Educational Stress Scale for Adolescents:

Development, Validity, and Reliability with Chinese Students

Jiandong Sun, Michael P. Dunne, and Xiang-yu Hou

Queensland University of Technology (QUT)

Ai-qiang Xu

Shandong Provincial Centre for Disease Control and Prevention (CDC)

Author Note

This study was sponsored by a Research Development Grant of Queensland University of Technology (QUT) and a QUT Grant-in-Aid. This work also gained in-kind support of Shandong Provincial Centre for Disease Control and Prevention (CDC) and Shouguang CDC. The authors thank all participating students and schools. Assistance with statistical analysis was provided by Dr. Jason Dixon and Dr. Cameron Hurst of QUT School of Public Health, and Dr. Cameron Newton from QUT Business School. The authors would like to thank Dr. Rebecca P. Ang of Nanyang Technological University, Singapore for her kind permission to use the AESI and the adoptions of AESI scales in our study.

Correspondence concerning this paper should be addressed to Prof. Michael P. Dunne, School of Public Health, Queensland University of Technology (QUT), Victoria Park Rd, Kelvin Grove, QLD 4059, Australia. E-mail: m.dunne@qut.edu.au

Full Word Count: 6400 (all texts, tables and references)

Abstract

This paper describes the development and initial validation of a new instrument to measure academic stress -- the Educational Stress Scale for Adolescents (ESSA). A series of cross-sectional questionnaire surveys were conducted with more than 2000 Chinese adolescents to examine the psychometric properties. The final 16-item ESSA contains five latent variables: Pressure from study, Workload, Worry about grades, Self-expectation, and Despondency, which together explain 64% of the total item variance. Scale scores showed adequate internal consistency, 2-week test-retest reliability, and satisfactory concurrent validity. A confirmatory factor analysis suggested the proposed factor model fits well in a different sample. For researchers who have a particular interest in academic stress among adolescents, the ESSA promises to be a useful tool.

Keywords: ESSA, validity, reliability, academic stress, Chinese adolescents

Educational Stress Scale for Adolescents:
Development, Validity, and Reliability with Chinese Students

Academic learning is among the most important sources of stress among young students worldwide and appears to be quite severe in Asian countries (Brown, Teufel, Birch, & Kancherla, 2006; Christie & MacMullin, 1998; Dodds & Lin, 1992; Gallagher & Millar, 1996; Huan, See, Ang, & Har, 2008; Tang & Westwood, 2007). Asian students usually have high academic burden (Lee & Larson, 2000), low satisfaction regarding their academic performance, and high expectations (Crystal *et al.*, 1994), and may suffer more academic stress (Ang & Huan, 2006a; Ang, Huan, & Braman, 2007) than their counterparts in English speaking countries. Academic stress is a significant contributor to a variety of mental and behavioral disorders, such as depression, anxiety and suicidal behavior (Ang & Huan, 2006b; Bjorkman, 2007).

In China, there has been a growing recognition of academic burden and its health impact among students as a public health and educational concern. A national survey conducted with 5040 adolescents and 6552 parents by the All-China Women's Federation (2008) reported that nearly half (49.1%) of the students in secondary schools spend at least 2 hours per day for homework assigned by their teachers. Another national survey (China Youth Social Service Center, 2008) found that most children and adolescents (66.7%) considered academic pressure as the biggest stress in their lives. Academic related factors, such as underachievement, pressure from transitional examinations and study workload are associated with poor mental health among Chinese adolescents (Li & Zhang, 2008; Liu & Tein, 2005; Zhang, Tao, & Zeng, 2001). High academic pressure may also lead to physical violence and many developmental problems (Lin & Chen, 1995).

A number of self-report instruments have been developed to assess the level of academic stress and associations with health problems among adolescents. These include the

Academic Stress Questionnaire (ASQ, Abouserie, 1994), Student Stress Inventory (SSI, Zeidner, 1992), Academic Stress Scale (ASS, Kohn & Frazer, 1986), Lakaev Academic Stress Response Scale (LASRS, Lakaev, 2009), Student-life Stress Inventory (SSI, Gadzella, 2001), High School Stressor Scale (HSSS, Burnett & Fanshawe, 1997), Academic Expectation Stress Inventory (AESI, Ang & Huan, 2006a), and Survey of Academic Stress (SAS, Bjorkman, 2007). Most of these scales were designed and used to measure academic stress among college or university students and only three have been used in surveys with secondary school students. All but the AESI (Ang & Huan, 2006a, 2006b) were developed and validated in western countries.

