Measurement Properties of the Chinese Weinstein Noise Sensitivity Scale

Daniel Y.T. Fong1, Naomi Takemura1, Pui-Hing Chau1, Sarah L.Y. Wan2, Janet Y.H. Wong1

1School of Nursing, The University of Hong Kong, 2Department of Psychological Studies and Centre for Psychosocial Health, The Education University of Hong Kong, Hong Kong

Abstract

Context: Noise sensitivity may mediate or moderate the influences of noise exposure on health, and it needs to be reliably evaluated. The 21-item Weinstein’s noise sensitivity scale (WNSS) has been the most popular scale for assessing noise sensitivity, but it is not yet available in traditional Chinese. Aims: This study aimed to conduct linguistic and psychometric performance of WNSS in Hong Kong (HK) Chinese. Settings and Design: A population-based telephone survey with 1-week follow-up. Materials and Methods: The HK Chinese WNSS was obtained after forward–backward translation from the original English version and cognitive debriefing in five Chinese adults. Its measurement properties were assessed in 569 Chinese adults aged 18 years or above. Statistical Analysis Used: The sample was randomly split into two halves. The first half was used to explore a scale structure of the WNSS by exploratory factor analysis (EFA) with the number of factors determined by the optimal comparison data technique and tested for being artifactual. The second half was used for confirmatory factor analysis. Convergent validity and test–retest validity were also assessed. Results: EFA identified two unipolar factors and removed three items poorly associated with the factors. The factors were likely artifactual and a unidimensional structure was assessed in CFA, which showed a satisfactory fit (root mean square error of approximation = 0.055; comparative fit index = 0.904; standardized root mean square residual = 0.061). The HK Chinese WNSS had good internal consistency (Cronbach’s α = 0.83) and test–retest reliability (intraclass correlation coefficient = 0.87). Furthermore, it confirmed the expected association with extraversion (r = −0.14, P < 0.001) and neuroticism (r = 0.28, P < 0.001). Conclusion: The 18-item HK Chinese WNSS was reliable and valid for assessing noise sensitivity in the Chinese population.

Keywords: Factor analysis, noise sensitivity, psychometrics

INTRODUCTION

Noise refers to sound that carries a negative connotation. Regardless of whether noise is wanted or not, it surrounds us every day in the environment, such as noises from transport vehicles, construction, community, and social activity. [1] There have been epidemiological and experimental studies supporting that prolonged exposure to noise may not only damage the sensitive structures in our inner ear and cause hearing loss, but it may also have nonauditory health effects, including cognitive impairment, sleep disturbance, annoyance, and cardiovascular diseases. [1] Noise sensitivity has been conceptualized as a stable, subjective attribute independent of noise exposure, but it influences personal reactions to environmental noise. [2] It has been noted that the same degree of exposure to noise in two individuals may not necessarily result in the same degree of annoyance or other nonauditory health effects. [3] Individuals who are less sensitive to noise are often less annoyed than individuals who are more sensitive to noise. Thus, noise sensitivity may mediate or moderate the influences of noise exposure on health. For example, noise sensitivity has been shown to be associated with hypertension, chest pain, and noise-induced sleep disturbance. [4,5]

Hence, noise sensitivity needs to be reliably evaluated before the noise-induced health effects can be adequately assessed.

Address for correspondence: Daniel Y.T. Fong, School of Nursing, The University of Hong Kong, 21 Sassoon Road, Pokfulam, Hong Kong, Hong Kong. E-mail: dytfong@hku.hk

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work noncommercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

