The Reliability and Validity of the Chinese Version of Abbreviated PAD Emotion Scales

Xiaoming Li[1, 2], Haotian Zhou[1, 2], Shengzun Song[1, 2], Tian Ran[1, 2], Xiaolan Fu[1]

1 State Key Laboratory of Brain and Cognitive Science, Institute of Psychology,
Chinese Academy of Sciences, Beijing 100101, China
lixm@psych.ac.cn, zhou-ht@mails.tsinghua.edu.cn,
songshenzun@163.com, rant@psych.ac.cn, fuxl@psych.ac.cn
2 Graduate School, Chinese Academy of Sciences, Beijing 100039, China

Abstract. The study aimed at testing the reliability and validity of the Chinese version of Abbreviated PAD Emotion Scales using a Chinese sample. 297 Chinese undergraduate students were tested with the Chinese version of Abbreviated PAD Emotion Scales; 98 of them were retested with the same scales after seven days in order to assess the test-retest reliability; and 102 of them were tested with SCL-90 at the same time which was intended as criteria for validity to assess the criterion validity. The results showed that the Chinese version of Abbreviated PAD Emotion Scales displayed satisfying reliability and validity on P (pleasure-displeasure), only moderate reliability and validity on D (dominance-submissiveness), but quite low reliability and validity on A (arousal-nonarousal).

1 Introduction

Nowadays affective computing has become the hotspot in computer science, and effective methods for annotating emotion have been regarded as of particular importance to the success of affective computing. Various methods have been used to annotate emotion, such as categorical annotation scheme [1] or 2 dimensional annotation scheme [2]. The problem with most of these tools in a view of affective computing was that they did not meet the requirement of psychological measurement for reliability and validity. The development of a more precise emotion annotation tool put forward based on many demonstration researches in psychology should be emphasized for the sake of affective computing.

The PAD Emotion Scales was build upon the PAD Emotional State Model [3] , It was an elaborate tool for measuring emotions developed by Dr Mehrabian of UCLA and had been used in many practical researches [4, 5, 6, 7]. There are three nearly independent dimensions that are used to describe and measure emotional states in the PAD emotional-state Model: pleasure-displeasure (P), arousal-nonarousal (A), and dominance-submissiveness (D). "pleasure-displeasure" distinguishes the positive-negative quality of emotional states, "arousal-nonarousal" refers to a combination of physical activity and mental alertness, and "dominance-submissiveness" is defined in

terms of control versus lack of control [3]. Specific emotional states can be visualized as points in a three-dimension PAD emotion space. When the PAD scale scores are standardized, each emotion can be described using corresponding values on the pleasure-displeasure, arousal-nonarousal, and dominance-submissiveness coordinates. For example, "anger" can be represented with (-.51, .59, .25) (scores on each PAD scale ranging from -1 to +1), which indicates that it is a highly unpleasant, highly aroused, and moderately dominant emotional state.

Although The full-length (34-item) PAD Emotion Scales was strongly recommended for most experimental applications [3], it was realized that researchers sometimes need to require participants rating a large number of stimuli and completing such a lengthy questionnaire for many times, and it could be an extremely irritating experience to the participant which might jeopardize the credibility of data obtained. Therefore, a 12-item abbreviated version had been prepared with each dimension consisting of four items.

Each item of the PAD Emotion Scales consists of a word pair that is separated by nine spaces. The two words on each line refer to feelings and highlight a special contrast between the two feelings. For items measuring P, the two words in each item differ only on P (but the same or similar on A and D), and the principle is the same with items for A and D. Participants are required to indicate which end of the scale is heavier or stronger (and by how much) as an accurate description of your feelings by placing a check-mark in one of the nine spaces. Check-marks in the left-most to the right-most spaces are transformed to scores that range from -4 to +4, and the check-mark in the middle space is coded as zero.

Half of the items on each dimension are inverted (worded negatively) to minimize response bias and to render the scale more opaque (i.e., less obvious) to respondents. In addition, the PAD Emotion Scales provides a set of equations for computing from the PAD scores the values of some specific emotion, such as exuberant, bored, anxious, fearful and so on, which can be weighed against each other to determine the dominant emotion under a given circumstances. The PAD Emotion Scales is a versatile psychological measuring instrument in accordance with the dimensional and categorical approach to emotion classification and is capable of adapting to a variety of applications including emotion annotation.

