Translation and Validation of a Chinese Version of the Cleveland Adolescent Sleepiness Questionnaire

Ya-Ting Yang 1,2
Han-Yun Chang 1
Chung-Yao Hsu 3
Jamie M Zeitzer 2,4

1Institute of Education, National Cheng Kung University, Tainan City, Taiwan; 2Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA; 3Department of Neurology, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; 4Mental Illness Research Education and Clinical Center, VA Palo Alto Health Care System, Palo Alto, CA, USA

**Purpose:** The Cleveland Adolescent Sleepiness Questionnaire was originally developed and published in English and has served as a valid and effective tool for the assessment of adolescents’ experiences with sleepiness in a variety of situations. To allow for comparisons between sleepiness in adolescents from different cultures, and with different linguistic backgrounds, reliable and valid measurement tools are necessary. The purpose of this study was to translate and validate a Chinese version of the Cleveland Adolescent Sleepiness Questionnaire (C-CASQ).

**Materials and Methods:** Sensitivity, specificity, internal consistency, and criterion validity data for the C-CASQ were tested using 458 adolescents in Taiwan. Data from 191 participants were used to establish internal consistency reliability and conduct exploratory factor analysis (EFA), while data from 267 participants were used to establish criterion validity and conduct confirmatory factor analysis (CFA). Initial criterion validity was established through a comparison of the C-CASQ with scores from the Chinese version of the Morningness-Eveningness Scale for Children, a measure of chronotype.

**Results:** EFA resulted in four factors, consistent with the original English version of the CASQ, while CFA established goodness of fit. The scale demonstrated acceptable to good internal consistency (α = 0.77–0.86). Initial criterion validity was evident as the total score and each of the subscale scores on the C-CASQ was significantly higher (greater sleepiness) in evening-types.

**Conclusion:** The C-CASQ appears to be a psychometrically sound measure to evaluate sleepiness in Chinese-speaking adolescents.

**Keywords:** teens, sleep, alertness, chronotype

**Introduction**

Over the past decade, adolescent sleep has become a major public health issue and exploration of the causes and consequences of insufficient sleep has gained increasing attention from researchers, clinicians, parents, and the general public. 1–3 According to the National Sleep Foundation,4 healthy teenagers need 8 to 10 hours of sleep every night. Numerous studies have reported that adolescent sleep duration, however, is significantly less than that which is recommended, particularly in Asian countries. 5,6 Sufficient sleep is not only necessary to avoid daytime fatigue but is crucial for adolescents’ physical growth, emotional stability, behavioral control, maintenance of cognitive function and academic performance. 7–9 Moreover, current studies have also found that the daytime sleepiness was related to
the morningness/eveningness dimension; therefore, students with more pronounced eveningness reported greater daytime sleepiness.\textsuperscript{10–12} Furthermore, adolescents who reach the recommended sleep duration demonstrate a significantly better performance in mathematics, weight control and depression compared to adolescents who have lower than the recommended sleep duration.\textsuperscript{13–15}

One of the most prominent, proximal consequences of sleep loss is daytime sleepiness.\textsuperscript{16} Daytime sleepiness can be defined by both objective measures (eg, the speed with which one can fall asleep if given the opportunity) and questionnaires that capture various aspects of the subjective phenomenology surrounding the catchall of “sleepiness”. One of the more commonly used tools to assess sleepiness is the Epworth Sleepiness Scale (ESS), which has been previously translated into Chinese.\textsuperscript{17} While Chen et al (2002) found high levels of internal consistency (Cronbach’s $\alpha = 0.81$), the ESS is validated for use with adult respondents. Another commonly used tool is the Pediatric Daytime Sleepiness Scale (PDSS), which also has been translated into Chinese,\textsuperscript{18} though this tool is designed to assess sleepiness in middle school-aged (11–15 years) children.\textsuperscript{19}