Among the three instruments used in secondary school settings, the HSSS (Burnett & Fanshawe, 1997) was developed with a sample of Australian students (year 8 through 12). It includes 35 items and 9 latent variables. One problem for this scale is the psychometric properties are less than satisfactory. For example, the GFI (.85) and AGFI (.82) based on the Confirmatory Factor Analysis (CFA) are below the threshold of an adequate fit (.90). The internal consistency for some factors was well below the threshold of a sufficient reliability (.70) for a new scale (Hinkin, 1998). The SAS (Bjorkman, 2007) is a 23-item scale developed with a US sample of junior high school students. However, its factor structure is problematic. For example, one of the four factors contains only 2 items which is less than the recommended minimum number of 3 (Costello & Osborne, 2005).

The AESI (Ang & Huan, 2006a) is the only one among the reviewed instruments that has been used in Asian countries. It was developed with Singapore students to measure the level of stress arising from academic expectations of both the students and significant others. Its psychometric profile has been well established and cross-cultural validity has also been tested with both Chinese and Hispanic students (Ang & Huan, 2006a, 2006b). However, the AESI authors acknowledged that the scale was limited to measurement of stress due to

academic expectations. The purpose of the present study was to extend the range of factors that contribute to the construct of educational stress.

In Chinese context, apart from high expectations, heavy burden of school work and homework, negative attitudes towards learning, such as dissatisfaction with grades, loss of interest, and difficulties in studying may also be important sources of pressure and stress among students (Lin & Chen, 1995; Lu, 2008). Thus, a more comprehensive tool seems necessary to study the nature and health effects of educational stress. Extensive review of both English and Chinese literature found no instrument that measures the multifactorial nature of academic stress among high school students in Asian countries.

Methods

Design

This study consisted of three cross-sectional questionnaire surveys with convenience samples of students (grades 7-12) from six secondary schools in three sites (the capital city, one county city and one rural town) in Shandong Province, China. These surveys were conducted in September and October 2009. The objective of the first survey was to explore the factor structure, internal consistency, concurrent validity and predictive validity of the draft ESSA scale. The second survey was to assess the test-retest reliability. The third survey was to evaluate the robustness of the factor structure established in the first survey.

Participants

Scale development sample. The first sample contained 364 grade 8 and 11 students. Data analysis was conducted with 347 students with a response rate of 95.3%. Of them, 44.8% were female and all were from the *Han* Chinese ethnic group. The age of the sample ranged from 12 to 18 (*Mean*=15.37, *SD*=1.69). Students from urban and rural families accounted for 43.2% and 56.8% of the sample, respectively.

Test-retest reliability sample. Two weeks later, a subset of the first sample (two classes, *N*=148) participated in a second survey with the same questionnaires. Data analysis on test-retest reliability was done with 135 (91.2%) respondents. The demographic characters were similar to the first sample.

CFA sample. A total of 1740 eligible participants (grades 7 through 12) from 36 classes were invited to participate in the third survey. Complete data were obtained from 1670 (95.8%) students and were included in the analysis. Of them, 44.6% were female and almost all (99.3%, 1659/1670) were ethnically *Han* Chinese. The age of the respondents ranged from 11 to 20 (*Mean*=15.44, *SD*=1.85). Urban and rural students accounted for 42.2% and 57.8% of the sample, respectively.

Measures

Educational stress. The preliminary version of the ESSA was used in the first two surveys. It contained 30 items derived from extensive review of both the English and Chinese literature and discussions with professionals in both public health and education in China. Six domains of stress consisting of five items each were predefined, including attitudes towards study and grades (such as "I am very dissatisfied with my academic grades"), perceived pressure (such as "I feel a lot of pressure in my daily studying"), perceived burden (such as "I feel that there is too much school work"), expectations from others (such as "I feel that I have disappointed my parents when my test/exam results are poor"), and self-expectation (such as "I feel stressed when I do not live up to my own standards"). Seven items were adapted from the AESI (Ang & Huan, 2006a) to form the last two dimensions. The response format used a 5-point Likert scale ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*) with a higher score indicating greater stress.

Items were initially created in English or adopted from other English scales. The Chinese version was then generated using the backward translation technique. Specifically,

two bilingual persons with Chinese background based at the Queensland University of Technology (QUT) independently translated the items into Chinese. The two Chinese copies were then sent to another bilingual professional based at Shandong University, China for review and translation into English. The back-translated scale was reviewed by an English native speaker at QUT to confirm its equivalence with the original. Revisions were made in the Chinese translation based on comments from the final reviewer.

After pilot testing, the scale was revised and a final 16-item version was used in the main survey. In the final scale, 5 items (item 9, 10, 14, 15, 16, Table 1) were adapted from the AESI (Ang & Huan, 2006a) with minor wording changes.

