negatively correlated with all of the HA subscales, and that the NS is positively correlated with aggression lend further support to the construct validity of the NS and HA dimensions. Fur-

ther research is needed to refine the RD scale of the TPQ.

ACKNOWLEDGMENT

The authors would like to thank Dr. C. R. Cloninger for granting Dr. Suo-Jeng Wang permission to translate the TPQ into Chinese, Dr. Tom Przybeck for providing us with the scoring keys of the TPQ, Dr. Lawrence Greenhill for granting us permission to translate the Brown-Goodwin Aggression Inventory (Revised Version) into Chinese, and Dr. Kuo-Liong Chien and the staff at the Health Screening Ward at National Taiwan University Hospital for helping us coordinate the recruitment of subjects. We also wish to acknowledge the excellent performance of our research assistants, Chao-Jung Chen, Shu-Tzu Pan, and Heuy-Lien Tsai.

REFERENCES

1. Cloninger CR. A unified biosocial theory of personality and its role in the development of anxiety states. *Psychiatr Dev* 1986;3:167-226.
2. Digman JM. Personality structure: emergence of the five factor model. *Ann Rev Psychol* 1990;41:417-440.
3. Eysenck HJ. *The Biological Basis of Personality*. Springfield, IL: C.C. Thomas, 1967.
4. Eysenck HJ. Dimensions of personality: 16, 5 or 3?—criteria for a taxonomic paradigm. *Pers Individ Differ* 1991;12:773-790.
5. Gray JA. *The Neuropsychology of Anxiety*. New York, NY: Plenum, 1982.
6. McCrae RR, Costa PT. Validation of the five-factor model of personality across instruments and observers. *J Pers Soc Psychol* 1987;56:586-595.
7. Cloninger CR, Adolphson R, Svrakic NM. Mapping genes for human personality. *Nat Genet* 1996;12:3-4.
8. Cloninger CR. A systematic method for clinical description and classification of personality variants: a proposal. *Arch Gen Psychiatry* 1987;44:573-588.
9. Sigvardsson S, Bohman M, Cloninger CR. Structure and stability of childhood personality: prediction of later social adjustment. *J Child Psychol Psychiatry* 1987;28:929-946.
10. Cloninger CR, Przybeck TR, Svrakic DM. The tridimensional personality questionnaire: U.S. normative data. *Psychol Rep* 1991;69:1047-1057.
11. Bagby RM, Parker JDA, Joffe RT. Confirmatory factor analysis of the Tridimensional Personality Questionnaire. *Pers Individ Differ* 1992;11:1245-1246.
12. Earleywine M, Finn PR, Peterson JB, Pihl RO. Factor structure and correlates of the tridimensional personality questionnaire. *J Stud Alcohol* 1992;53:233-238.
13. Giancola PR, Zeichner A, Newbolt WH, Stennett RB. Construct validity of the dimensions of Cloninger's tridimensional personality questionnaire. *Pers Individ Differ* 1994;17:627-636.
14. Otter C, Huber J, Bonner A. Cloninger's tridimensional personality questionnaire: reliability in an English sample. *Pers Individ Differ* 1995,18:471-480.
15. Takeuchi M, Yoshino A, Kato M, Ono Y, Kitamura T. Reliability and validity of the Japanese version of the Tridimensional Personality Questionnaire among university students. *Compr Psychiatry* 1993;34:273-279.
16. Svrakic DM, Przybeck TR, Cloninger CR. Further contribution to the conceptual validity of the unified biosocial model of personality: US and Yugoslav data. *Compr Psychiatry* 1991;32:195-209.
17. Waller NG, Lilienfeld SO, Tellegen A, Lykken DT. The Tridimensional Personality Questionnaire: structural validity and comparison with the multidimensional personality questionnaire. *Multivariate Behav Res* 1991;26:1-23.
18. Kozeny J, Kubicka L, Prochazkova Z. Psychometric properties of the Czech version of Cloninger's three-dimensional personality questionnaire. *Pers Individ Differ* 1989;10:1253-1259.
19. Krebs H, Weyers P, Janke W. Validation of the German version of Cloninger's TPQ: replication and correlations with stress coping, mood measures and drug use. *Pers Individ Differ* 1998;24:805-814.
20. Weyers P, Krebs H, Janke W. Reliability and construct validity of the German version of Cloninger's Tridimensional Personality Questionnaire. *Pers Individ Differ* 1995;19:853-861.
21. Brandstrom S, Schlette P, Przybeck TR, Lundberg M, Forsgren T, Sigvardsson S, et al. Swedish normative data on personality using the Temperament and Character Inventory. *Compr Psychiatry* 1998;39:122-128.
22. Wang SJ. Personality features of alcoholic patients (pre-doctoral independent study). Department of Psychology, National Taiwan University, 1990.
23. Brown GL, Goodwin FK, Ballenger JC, Goyer PF, Major LF. Aggression in human correlates with cerebrospinal fluid amine metabolites. *Psychiatry Res* 1979;1:131-139.
24. Limson R, Goldman D, Roy A, Lamparski D, Ravitz B, Adinoff B, et al. Personality and cerebrospinal fluid monoamine metabolites in alcoholics and controls. *Arch Gen Psychiatry* 1991;48:437-441.
25. Muthén B, Kaplan D. A comparison of some methodologies for the factor analysis of non-normal Likert variables: a note on the size of the model. *Br J Math Stat Psychol* 1992;45:19-30.
26. Cole DA. Utility of confirmatory factor analysis in test validation research. *J Consult Clin Psychol* 1987;55:584-594.
27. Nixon SJ, Parsons OA. Cloninger's tridimensional theory of personality: construct validity in a sample of college students. *Pers Individ Differ* 1989;10:1261-1267.