The primary goal of the present study was to test the reliability and validity of the Chinese version of Abbreviated PAD Emotion Scales in a Chinese sample.

2 Methods

2.1 Participants

297 (male 132; female 165) undergraduate students in Beijing and Hebei were tested with the Chinese version of Abbreviated PAD Emotion Scales. 98 of them (male 53; female 45) were retested with the same scales after seven days in order to assess the test-retest reliability. To assess the criterion validity, 102 of them (male 57; female 45)

were tested with SCL-90 at the same time which was intended as criteria for validity test.

2.2 Testing tools

(1) The Chinese version of Abbreviated PAD Emotion Scales. Firstly, a prototype of the Chinese version of Abbreviated PAD Emotion Scales was obtained by translating the original English version through the collaboration of one psychology professional and three graduate students of psychology major. Then this prototype was back-translated into English by an expert in English to be compared with the original English version. The final version used in this study was formed through appropriate subsequent modification; (2) SCL –90. SCL –90 was intended as criteria for validity test which has eleven statistic indexes including Somatization, Obsessive-Compulsive, Interpersonal sensitivity, Depression, Anxiety, Hostility, Photic anxiety, Paranoid ideation, Psychoticism, Others, and GSI (General Symptomatic Index).

3 Results

3.1 Reliability Analysis

Cronbach's was computed to gauge the inter-item consistencies for P, A, and D. Correlation analysis was used to analyzing the test-retest reliabilities for P, A, and D. The detailed results were displayed in Table1.

Table 1. The Inter-item Consistencies and Test-retest Reliabilities for P, A, and D

Dimension	the Inter-item Consistency (Cronbach's)	Test-retest Reliability
P pleasure-displeasure	.692	.427**
A arousal-nonarousal	.235	.380**
D dominance-submissiveness	.467	.565**

Note: **means *P*< 0.01

3.2 Validity Analysis

3.2.1 Construct-related Validity

The construct-related validity was assessed using exploratory factor analysis by Vari-max – rotated method, and the detailed results were displayed in Table 2. The three

factors accounted for 22.47%, 12.90%, and 10.58% of the total variance respectively. All 4 items of P had their highest loadings on factor 1(P); three items of D and one item of A (item2) had their highest loadings on factor 2(D); and three items of A and one item of D (item9) had their highest loadings on factor3 (A).

Table 2. Exploratory Factor Analysis for 12 items

Item	Factor		
	1 P	2 D	3 A
10	**_.758_**	.192	.100
1	**_.747_**	.207	.130
7	**_.698_**	.213	.004
4	**_.545_**	-.019	-.214
6	-.002	**_.726_**	.066
2	.146	**_.628_**	-.071
3	.212	**_.621_**	-.143
12	.187	**_.554_**	.110
5	.035	-.136	**_.713_**
11	-.094	.039	**_.634_**
9	-.284	.387	**_.502_**
8	.244	-.012	**_.400_**

3.2 .2 Criterion-related Validity

SCL—90 served as criteria for criterion-related validity test in the present study, and correlation analysis was used to test the criterion-related validity. The detailed results were displayed in Table3.

Table 3. Coefficients of Correlation between P, A, D and SCL-90

	P	A	D
Somatization	-.343**	-.093	-.298**
Obsessive-Compulsive	-.312**	-.162	-.250*
Interpersonal sensitivity	-.380**	-.177	-.348**
Depression	-.448**	-.205*	-.354**
Anxiety	-.416**	-.093	-.245*
Hostility	-.419**	-.165	-.172
Photic anxiety	-.363**	-.181	-.276**
Paranoid ideation	-.461**	-.271**	-.203*
Psychoticism	-.282**	-.188	-.277**
Others	-.398**	-.055	-.303**
GSI	-.484**	-.203*	-.359**

Note: *means $P < 0.05$ level; **means $P < 0.01$ level

4 Discussion

For the inter-item consistency, the Cronbach's for P and D met the conventional requirement for the inter-item consistency of common psychological measurement tools, whereas the Cronbach's for A was not good enough. The results differed drastically from those of Mehrabian's study in which Alpha internal consistency coefficients for P, A, and D were 0.95, 0.83, and 0.78 respectively [3]. As for the test-retest reliability, the coefficients of correlation of P, A, and D between the two sessions with a time interval of seven days were all significant at .01level. Considering the fact that emotion is more variable than personality and intelligence, the test-retest reliabilities for P, A, and D were acceptable.