Academic expectation stress. The original (English) AESI (Ang & Huan, 2006a) was translated into Chinese following the same procedure for the ESSA and was used in this study. This 9-item scale has two subscales, Expectations of Parents/Teachers (five items) and Expectations of Self (four items). Respondents rated each statement on a 5-point Likert scale ranging from 1 (*never true*) to 5 (*almost always true*). The possible total score ranges from 9 to 45, with higher scores indicating greater stress. It gained good internal consistency (Cronbach's α = .89 for the total scale, .84-.85 for two factors) in the original study (Ang & Huan, 2006a).

Depressive symptoms. The Chinese version of Centre for Epidemiological Studies – Depression Scale (CCES-D, Radloff, 1977; Liu, 1999) is a 20-item self-report instrument for depressive symptoms. Items were rated using a 4-point scale from *Rarely or none of the time* (*less than 1 day*) to *Most or all of the time* (*5-7 days*) during the past week and were scored either 0-3 or 3-0, with a total range of 0-60, where higher scores indicate greater frequency and number of symptoms. It has 4 separate factors: depressive affect, somatic symptoms, positive effect, and interpersonal relations. The CES-D has good internal consistency with alphas of .85 for the general population and .90 for a psychiatric population (Radloff, 1977).

Suicidal thoughts. Suicidal thoughts were measured using one question "In the past 12 months, have you ever seriously considered attempting suicide?", adopted from the youth risk behaviour survey (YRBS) questionnaire (Eaton *et al.*, 2008). Students who responded "Yes" to this question were defined as having suicidal thoughts. This question has been widely used in the US (Brener et al., 2002, Eaton *et al.*, 2008), China (Chen, Dunne, & Han, 2006) and elsewhere.

Academic grades. Participants were asked to rank their average grades during the past 12 months into one of the five categories: *Very poor, Poor, Middle, Good,* and *Very good*. In data analysis, the first two and last two groups were combined as *Very poor/Poor* and *Good / Very good*, respectively, resulting in three categories coded from 1 to 3.

Procedure

All surveys were conducted in schools during self-study sessions. A brief introduction was given by the investigator and followed by the distribution of assent forms and information sheets. The survey questionnaire was then administered to the students who signed assent forms. On average, it took 30 minutes for the students to complete the questionnaires. For the test-retest survey, a technique reported by Brener and colleagues (1995; 2002) was followed to assure anonymity and obtain matching data from participants.

Data Analyses

Analyses were conducted using SPSS for Windows 17.0 (SPSS Inc, Chicago, IL) and Amos 7.0 (SPSS Inc, Chicago, IL). All statistical tests were two-sided and significance level was defined as α=.05.

Ethics Approval

This project obtained ethics approval from the University Human Research Ethics Committee of QUT and the Preventive Medicine Ethics Committee of Shandong Provincial Centre for Disease Control and Prevention (CDC). Participation was entirely voluntary and

anonymous. Before the data collection, a written approval was given by the principal/vice principal of each participating school and a standard assent was gained from each student. Passive consent was also obtained from parents.

Results

EFA

The factor structure of the preliminary 30-item ESSA was identified using Exploratory Factor Analysis (EFA) with Principal Axis Factoring (PAF) as the extraction method. The Promax method was used for rotation because the factors were thought to be correlated. The Kaiser–Meyer–Olkin's measure (KMO) of sampling adequacy (acceptable level >.50) (Kaiser, 1970) and Bartlett's test of sphericity (Bartlett, 1950) were calculated to verify the appropriateness of an EFA. The number of factors was determined using Parallel Analysis (PA) performed with SPSS syntax developed by O'Connor (2000). Only items with a strong loading (.50 or higher) on one factor and <.30 on any other factors were retained to form latent variables because large loadings on factors other than the primary factor could result in serious flaws in the factor structure (Costello & Osborne, 2005).

The EFA with data from the 30-item ESSA scale indicated that the sample and correlation matrix were appropriate for factor analysis (KMO index = .88, and Bartlett's test of Sphericity was significant (χ^2 (435, N = 347) = 3675.80, p<.001)). Five factors were extracted based on PA analysis that cumulatively accounted for 52.1% of the total variance. However, several items were found to have a poor loading (<.5) on all factors or had crossed loadings (loading ≥.3 in two or more factors). These items were then dropped and a repeated EFA was conducted with remaining items using the same method. This procedure was replicated until all retained items met the criteria. Finally, a 16-item scale was generated with all items having a strong loading but on the primary factor, but not on the other factors (Table 1).