How to cite this article: Fong DY, Takemura N, Chau PH, Wan SL, Wong JY. Measurement properties of the chinese weinstein noise sensitivity scale. Noise Health 2017;19:193-9.
Subjective noise sensitivity can be measured by using a self-reported questionnaire. Several scales for assessing noise sensitivity have been developed, but the most popular and thoroughly tested one is the Weinstein’s noise sensitivity scale (WNSS).[3] The WNSS is a 21-item scale, with each item responded on a 6-point (0–5) Likert scale ranging from “disagree strongly” to “agree strongly”. Higher agreement on a statement indicates higher noise sensitivity. Since the development of the WNSS, it has been translated and validated into different languages, including German,[8] Italian,[9] Japanese,[9] Persian,[10] simplified Chinese,[11] and Swedish.[12] However, the WNSS has not been translated and validated in the traditional Chinese context. Moreover, a one-factor structure was found except for the Italian and Persian versions. The Italian version identified two unipolar factors, that is, one factor comprised all positively worded items and the other factor included all negatively worded items, which may be likely due to method effects rather than driven by the underlying constructs. For the Persian version, four factors were identified but its appropriateness for the Chinese version was uncertain. Therefore, this study aimed to translate the WNSS into Hong Kong (HK) Chinese and assess its measurement properties in a Chinese population.

Materials and methods

Participants

Five-hundred Chinese patients who were 18 years old or above and could understand traditional Chinese were planned to be recruited by a telephone call by using random digit dialing. Informed consent was sought orally before an eligible patient participated in the study. Each participant was invited to respond to a battery of questionnaires twice about 1 week apart. Ethics approval for the study protocol and the consent process were sought from a local institutional review board.

The sample size was tailored to meet the needs of conducting the exploratory and confirmatory factor analyses on two random halves of the data. For each factor analysis, we adopted the usual rule of 10 patients per item. Thus, the 21-item WNSS required a total of 420 patients. Accounting for a small portion of unexpectedly unusable questionnaires, we planned to recruit 500 patients.

Measures

The HK Chinese 21-item WNSS was obtained from the original English version by standard forward–backward translation. Specifically, two bilingual persons independently translated the WNSS into traditional Chinese. The two forward Chinese versions were discussed in a consensus meeting among the two translators and a researcher with prior experience of the cultural adaption process. A consensus of Chinese version was obtained and back translated into English by a third translator who was unaware of the original English version. The two English versions were compared, and on the basis of this comparison, a revised Chinese version was obtained and assessed for clarity in five individuals who were aged at least 18 years and could read traditional Chinese. A small revision was made and the final HK Chinese version was obtained for psychometric evaluation.

The extraversion (12 items) and neuroticism (12 items) scales of the Chinese NEO-Five Factor Inventory (NEO-FFI) were also administered. Each item was responded on a 5-point Likert scale from 0 = strongly agree to 4 = strongly disagree. The two scales were used for assessment as they had been previously shown to be associated with noise sensitivity.[14,15]

In addition, age, gender, marital status, and educational level were also assessed.

Statistical analyses

Sample characteristics were summarized by descriptive statistics. To determine the scale structure of the Chinese WNSS, the sample was randomly split into two halves. The first half served as a training set in which an exploratory factor analysis (EFA) with the oblimin oblique rotation was performed to identify a scale structure of the items. Factor loadings were obtained by principal axis factoring, with the number of factors determined by the comparison data technique that has been shown to outperform other existing approaches in terms of accuracy and robustness.[16] Items with all factor loadings less than 0.4 in magnitude were removed. To assess the robustness of the identified scale structure, the EFA was repeated by using maximum likelihood estimation. We identified two unipolar factors, that is, one factor comprised positively worded items and the other represented by negatively worded items, which can be artifactual rather than representing two distinct underlying constructs. This may have happened as the assumption that people who agree on items of one type would most likely disagree on items of the opposite type is violated. We assessed this assumption through multiple approaches.[17] First, for each respondent, we obtained the percentage of positively worded items having an “agree” response and the percentage of negatively worded items having a “disagree” response. If the assumption holds, the distribution of the percentage differences would be unimodal, symmetric, and highly dense at 0. Second, we assessed the normality of items through the Shapiro–Wilk test and normal probability plots. Any deviations from normality would indicate the possibility of an artifact. Third, we examined if the rank of the correlation matrix of the items was larger than the double-centered data matrix by 1, which again would imply a possible artifact.[18]