The Cleveland Adolescent Sleepiness Questionnaire (CASQ)\textsuperscript{20} was specifically designed to evaluate adolescents’ experiences of sleepiness in a variety of situations, including in school, at home during the evening, and while in transit. The CASQ has shown acceptable levels of internal consistency reliability, as well as criterion or concurrent validity. The generalizability of the CASQ to other populations or cultures, however, must be assessed\textsuperscript{21} as it was developed based on adolescents in the US. As such, reliability and validity may be limited in terms of assessing daytime sleepiness among Chinese adolescents, as different cultures have divergent views on the nature, purpose, and importance of sleep.\textsuperscript{22} There is currently, however, no validated translation of the CASQ into Chinese. In order for a cross-cultural adaptation of a questionnaire to reach equivalence between the original source and target versions of the questionnaire, it is recommended that the items must not only be translated well linguistically but also must be adapted culturally to maintain the content validity of the instrument at a conceptual level across different cultures.\textsuperscript{23} Thus, there is particular value for cross-cultural studies to develop a validated Chinese version of the CASQ.

In light of the lack of a Chinese version of the CASQ, the purpose of this study was to (1) develop a Chinese translation of the CASQ and (2) establish the sensitivity, specificity, internal consistency, and basic criterion validity data of the Chinese version of the CASQ (C-CASQ). For the translation and cross-cultural adaptation from American English to Chinese, we based our procedures on methodologic guidelines from psychology and sociology.\textsuperscript{24}

**Materials and Methods**

Prior to participation, all participants signed an informed consent and parents provided informed consent for their child who was under the age of 18 to participate. The study was approved by the National Cheng Kung University Institutional Review Board (IRB). All procedures were conducted in accordance with the Declaration of Helsinki.

**Translation of the CASQ Questionnaire Translation**

The aim of the translation process was to develop a conceptually equivalent version of an English instrument for use among Chinese respondents. As such, it was necessary that the instrument be intuitive, appropriate, and comprehensible in order for the Chinese version to perform practically in the same way as the original. A well-established method for developing conceptually equivalent and linguistically and culturally appropriate versions of a scale is to use both forward translation and back translation.\textsuperscript{23} The translation was performed using a decentered approach, which seeks to ensure that both the original and translated versions remain meaningful in their respective languages, and which allows revisions to be made in order to ensure the consistency and meaningfulness of the wording in the cultural contexts of both languages. The CASQ was translated into Chinese by two native Taiwanese speakers who were college graduates with degrees in Chinese Literature and familiar with research related to health education. The initial translations by both translators were conducted independently without any communication among the translators. The main researchers and the two translators then analyzed the resulting translations together. After discussing variations in each translation (T1 and T2), a single version (T3) was agreed upon. For example, in the original question, “I feel alert during my classes,” the word “alert” could have multiple translations into Chinese. The translation of T1 tended towards a connotation of vigilance while the translation of T2 tended towards the connotation of concentration.
The T3 agreed upon the “concentration” translation as being most representative of the original meaning.

**Back Translation**
From the synthesized translation (T3), two back translations (BT1 and BT2) were completed by two native English speakers who had no knowledge of or access to the original instrument. One translator was a native-English speaking professor from a Foreign Languages department, with 15 years of Chinese language learning experience. The second translator was a college professor who is a health expert, a resident of the United States and bilingual in English and Chinese. Both back-translations were conducted individually at first, with the main researchers and the two back translators analyzing the back translations together after their completion. After discussing variations between the back translations (BT1 and BT2), a single version (BT3) was agreed upon. A revised Chinese version (T4) was developed based on BT3. For example, the original question, “During the school day, there are times when I realize that I have just fallen asleep,” was translated by BT1 as: “At school, there are times when I find that I have just fallen asleep”, while the translation of BT2 was “There have been a few times I found myself fell asleep during the school day”. As there is no difference between past tense or present tense in Chinese, BT3 decided to ignore the difference between “find” of BT1 and “found” of BT2.