The appropriateness of the EFA for the revised 16-item ESSA scale was again justified (KMO =.81; Bartlett's test: χ^2 (120, N = 347) = 1495.83, p<.001). The number of factors remained the same based on a repeated PA. Each factor contained at least 3 items (Table 1). The initial eigenvalues of five factors were 4.26, 2.30, 1.31, 1.22, and 1.07. The post-rotation traces were 3.20, 2.25, 1.96, 2.06 and 2.36, respectively. These latent variables explained 26.6%, 14.4%, 8.2%, 7.6%, and 6.7% of variance respectively, and together 63.6% of the total variance. Interfactor correlations ranged from .04 to .57. After carefully examining the meaning, these factors were labelled as Pressure from study, Workload, Worry about grades, Self-expectation, and Despondency (Table 1).

Reliability

The internal consistency reliability was assessed using Cronbach's alpha and average inter-item correlation. A Cronbach's alpha of .70 or higher, or an average inter-item correlation of .30 or higher indicates acceptable reliability (Robinson, Shaver, & Wrightsman, 1991). The test-retest reliability was assessed with Intraclass Correlation Coefficients (ICCs, Koch, 1982). An ICC of .2 and lower indicates "poor"; .21-.40 "fair"; .41-.60 "moderate"; .61-.80 "good"; and .80 or higher as "almost perfect" reliability (Landis & Koch, 1977).

Based on the data from the first survey (N = 347), the Cronbach's alpha for the total 16-item ESSA scale was .81 indicating good internal consistency. The coefficient alpha for each factor ranged from .66 - .75 and most were above the criteria for an acceptable level of reliability (Table 2). The average inter-item correlations for the five factors were .47, .50 .47, .39, and .39.

Using the data from the second sample (N = 135), the ICC for the total ESSA score was .78, and for the five factors was .75, .61, .70, .59, and .62, respectively, with the majority suggesting good test-retest reliability. The ICC for each of the 16 items varied from .44 to .67 suggesting moderate to good reliability over two weeks.

Concurrent and Predictive Validity

The AESI (Ang & Huan, 2006a) served as a criterion measure to assess the concurrent validity of the ESSA. Scores from the two scales were hypothesised to be correlated because expectations are an important source of academic stress among adolescents. Academic grades were also hypothesized to be associated with educational stress with lower level students having more stress. To assess predictive validity, depression (CES-D score) and suicidal thoughts were used as criterion measures because of their known associations with academic stress (Ang & Huan, 2006b; Bjorkman, 2007; Liu & Tein, 2005). The expected relationships were analysed using Pearson correlation for continuous variables, point-biserial correlation for associations between ESSA scores and suicidal thoughts, and Spearman correlation for relationships between academic grades and others.

As expected, the ESSA total score was significantly correlated with the AESI scores (Table 2). Three ESSA factors, i.e., Pressure from study, Worry about grades and Self-expectation were also significantly correlated with AESI total and subscales (Table 2). However, there were no significant correlations between other two ESSA factors (Workload and Despondency) and AESI scores (Table 2).

The overall ESSA score was negatively correlated with academic grades (Spearman $r = -.20$, $p < .001$), indicating that students with low academic achievements have more stress. However, only two of the five factors (Workload and Despondency) showed significant correlations with self-reported academic grades (Table 2).

Total academic stress and all factors were positively correlated with CES-D score (Table 2). The coefficient for overall stress (.47) approached a moderate effect size according to Cohen's (1988) criteria ($r = .5$). There were also significant correlations between suicidal thoughts and total ESSA score and two of the factors (Pressure from study and Despondency, Table 2).

CFA

Using Amos 7.0 (SPSS Inc, Chicago, IL), a CFA was conducted to assess the fit of the exploratory model to the data from the third sample (N=1670). The maximum likelihood method was used to estimate these parameters (Byrne, 1994). Missing data were rare (all items <1%) and assumed to be missing at random. A set of goodness of fit indices were calculated, including the traditional chi-square fit index, comparative fit index (CFI), Bentler-Bonett normed fit index (NFI), incremental fit index (IFI), goodness-of-fit index (GFI), root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). A value of .90 or higher for CFI, NFI, IFI, and GFI, a RMSEA of .06 or lower, and a SRMR of .05 or lower were served as the indicators for an adequate fit (Byrne, 1994; Hu & Bentler, 1999; MacCallum, Brown, & Sugawara, 1996).

According to the results of the traditional Chi-square fit index (χ^2 (94, N = 1670) = 604.59, p < .001), the observed model was significantly different from the expected model. However, this may be related to the large sample size. All other indices, including the CFI (.93), NFI (.92), IFI (.93), GFI (.96), RMSEA (.06) and SRMR (.05) suggested an adequate fit to the original factor structure. The factor loading for each item on the corresponding factor in the CFA were similar to the results from the EFA and all were above .50 (Table 1).