The second half served as a validation set in which a confirmatory factor analysis (CFA) was performed to assess the fitness of the scale structure identified from EFA.[19] Assessment of goodness-of-fit was based on the fit indices: the root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), and the comparative fit index (CFI). They covered parsimony...
correction, absolute fit, and comparative fit, as recommended. A CFA model was considered reasonable when RMSEA was close to 0.06 or lower, SRMR was close to 0.08 or lower, and CFI were close to 0.95 or greater. When there was inadequate fit, error covariance with the largest modification index and substantive rationale was incorporated. Appropriateness of the Persian four-factor structure that comprised the scales: becoming sensitive to noise (items Q2, Q4–Q7, Q10, Q13, Q18, Q19, and Q21), disturbance in concentration (items Q8, Q11, Q13–Q15, and Q18), attitude to noise in where they live (items Q1, Q2, Q8, Q9, and Q12), and attitude to noise control (items Q3, Q13, Q16, Q17, Q20, and Q21) were also assessed by CFA.

By using the full dataset, each participant had the WNSS scored as the average of the observed item responses, standardized onto the 0 to 100 scale, when at least 50% of the items received responses. The scaling properties of the HK Chinese WNSS were assessed by calculating the floor percentage, ceiling percentage, and Cronbach’s α. Convergent validity was assessed by Spearman’s rank correlation between the HK Chinese WNSS and the extraversion and neuroticism scales of the NEO-FFI. Test–retest reliability was assessed by the intraclass correlation coefficient (ICC).

The nominal level of significance was 5% in all significance tests, and each estimate was accompanied by a 95% confidence interval (CI), where appropriate. The R package (R Foundation, Vienna, Austria) and SPSS software (IBM Corp., Armonk, New York, United States) were used for the analysis.

Results
We recruited 569 Chinese adults of mean age 37 years (range: 18–91). Their characteristics are summarized in Table 1. Sixteen participants did not respond to one or two items of the 21-item Chinese WNSS, and the remaining 553 participants were randomly split for EFA and CFA.

The first half comprised 287 participants, based on which, the comparison data technique identified that two factors would be most appropriate. They were extracted by principal axis factoring, which explained 20 and 8% of total variance for the first and second factors, respectively. Three items (Q8, Q9, and Q15) had all rotated factor loadings smaller than 0.4 with communalities less than 0.1 and were removed. The first factor included all items worded toward noise sensitive (positively worded), whereas the second factor comprised all items worded toward insensitive to noise (negatively worded). The use of maximum likelihood estimation identified the same factor structure. The distribution of the difference between the agree percentage among the positively worded items and that among the negatively worded items is shown in Figure 1. Both the normal probability plot and Shapiro–Wilk test \(P = 0.183\) did not show deviation from normality. However, respondents were more likely to agree on positively worded items than to disagree on negatively