**Expert Committee**
The translations and back translations (T1-T4 and BT1-BT3) were assessed by two expert clinicians specializing in sleep medicine, with 15 years and 31 years of experience, respectively. The clinicians examined the various versions and relevant annotations from the translators and discussed the items from the questionnaire to ensure a clear final version, equivalent to the original in terms of the connotations of the language used, naturalness of the language, and aspects of culture and belief. The clinicians drew attention to some of the questionnaire’s items and recommended that the reliability and validity of the final version be verified.

**Research Design and Participants**
Reliability and validity were assessed using a two-stage research design. Stage 1 involved examining internal consistency, and conducting exploratory factor analysis (EFA). Stage 2 involved examining criterion validity (by comparing differences in responses from individuals with different chronotypes) and conducting confirmatory factor analysis (CFA). To satisfy the basic requirement of a sample size of 5–10 participants for each of the 16 items in EFA, the researchers recruited a sample of 191 Taiwanese secondary students (grades 9–12). To satisfy a sample size of at least 200 for CFA, the researchers recruited a separate sample of 267 Taiwanese secondary students (grades 9–12). All participants were full-time students and not preselected for any type of sleep disorder or other health problem. Students were tested randomly throughout the day.

**Measures**
**Cleveland Adolescent Sleepiness Questionnaire (CASQ)**
The CASQ was designed to measure daytime sleepiness in adolescents from 11 to 17 years of age. The CASQ consists of a total of 16 items scored using a 5-point Likert scale (with five of the statements worded for reversed scoring) resulting in a scoring range of 16–80. Higher scores indicate increased daytime sleepiness. The four factors of the CASQ are sleepiness in school (five questions), alertness in school (five questions), sleepiness during the evening (three questions), and sleepiness during transport (three questions). The Cronbach’s α for the original English CASQ version is 0.89. The results of exploratory factor analysis explained 55% of the total variance. The Root Mean Square Error of Approximation, Tucker–Lewis Index and Comparative Fit Index of confirmatory factor analysis for the original English CASQ version were 0.06, 0.94 and 0.95, respectively.

**Morningness-Eveningness Scale for Children (MESC)**
A Chinese version of the MESC was administered to participants in order to determine chronotype, a measure of the degree to which respondents prefer mornings or evenings. This measure, initially validated in 11 to 12 year-old participants, has also frequently been used with adolescent samples. The test-retest reliability of the Chinese version is 0.75. The internal consistency reliability of the sample evaluated by this study is 0.74. The current study used cut-off values at the 25th and 75th percentile to define chronotypes, with resulting values of 10–23 corresponding to evening-type chronotypes, 24–28 for intermediate-type chronotypes, and 29–43 for morning-type chronotypes. As several studies have shown...
that evening-type adolescents suffer from greater daytime sleepiness than morning-types.\textsuperscript{29,30} Chronotype was used as a measure for establishing initial criterion validity.

**Data Analysis**

Internal consistency, EFA, and criterion validity were conducted using SPSS 20.0 (IBM, Armonk NY). CFA was conducted using AMOS 24.0. Internal consistency was measured using Cronbach’s $\alpha$ for the overall score and subscale scores. Cronbach’s $\alpha \geq 0.7$ is considered acceptable with $\alpha \geq 0.8$ considered good, but $\alpha \geq 0.9$ suggesting potential redundancy among scale items.\textsuperscript{31} The KMO statistic and Bartlett’s test were first evaluated to determine the suitability of the data for EFA. KMO should be 0.60 or higher to proceed with factor analysis.\textsuperscript{32} Criterion validity was measured by independent sample $t$-tests for overall score and subscale scores, comparing sleepiness differences by chronotype. CFA was evaluated by multiple goodness-of-fit indices:\textsuperscript{33,34} [1] $\chi^2$/df, [2] Root Mean Square Error of Approximation (RMSEA), [3] Comparative Fit Index (CFI), [4] Tucker–Lewis Index (TLI), and [5] Standardized Root Mean square Residual (SRMR).