Considering the large correlation between the overall score and each subscale (Table 2), there is likely a second-order factor. We thus conducted a repeated CFA to test the alternative model including a second-order factor. Compared to the first model, all indices slightly changed and some of those fell below the adequacy criteria (Chi-square fit index: χ^2 (99, N = 1670) = 815.57, p < .001; CFI= .90; NFI = .89; IFI = .90; GFI = .94; RMSEA = .07; and SRMR = .07). However, three indices, CFI, IFI and GFI were still above the criteria for an adequate fit. The factor loadings for the five first-order factors on the second-order factor were .97, .64, .40, .44 and .73, respectively.

Discussion

A new instrument for academic stress was developed and validated in this study with over 2000 adolescents from urban and rural areas of Shandong, China. The final scale contains 16 items and five latent variables, i.e., Pressure from study, Workload, Worry about grades, Self-expectation and Despondency. Scores from this scale exhibit satisfactory psychometric properties in terms of internal and test-retest reliability and concurrent and predictive validity.

Application of a relatively high criterion to retain items in the final scale (Costello & Osborne, 2005) resulted in nearly half (14 items) of the initial items being dropped because of poor loading on the primary factor and high loadings on the other factors. By doing this, there may be a risk that the subscales suffer from construct underrepresentation, which might also be related to the slightly low internal consistency of some factors. However, given that the application of this instrument is for school or community based survey of students to examine the magnitude and health associations of academic stress, rather than for clinical diagnostic purposes, a brief scale with clear factor structure may be preferable than a lengthy but more accurate one. The relatively low Cronbach's alphas for some factors are likely to be caused by the small numbers of items per scale. The average inter-item correlations for all factors are well above the criterion ($\geq$.30) for an acceptable internal consistency (Robinson, Shaver, & Wrightsman, 1991), indicating the items within each subscale are highly correlated. In addition, the number of items in each factor meets the minimum number of 3 items for best practice in factor analysis (Costello & Osborne, 2005). More importantly, the number of factors determined using Parallel Analysis did not change even after dropping poor or cross loaded items, indicating minimal change to the factor structure.

Confirmatory factor analysis is essential in the development of a new scale and should be conducted with data from in a different sample from the EFA (Hinkin, 1998). Using a

large, separate sample (*N*=1670), we tested the fit of the factor model of the 16-item ESSA and found an adequate fit according to a range of indices. All indices except the chi-square fit index meet the recommended thresholds for an adequate fit. The high value of the chi-square index is likely related to the large sample size (*N*=1670). Although all indices negatively changed in a repeated CFA involving a second-order factor, there is still evidence to suggest that the revised model has adequate fit and the construction of an aggregated ESSA total score is appropriate. However, the factor loadings for two first-order factors, i.e., Worry about grades and Self-expectation are relatively low (.40 and .44), suggesting further analysis is necessary to examine the convergent and discriminant validity of the ESSA scale.

Academic expectations are an important source of stress among Asian students (Ang & Huan, 2006a). As expected, the ESSA score was significantly correlated with the AESI score, indicating a satisfactory concurrent validity. Two factors in the ESSA, Worry about grades and Self-expectation obtained relatively high correlation coefficients with the AESI-Other expectations and AESI-Self-expectations (*r*=.53 and .96, Table 2). This is because two items (item 9 and 10, Table 1) of Worry about grades and all items (items 14-16, Table 1) were adapted from the AESI with minor changes, although the response format (from *Strongly disagree* to *Strongly agree*) differs from the AESI (from *Never true* to *Almost always true*). Two other factors, Despondency and Workload are not captured by the AESI which supports the value of development of a new multifactorial scale.

Academic stress is recognized as a risk factor for depression and suicidal behavior (Ang & Huan, 2006b; Bjorkman 2007). In this study, the ESSA scores show significant associations with these health problems. Compared to the AESI, the ESSA scores appear to be more predictive of depression and suicidality (Table 2). This is probably because the ESSA captures more elements of academic stress than the AESI. In relation to the Suicidal thoughts and Academic grades, the ESSA scales have relatively poor predictive validity (see

Table 2). One explanation is that academic stress is only one of the correlates that are associated with these two variables. Many other factors, such as loss of loved ones, conflicts with parents, teachers and peers and significant physical diseases may have important effects on adolescent suicidality (Liu & Tein, 2005). Similarly, while poor academic grades generally predict high educational stress, the discrepancy between expected and actual grades may play a more important role in the development of psychological distress and other mental health problems (Lin *et al.*, 2008). In addition, this could be also related to the poorly measured criterion variables, as suicidal thoughts does not include an academic component and academic grades was not very precise given its subjective nature. More suitable criterion measures should be used in future research.