### Table 1: Sample characteristics of 569 subjects

| Age, mean (SD) in years | 37 (13) |
|-------------------------|--------|
| Gender                  |        |
| Male                    | 208    | 36.6 |
| Female                  | 361    | 63.4 |
| Marital Status          |        |
| Single                  | 228    | 40.1 |
| Cohabitated             | 14     | 2.5  |
| Married                 | 305    | 53.7 |
| Widowed                 | 11     | 1.9  |
| Divorced                | 10     | 1.8  |
| Education Level         |        |
| Primary or below        | 50     | 8.8  |
| Secondary               | 259    | 45.5 |
| Associate/Bachelor      | 222    | 39.0 |
| Master/Doctoral         | 38     | 6.7  |
| Occupation              |        |
| Managers and executives | 28     | 4.9  |
| Professionals           | 68     | 12.0 |
| Associate Professionals | 35     | 6.2  |
| Clerical support workers| 88     | 15.5 |
| Service and sales workers| 78   | 13.7 |
| Craft and related workers| 6      | 1.1  |
| Plant and machine operators and assemblers| 20 | 3.5 |
| Elementary              | 13     | 2.3  |
| Housewives              | 91     | 16.0 |
| Retired                 | 63     | 11.1 |
| Job seeking             | 15     | 2.6  |
| Students                | 52     | 9.2  |
| Others                  | 11     | 1.9  |
| Monthly household income (HK$) | | |
| $<5,000                 | 50     | 8.9  |
| $5,000–9,999            | 20     | 3.6  |
| $10,000–14,999          | 57     | 10.2 |
| $15,000–19,999          | 57     | 10.2 |
| $20,000–24,999          | 60     | 10.7 |
| $25,000–29,999          | 47     | 8.4  |
| $30,000–34,999          | 63     | 11.3 |
| $35,000–39,999          | 40     | 7.2  |
| $40,000–44,999          | 58     | 10.4 |
| $45,000–49,999          | 37     | 6.6  |
| $50,000–59,999          | 70     | 12.5 |
| Household size, median (range) | 3 (1–9) | |
| Medical history         |        |
| None                    | 474    | 83.4 |
| Heart disease           | 4      | 0.7  |
| Hypertension            | 38     | 6.7  |
| Diabetes mellitus       | 17     | 3.0  |
| High cholesterol        | 23     | 4.0  |
| Stroke                  | 4      | 0.7  |
| Obese                   | 2      | 0.4  |
| Renal disease           | 1      | 0.2  |
| Gastrointestinal disease| 11     | 1.9  |
| Mood disorder           | 1      | 0.2  |
| Mental disorder         | 2      | 0.4  |
| Cancer                  | 3      | 0.5  |

(Continued)
worded items by 13.3% (95% CI = 9.2–17.6%, \( P < 0.001 \) by \( t \) test). All items revealed significant deviations from normality (\( P < 0.01 \) by Shapiro–Wilk test). Responses to positively worded items were generally left skewed, whereas those to negatively worded items did not show much skewness. The rank of the item correlation matrix was 21, whereas that of the double-centered data matrix was 20. Therefore, the two-factor structure was likely artifactual and a unidimensional 18-item structure was adopted for further analysis.

By using the second half of 266 participants, fit indices of the one-factor 18-item HK Chinese WNSS CFA model and other models are presented in Table 2. The one-factor 18-item model showed the best overall fit among the other CFA models. Its unstandardized coefficients (\( P < 0.007 \)) are shown in Figure 2. The one-factor structure of 21 items Chinese WNSS and the four-factor structure of the 21-item Persian WNSS did not show a satisfactory fit overall (Table 2).

From the full sample of 569 participants, the mean 18-item Chinese WNSS score was 50.1 (SD = 12.8, range = 21.1–96.7). The normal probability plot did not reveal a substantial departure from normality. The instrument had 0% floor and ceiling effects, with a Cronbach’s \( \alpha \) of 0.83 (95% CI = 0.81–0.85). The corrected item total correlation ranged between 0.24 and 0.71. A total of 311 participants were successfully followed at 1 week; based on which, the ICC was 0.87 (95% CI = 0.85–0.89). In addition, the 18-item Chinese WNSS was significantly associated with both extraversion (\( r = -0.14, P < 0.001 \)) and neuroticism (\( r = 0.28, P < 0.001 \)).

**DISCUSSION**

We conducted rigorous linguistic validation of the HK Chinese WNSS and utilized a territory-wide sample of 569 individuals to assess its measurement properties. The sample size was larger than the next largest study of 413 individuals.[2,6] A unidimensional 18-item scale was identified and validated in an independent sample. It possessed a satisfactory scale structure, good internal consistency, high test–retest reliability, and acceptable convergent validity.