**Results**

Participants ($N = 458$; 51% female, 15.5 ($SD = 1.14$) years old) from grades 9 through 12 completed the questionnaires. They were randomly assigned to one of the two groups, with the data from the first group used to test for internal consistency and conduct EFA (stage 1, $n = 191$; 55% female, 15.7 ($SD = 1.14$) years old), and data from the second group used to establish criterion validity and conduct CFA (stage 2, $n = 267$; 48% female, 15.3 ($SD = 1.12$) years old). There were no differences between stage 1 and stage 2 groups regarding gender ($\chi^2 = 2.80, p = 0.09$) or age ($\chi^2 = 2.10, p = 0.55$) by chi-square test for homogeneity.

**Internal Consistency Reliability**

The internal consistency for the overall instrument was $\alpha = 0.85$. All subscales were found to demonstrate acceptable to good internal consistency: 0.77 (sleepiness during the evening), 0.77 (sleepiness during transport), 0.79 (alertness in school), and 0.86 (sleepiness in school).

**Exploratory Factor Analysis (EFA)**

Principal component analysis with the varimax rotation was utilized to extract the factor structure. EFA was performed on each of the 16 items, with a KMO value of 0.82 and Bartlett’s test of sphericity $= 1311.29$ ($p = 0.00$), indicating that the use of factor analysis was appropriate. The four factors each had eigenvalues greater than 1 and explained 65.01% of the total variance.

**Criterion Validity**

Comparing CASQ-measured sleepiness between the two chronotypes revealed that, consistent with the literature,\textsuperscript{10–12} evening-types demonstrated higher levels of daytime sleepiness on the CASQ overall score and all subscale scores (Table 1).

**Confirmatory Factor Analysis (CFA)**

The CFA model demonstrated a good fit of the data to the CASQ model in terms of each of the five summary metrics (Table 2; Figure 1).

| Table 1 Sleepiness Differences by Chronotype |
|---------------------------------------------|
| ![Image of Table 1](https://via.placeholder.com/150) |

*Abbreviations: E, evening; M, morning.*
The CASQ has been widely used in the area of sleep medicine, children’s health care, nursing research, and psychology and education. The purpose of this work was to translate, adapt, and evaluate the psychometric properties of a Chinese version of the CASQ (C-CASQ) using a sample of adolescents from Taiwan. We found that the C-CASQ had good psychometric properties and was consistent in measuring sleepiness differences in morning and evening chronotypes.

The Cronbach’s α for the original English CASQ version was 0.89, while the Cronbach’s α for the C-CASQ designed and evaluated in this study was 0.85. Furthermore, no items from the original CASQ scale were removed, indicating that the scales have similar psychometric properties for different populations.

### Discussion

The CASQ had been widely used in the area of sleep medicine, children’s health care, nursing research, and psychology and education. The purpose of this work was to translate, adapt, and evaluate the psychometric properties of a Chinese version of the CASQ (C-CASQ) using a sample of adolescents from Taiwan. We found that the C-CASQ had good psychometric properties and was consistent in measuring sleepiness differences in morning and evening chronotypes.

The Cronbach’s α for the original English CASQ version was 0.89, while the Cronbach’s α for the C-CASQ designed and evaluated in this study was 0.85. Furthermore, no items from the original CASQ scale were removed, indicating that the scales have similar psychometric properties for different populations.

### Table 2 Results of Confirmatory Factor Analysis

| Measure   | Results         | Recommended Cutoff |
|-----------|-----------------|--------------------|
| $\chi^2$/df | 2.215           | < 5                |
| RMSEA     | 0.068           | < 0.08             |
| CFI       | 0.918           | > 0.90             |
| TLI       | 0.902           | > 0.90             |
| SRMR      | 0.066           | < 0.08             |

**Abbreviations:** RMSEA, root mean square error of approximation; CFI, Comparative Fit Index; TLI, Tucker–Lewis Index; SRMR, standardized root mean square residual.