This study has some implications for educational policy and practice, including school counseling. A brief tool with sound psychometric properties could be used to examine the nature and magnitude of the phenomenon in many educational contexts, to inform the design and implementation of interventions to reduce educational stress in schools. Students' mental health and wellbeing has been drawing increasing attention in China where school counseling has been made available only in recent years. School Counselors should have a good understanding of the multifactorial nature of educational stress and its links to common mental and behavioral problems among students to inform best practice in counseling.

This study has some limitations. First, the development of the items in the ESSA was mainly based on review of recent Chinese and English literature plus informal discussions with experts. No attempt was made to more comprehensively map the construct using grounded theory to explore an underlined model. Second, despite the identification of five factors with just 16 items, the ESSA cannot capture all facets of educational stress. More work should be done to further investigate the multidimensional nature. Third, the ESSA was only tested with Chinese adolescents in Shandong and cross-cultural suitability is yet to be

established. Therefore, this work should be viewed as a starting point of a continuous process of validation and revision. Fourth, we used a single question to measure self-report academic grades but we do not know if there are disparities between perceived grades and actual grades. Actual scores are ideal but very difficult to obtain in a self-report anonymous survey. Further research is needed to examine the difference between self-report and actual grades and their relationships with stress and other outcomes. Sixth, information in this study was collected solely relying upon self-report of students and hence some recall bias cannot be avoided.

Nevertheless, this newly developed scale demonstrates satisfactory psychometric properties and is suitable to be used in further research into academic-related stress among secondary school adolescents. The ESSA promises to be a useful tool at least with Chinese populations and in other Asian countries, and possibly useful in different social and cultural contexts.

References

Abouserie, R. (1994). Sources and levels of stress in relation to locus of control and self esteem in university students. *Educational Psychology, 14*, 323-330. doi 10.1080/0144341940140306

All-China Women's Federation. (2008). National juvenile family education status sampling survey report. Retrieved from http://www.women.org.cn/allnews/02/1986.html

Ang, R. P., & Huan, V. S. (2006a). Academic Expectations Stress Inventory: Development, factor analysis, reliability, and validity. *Educational and Psychological Measurement, 66*, 522-539. doi: 10.1177/0013164405282461

Ang, R. P., & Huan, V. S. (2006b). Relationship between academic stress and suicidal ideation: Testing for depression as a mediator using multiple regression. *Child Psychiatry & Human Development, 37*, 133-143. doi: 10.1007/s10578-006-0023-8

Ang, R. P., Huan, V. S., & Braman, O. R. (2007). Factorial structure and invariance of the Academic Expectations Stress Inventory across Hispanic and Chinese adolescent samples. *Child Psychiatry & Human Development, 38*, 73-87. doi: 10.1007/s10578-006-0044-3

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology 3*, 77-85.

Bjorkman, S. M. (2007). *Relationships among academic stress, social support, and internalizing and externalizing behavior in adolescence*. (Unpublished Ph.D's thesis). Northern Illinois University, Illinois, United States.

Brener, N. D., Collins, J. L., Kann, L., Warren, C. W., & Williams, B. I. (1995). Reliability of the youth risk behavior survey questionnaire. *American Journal of Epidemiology, 141*, 575-580.

Brener, N. D., Kann, L., McManus, T., Kinchen, S. A., Sundberg, E. C., & Ross, J. G. (2002). Reliability of the 1999 youth risk behavior survey questionnaire. *Journal of Adolescent Health, 31*, 336-342. doi: 10.1016/S1054-139X(02)00339-7

Brown, S. L., Teufel, J. A., Birch, D. A., & Kancherla, V. (2006). Gender, age, and behavior differences in early adolescent worry. *Journal of School Health, 76*, 430-437. doi: 10.1111/j.1746-1561.2006.00137.x

Burnett, P. C., & Fanshawe, J. P. (1997). Measuring school-related stressors in adolescents. *Journal of Youth and Adolescence, 26*, 415-428. doi: 10.1023/A:1024529321194

Byrne, B. M. (1994). Testing for the factorial validity, replication, and invariance of a measuring instrument: Aparadigmatic application based on the Maslach Burnout Inventory. *Multivariate Behavioral Research, 29*, 289-311. doi: 10.1207/s15327906mbr2903_5

Chen, J., Dunne, M. P., & Han, P. (2006). Child sexual abuse in Henan province, China: Associations with sadness, suicidality, and risk behaviors among adolescent girls. *Journal of Adolescent Health*, 38, 544-549. doi:10.1016/j.jadohealth.2005.04.001

China Youth Social Service Center. (2008). 2007 China national juvenile internet use survey report. Retrieved from http://news.qq.com/a/20081020/001953.htm

Christie, E., & MacMullin, C. (1998). What do children worry about? *Australian Journal of Guidance and Counselling, 8*, 9-24.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation, 10*, 1-7. Retrieved from http://pareonline.net/getvn.asp?v=10&n=7

Crystal, D. S., Chen, C., Fuligni, A. J., Stevenson, H. W., Hsu, C.C., Ko, H.J., … Kimura, S. (1994). Psychological maladjustment and academic achievement: A cross-cultural study of Japanese, Chinese, and American high school students. *Child Development, 65*, 738-753. doi: 10.1111/j.1467-8624.1994.tb00780.x

Dodds, J., & Lin, C.D. (1992). Chinese teenagers' concerns about the future: A cross-national comparison. *Adolescence, 27,* 481-486.