We removed three items from the original 21-item version of WNSS. They are “Q8. I get used to most noises without much difficulty,” “Q12. It wouldn’t bother me to hear the sounds of everyday living from neighbors (footsteps, running water, etc.),” and “Q15. In a library, I don’t mind if people carry on a conversation if they do it quietly.” For item Q8, one may have considered the type of noise the person is living with which they may have gotten used to irrespective of the degree one tolerates the noise. Item Q12 may not be highly applicable in the HK setting. Most apartment units in HK are separated by concrete walls and floors, and most people would not hear the footsteps and running water from neighbors. Item 15 can be a mixture of responding to the degree one tolerates the noise and compliance to the library rule that everyone should keep quiet. Consequently, the three items did not have substantial association with the others that were more indicative of noise sensitivity.
Although the WNSS was originally developed as a unidimensional tool to assess noise sensitivity, two- and four-factor structures of the WNSS have been previously identified.\[2,10\] The difference in factor structure can be a result of using a suboptimal method to determine the number of factors or the presence of artifactual factors. Most, if not all, previous studies assessed the scree plot for determining the number of factors, which has been the most common approach to determine the number of factors in an EFA since it was developed.\[23\] However, a recent simulation study has shown that the scree plot identified the correct number of factors only 42% of the time.\[24\] As a result, other approaches have been developed, among which, we adopted the comparison data approach that has been shown to have the highest correct-identification rate at 87%.\[16,24\] Indeed, our CFA confirmed the four-factor structure had an inadequate fit. Having said that, factors that have been correctly identified can still be artifactual.\[17\] The two unipolar

Figure 2: Path diagram with unstandardized coefficients of the 1-factor structure of the 18-item Chinese Weinstein Noise Sensitivity Scale
factors in the Chinese WNSS are also found in the Italian version.\textsuperscript{[21]} The Italian WNSS was, however, concluded as unidimensional on the ground that the two factors were highly associated. As highly correlated factors may also imply a bifactor model that still incorporates a two-factor structure and there was no conceptual basis for two unipolar factors of the WNSS, we used multiple statistical procedures to conclude that the unipolar factors were indeed likely to be artifactual. In addition, the first factor in EFA explained substantially more of the total variance (20\% vs. 8\%). This was also observed in previous studies that reported 21 to 27\% for the first factor and 5 to 8\% for the second factor.\textsuperscript{[2,10-12]} Hence, a unidimensional structure for the Chinese WNSS is deemed to be appropriate.

Error covariances between items were identified due to item wordings. Specifically, items Q1 and Q3 are both worded toward noise insensitive with a stronger agreement corresponding to a lower noise sensitivity score, whereas other items were worded toward noise sensitive with a stronger agreement corresponding to a higher noise sensitivity score. The other items shared similar contexts. For example, item Q5 that asks about the ease of awakening by noise and item Q9 that asks about the location of apartment share the context of the environmental consideration of where one is living. Such error covariation due to methods effects was considered in our CFA, which however cannot be catered in EFA. The HK Chinese WNSS had satisfactory internal consistency of 0.82, which lies within the currently reported range of 0.78 to 0.87 in other language versions.\textsuperscript{[2,10-12]} It is not too high to indicate item redundancy. Moreover, the test–retest reliability of 0.87 is also high and comparable with those of other language versions that were reported in the range of 0.66 to 0.87.\textsuperscript{[8,10,11]}

The 18-item HK Chinese WNSS had no obvious departure from normality that facilitates the use of standard statistical methods in the analysis. Moreover, it had a norm value of 50, the middle of the standardized scale of 0 to 100, without any ceiling and floor effects. This makes the scores of the instrument more interpretable. Moreover, the different language versions of the WNSS have varied lengths, which may possibly be a result of cultural differences.\textsuperscript{[2,6,10,12]} Standardizing the WNSS scores on the 0 to 100 scale may facilitate cross-cultural comparisons, although further work on measurement in variance across different cultural groups would be desirable.

Despite our efforts to ensure that the study was properly designed and conducted, a number of limitations deserve attention. First, we have not attempted to reduce the length of the HK Chinese WNSS, which would preferably be explored by detailed item analysis. Second, we have not assessed noise annoyance, and hence its association with the HK Chinese WNSS has not been ascertained. Further study would be desirable.

Conclusively, the 18-item HK Chinese WNSS is reliable and valid to assess the noise sensitivity of the Chinese population. Further item analysis to develop a shorter version would be desirable.

Financial support and sponsorship

Small Project Funding, The University of Hong Kong (Project number: 104002891).