The results of exploratory factor analysis for the construct-related validity illustrated that one item of A (item2: Wide-awaken----Sleepy) were factored into factor2 (D), and one item of D (item9: Humble ----Superior) were factored into factor3 (A) which were contrary to expectation. As mentioned above, for any item of a particular dimension, two kinds of emotions described by the two words are of similar values on the other two dimensions. Since the PAD scores for Sleepy is (.20,-.70,-.44), the value of D for Wide-awaken should also be negative (submissive). In the same way, the terms "Humble" and "Superior" should be similar in A. The unexpected outcomes may be caused by the distortion in meanings resulted from the translation. The corresponding Chinese translation ("qingxingde") for Wide-awaken might have been considered to be more dominant than its counterpart for Sleepy("kunjuande") by Chinese participants. Similarly, the Chinese terms for "Humble" ("qianbeide") and "Superior" ("gaoaode") might also have been rated differently on A with "Superior" judged to be more aroused.

SCL-90 is a 90-item self-report symptom inventory designed primarily to reflect the psychological symptom patterns of psychiatric and clinical patients. All of the eleven indexes of SCL-90 represent relatively negative emotions. The results of the correlation analysis between P and SCL-90 showed that P had significantly negative correlations with all the eleven statistic indexes of SCL-90, meaning that the criterion-related validity for P was good. D was found to be significantly negatively correlated with all the indexes of SCL-90 but Hostility. Therefore, D was moderately valid in terms of criterion-related validity. However, A had poor criterion-related validity indicated by the absence of statistically significant correlation between A and SCL-90.

5 Conclusion

In conclusion, the Chinese version of Abbreviated PAD Emotion Scales had satisfying reliability and validity on P, only moderate reliability and validity on D, but quite low reliability and validity on A. The disparity between the present study and the study of Mehrabian might be caused by the difference in testing methods selected. According to Mehrabian, the most desirable way to check the PAD Emotion Scales was to administer the questionnaire under as many circumstances as possible. A systematic list of 80 such circumstances had been provided in [5] and [7]. These descrip-

tions constituted a balanced representation of a large variety of everyday settings, and included 10 replications of eight categories of PAD combinations (+P +A +D, +P +A–D, +P –A +D, +P –A –D, –P +A +D, –P +A –D, –P –A +D, –P –A –D). Further work needs to be done before any credible conclusion can be drawn about the applicability of the Chinese version of Abbreviated PAD Emotion Scales.

References

1. Craggs, R., Wood, M.M.: A 2 dimensional annotation scheme for emotion in dialogue. In: Proceedings of AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications, Stanford University (2004)
2. Craggs, R., Wood, M.M.: A categorical annotation scheme for emotion in the linguistic content of dialogue. In: Elisabeth André, Laila Dybkjaer, Wolfgang Minker and Paul Heisterkamp (eds.): Affective Dialogue System, Vol. 3068. Springer. (2004) 89-100
3. Mehrabian, A.: Framework for a comprehensive description and measurement of emotional states. Genetic, Social, and General Psychology Monographs. 121 (1995) 339-361
4. Mehrabian, A.: Comparison of the PAD and PANAS as models for describing emotions and for differentiating anxiety from depression. Journal of Psychopathology and Behavioral Assessment. 19 (1997) 331-357
5. Mehrabian, A.: Correlations of the PAD emotion scales with self-reported satisfaction in marriage and work. Genetic, Social, and General Psychology Monographs. 124 (1998) 311-334
6. Mehrabian, A., Blum, J.S.: Physical appearance, attractiveness, and the mediating role of emotions. Current Psychology: Developmental, Learning, Personality, Social. 16 (1997) 20-42
7. Mehrabian, A., Wihardja, C., Ljunggren, E.: Emotional correlates of preferences for situation-activity combinations in everyday life. Genetic, Social, and General Psychology Monographs. 123 (1997) 461-477