Eaton, D. K., Kann, L., Kinchen, S., Shanklin, S., Ross, J., Hawkins, J., … Wechsler, H. (2008). Youth Risk Behavior Surveillance -- United States, 2007. *MMWR Surveillance Summaries, 57(SS04),* 1-131. Retrieved from http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5704a1.htm

Gadzella, B. M. (2001). Confirmatory factor analysis and internal consistency of the Student-life Stress Inventory. *Journal of Instructional Psychology, 28*, 84-94.

Gallagher, M., & Millar, R. (1996). A survey of adolescent worry in Northern Ireland. *Pastoral Care in Education, 14*, 26-32. doi: 10.1080/02643949609470963

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1*, 104-121. doi: 10.1177/109442819800100106

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling, 6*, 1-55. doi: 10.1080/10705519909540118

Huan, V. S., See, Y. L., Ang, R. P., & Har, C. W. (2008). The impact of adolescent concerns on their academic stress. *Educational Review, 60*, 169-178. doi: 10.1080/00131910801934045

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika, 35*, 401-416. doi: 10.1007/BF02291817

Koch, G. G. (1982). Intraclass correlation coefficient. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of Statistical Sciences. 4.* (pp. 213-217): New York: John Wiley & Sons.

Kohn, J. P., & Frazer, G. H. (1986). An academic stress scale: Identification and rated importance of academic stressors. *Psychological Reports, 59*, 415-426.

Lakaev, N. (2009). Validation of an Australian Academic Stress Questionnaire. *Australian Journal of Guidance & Counselling, 19*, 56-70. doi: 10.1375/ajgc.19.1.56

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Lee, M., & Larson, R. (2000). The Korean 'examination hell': Long hours of studying, distress, and depression. *Journal of Youth and Adolescence, 29*, 249-271. doi: 10.1023/A:1005160717081

Li, H., & Zhang, Y. (2008). Factors predicting rural Chinese adolescents' anxieties, fears and depression. *School Psychology International, 29*, 376-384. doi: 10.1177/0143034308093676

Lin, H. C., Tang, T. C., Yen, J. Y., Ko, C. H., Huang, C. F., Liu, S. C., … Yen, C,F. (2008). Depression and its association with self-esteem, family, peer and school factors in a population of 9586 adolescents in southern Taiwan. *Psychiatry and Clinical Neurosciences, 62*, 412-420. doi: 10.1111/j.1440-1819.2008.01820.x

Lin, J., & Chen, Q. (1995). Academic pressure and impact on students' development in China. *McGill Journal of Education, 30*, 149-168.

Liu, P. (1999). Center for Epidemiologic Studies Depression scale (Chinese version, translated from Radloff, 1977). In X.-D. Wang, X.-L. Wang & H. Mang (Eds.), *Rating Scales For Mental Health* (pp. 200-202). Beijing: Chinese Mental Health Journal Publishing Company.

Liu, X., & Tein, J.Y. (2005). Life events, psychopathology, and suicidal behavior in Chinese adolescents. *Journal of Affective Disorders, 86*, 195-203. doi: 10.1016/j.jad.2005.01.016

Lu, H.D. (2008). Focus on learning stress of Chinese children: The puzzledom and the way out. *Journal of Northeast Normal University (Philosophy and Social Sciences), 6*, 24-28.

MacCallum, R. C., Brown, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149.

O'connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*, 396-402.

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

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). *Criteria for scale selection and evaluation in measure of personality and social psychological attitudes*: Calif: Academic Press, San Diego.

Tang, N. Y. Y., & Westwood, P. (2007). Worry, general self-efficacy and school achievement: An exploratory study with Chinese adolescents. *Australian Journal of Guidance and Counselling, 17*, 68-80. doi: 10.1375/ajgc.17.1.68

Zeidner, M. (1992). Sources of academic stress: The case of 1st year Jewish and Arab college-students in Israel. *Higher Education, 24*, 25-40. doi: 10.1007/BF00138616

Zhang, H., Tao, F., & Zeng, G. (2001). Depression and its correlates among middle school students in Anhui province. *Chinese Journal of School Health, 22*, 497-498.