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. WHO. Burden of disease from environmental noise: Quantification of healthy life years lost in Europe. In: Theakston F, editor. Copenhagen: WHO Regional Office for Europe; 2011 p. 106.
2. Senese VP, Ruotolo F, Ruggiero G, Iachini T. The Italian Version of the Weinstein noise sensitivity scale measurement invariance across age, sex, and context. Eur J Psychol Assess 2012;28:118-24.
3. Weinstein ND. Individual differences in reactions to noise: A longitudinal study in a college dormitory. J Appl Psychol 1978;63:458-66.
4. Fyhri A, Klaeboe R. Road traffic noise, sensitivity, annoyance and self-reported health – A structural equation model exercise. Environ Int 2009;35:91-7.
5. Aasvang GM, Moun T, Engdahl B. Self-reported sleep disturbances due to railway noise: Exposure-response relationships for nighttime equivalent and maximum noise levels. J Acoust Soc Am 2008;124:257-68.
6. Kishikawa H, Matsu T, Uchiyama I, Miyakawa M, Hiramatsu K, Stansfeld SA. The development of Weinstein’s noise sensitivity scale. Noise Health 2006;8:154-60.
7. Schutte M, Marks A, Wenning E, Griefahn B. The development of the noise sensitivity questionnaire. Noise Health 2007;9:15-24.
8. Zimmer K, Ellermeier W. Psychometric properties of four measures of noise sensitivity: A comparison. J Environ Psychol 1999;19:295-302.
9. Kishikawa H, Matsu T, Uchiyama I, Miyakawa M, Hiramatsu K, Stansfeld SA. Noise sensitivity and subjective health: Questionnaire study conducted along trunk roads in Kusatsu, Japan. Noise Health 2009;11:111-7.
10. Alimohammadi I, Nassiri P, Azkhosh M, Sabet M, Hosseini M. Reliability and validity of the Persian translation of the Weinstein noise sensitivity scale. Psychol Res 2006;9:74-87.
11. Han T, Wu J. Revise of the noise sensitive scale and its reliability and validity. China J Health Psychol 2015;23:196-200.
12. Ekehammar B, Dornic S. Weinstein’s noise sensitivity scale: Reliability and construct validity. Percept Mot Skills 1990;70:129-30.
13. Costa PT, McCrea RR. Psychological Assessment Resources Inc. Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI). Odessa, FL: Psychological Assessment Resources Inc. 1992.
14. Dornic S, Ekehammar B. Extraversion, neuroticism, and noise sensitivity. Pers Individ Dif 1990;11:989-92.
15. Campbell JB. Extraversion and noise sensitivity: A replication of Dornic and Ekehammar’s study. Pers Individ Dif 1992;13:953-5.
16. Ruscio J, Roche B. Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. Psychol Assess 2012;24:282-92.
17. Spector PE, Katwyk PTV, Brannick MT, Chen PY. When two factors don’t reflect two constructs: How item characteristics can produce artifactual factors. J Manage 1997;23:659-77.
18. van Schuur W, Kiers HAL. Why factor analysis often is the incorrect model for analyzing bipolar concepts, and what model to use instead. Appl Psych Meas 1994;18:97-110.
19. Fong DY, Lam CL, Mak KK, Lo WS, Lai YK, Ho SY. et al. The Short Form-12 Health Survey was a valid instrument in Chinese adolescents. J Clin Epidemiol 2010;63:1020-9.
20. Hu LT, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychol Methods 1998;3:424-53.
21. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Model 1999;6:1-55.

22. Benfield JA, Nurse GA, Jakubowski R, Gibson AW, Derrick TB, Newman P, et al. Testing noise in the field: A brief measure of individual noise sensitivity. Environ Behav 2012;1-20.

23. Cattell RB. The scree test for the number of factors. Multivariate Behav Res 1966;1:245-76.

24. Courtney MGR. Determining the number of factors to retain in EFA: Using the SPSS R-Menu v2.0 to make more judicious estimations. Pract Assess Res Eval 2013;18:1-14.