### Figure 1 Results of confirmatory factor analysis for the C-CASQ.
parameter estimates for the relationships among the CASQ factors provide confirmatory evidence that the four factors reflect separate, yet conceptually related constructs. Thus, the C-CASQ scale developed in this study has the potential to allow researchers to meaningfully compare the responses of individuals from different cultures using two versions (languages) of the same scale.

Our study does have several limitations. Age and gender were considered as potential confounding variables but as with all studies, especially those that are not randomized, confounding variables could have obscured an underlying association. As such, it is possible that there is either a sex or age difference in our translation that went undetected. Future studies would benefit by examining the factor structure of the C-CASQ among different populations, determining whether different structures exist based on different sociocultural contexts or clinical samples. Moreover, the test-retest reliability of the scale can be further established. Although the internal consistency of the C-CASQ was adequate, further studies using data collected at different time points to assess the reliability of the scale across time intervals are required.

The differences in subjective and objective sleep measurements have been increasingly discussed. An important limitation in the current study that could be rectified in future studies would be to establish stronger criterion convergent validity by examining the relationship between the construct of subjective sleepiness, which could be measured by the C-CASQ, and other subjective or objective measures of sleepiness. As there is likely to be a degree of non-overlap in the variance explained by different measures of sleepiness, it will be important to understand how much is due to culture or language and how to adequately address these differences using properly designed questionnaires.

**Conclusion**

In sum, the newly translated Chinese version of the CASQ is deemed to be a feasible and acceptable instrument that satisfies the intentions of the study aiming to measure the extent of sleepiness in Chinese adolescent populations. This study has presented statistical evidence that the C-CASQ translation is accurate in a linguistic sense, as well as reliable and valid. It has psychometric features that make it suitable for use as a standardized test of sleepiness for screening or clinical trials among Chinese adolescents. The C-CASQ provides a simple, inexpensive method for monitoring adolescent responses in order to evaluate daytime sleepiness and sleep-related problems, such as insomnia.

**Abbreviations**

BT1 … 3, back translation version 1 … 3; CASQ, Cleveland Adolescent Sleepiness Questionnaire; C-CASQ, Chinese version of the Cleveland Adolescent Sleepiness Questionnaire; CFA, confirmatory factor analysis; CFI, Comparative Fit Index; EFA, exploratory factor analysis; ESS, Epworth Sleepiness Scale; MESC, Morningness-Eveningness Scale for Children; PDSS, Pediatric Daytime Sleepiness Scale; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; T1 … 4, translation version 1 … 4; TLI, Tucker–Lewis Index.

**Acknowledgments**

We thank Director Cheng-Yu Lin from the Sleep Center of the National Cheng Kung University Hospital who provided translation expertise.

**Author Contributions**

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

**Funding**

This study was supported by the Ministry of Science and Technology, Taiwan, R.O.C. [grant #106-2918-I-006-002, YTY]. The funding agency played no role in the design of the study, the collection, analysis, and interpretation of data, or in the writing of the manuscript.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**

1. Mei X, Zhou Q, Li X, Jing P, Wang X, Hu Z. Sleep problems in excessive technology use among adolescent: a systemic review and meta-analysis. Sleep Sci Pract. 2018;2:9. doi:10.1186/s41606-018-0028-9

2. Short MA, Booth SA, Omar O, Ostlundh L, Arora T. The relationship between sleep duration and mood in adolescents: a systematic review and meta-analysis. Sleep Med Rev. 2020;52:101311. doi:10.1016/j.smrv.2020.101311

3. Short MA, Weber N. Sleep duration and risk-taking in adolescents: a systematic review and meta-analysis. Sleep Med Rev. 2018;41:185–196. doi:10.1016/j.smrv.2018.03.006