Table 1

Rotated Factor Loadings and Communalities (h^2) for the ESSA in the EFA (N=347) and Factor Loadings in the CFA (N = 1670)

	EFA						CFA
	Factor loadings					h^2	Factor loadings
	1	2	3	4	5		
6. I feel a lot of pressure in my daily studying.	**.77**	-.01	-.02	-.09	.03	.56	.80
11. There is too much competition among classmates which brings me a lot of academic pressure.	**.76**	.06	.03	.05	-.21	.51	.66
4. Future education and employment bring me a lot of academic pressure.	**.68**	-.09	-.02	.04	.08	.51	.72
5. My parents care about my academic grades too much which brings me a lot of pressure.	**.54**	.13	.04	-.08	.14	.36	.64
3. I feel there is too much homework.	.06	**.81**	-.08	.06	-.07	.68	.70
2. I feel that there is too much school work.	-.11	**.70**	.10	-.07	.15	.52	.68
7. I feel that there are too many tests /exams in the school.	.07	**.59**	.04	.03	-.02	.39	.69
10. I feel that I have disappointed my teacher when my test/exam results are not ideal.	-.10	.09	**.93**	-.09	-.10	.74	.67
9. I feel that I have disappointed my parents when my test/exam results are poor.	.15	.01	**.59**	.15	.04	.56	.85
8. Academic grade is very important to my future and even can determine my whole life.	.13	-.13	**.53**	.12	.14	.37	.52
14. I feel stressed when I do not live up to my own standards.	-.02	.08	-.10	**.84**	.01	.66	.71
15. When I fail to live up to my own expectations, I feel I am not good enough.	-.01	.01	.05	**.59**	-.06	.35	.56
16. I usually cannot sleep because of worry when I cannot meet the goals I set for myself.	-.04	-.08	.10	**.52**	.08	.32	.58
12. I always lack confidence with my academic scores.	.04	.03	-.03	-.04	**.68**	.49	.52
1. I am very dissatisfied with my academic grades.	-.15	-.01	.04	.05	**.66**	.37	.68
13. It is very difficult for me to concentrate during classes.	.17	.08	-.08	.00	**.51**	.39	.67

Note. ESSA = Educational Stress Scale for Adolescents; EFA = Exploratory Factor Analysis; CFA = Confirmatory Factor Analysis; Factor 1= Pressure from study; Factor 2 = Workload; Factor 3 = Worry about grades; Factor 4 = Self-expectation; Factor 5 = Despondency; h^2 = Communalities

Factors loadings in the CFA are the standardised regression weights for each item with the corresponding factor.

Table 2

Mean (SD), Alpha Coefficients and Inter-variable Correlation Coefficients in the First Survey (N=347) [a]

	M (*SD*)	α[b]	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1. ESSA total	54.14 (9.32)	.81	1											
2. Pressure from study	13.99 (3.56)	.74	.81**	1										
3. Workload	9.51 (2.90)	.75	.58**	.39**	1									
4. Worry about grades	11.38 (1.25)	.71	.57**	.34**	.05	1								
5. Self-expectation	9.91 (2.56)	.66	.55**	.27**	.04	.35**	1							
6. Despondency	9.31 (2.83)	.66	.67**	.44**	.33**	.14**	.19**	1						
7. AESI total	30.61 (6.46)	.85	.51**	.29**	-.03	.52**	.83**	.07	1					
8. AESI-Other expectations	17.22 (3.86)	.81	.41**	.26**	-.06	.53**	.58**	.01	.92**	1				
9. AESI-Self expectations	13.38 (3.30)	.73	.52**	.26**	-.01	.39**	.96**	.13*	.88**	.62**	1			
10. CES-D (Depression)	15.34 (8.93)	.87	.47**	.38**	.25**	.15**	.24**	.44**	.19**	.13*	.22**	1		
11. Suicidal thoughts	N/A	N/A	.17**	.12*	.13*	.04	.03	.21**	.01	.02	.04	.42**	1	
12. Academic grades	N/A	N/A	-.20**	.10	-.13*	.03	.02	-.43**	.16**	.21**	.07	-.17**	-.07	1

Note. *M* = Mean; *SD* = Standard deviation; ESSA = Academic Stress Scale for Adolescents; AESI = Academic Expectation Stress Inventory; CES-D = Centre for Epidemiological Studies – Depression Scale; N/A = Not applicable

[a] Pearson correlation coefficients for continuous variables; Point-biserial correlation coefficients for correlations between suicidal thoughts and others; Spearman *r* for relationships between academic grades and others.

[b] Cronbach's α coefficient

* $p<.05$; ** $p<.01$