L'Encéphale xxx (xxxx) xxx–xxx

Disponible en ligne sur
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com

Original article

French Sleepiness Scale for Adolescents-8 items: A discriminant and diagnostic validation

Échelle française de somnolence pour adolescents : une validation discriminante et diagnostique

suppl. Informations

M.-P. Gustin[a,b], B. Putois[c], A. Guyon[d,e,f], M. Lecendreux[e,g], M.-J. Challamel[d], S. Plancoulaine[h], S. Bioulac-Rogier[i], C. Schroder[j], S. Royant-Parola[k], S. Huguelet[c], P. Franco[d,e,f,*]

[a] *Emerging Pathogens Laboratory–Fondation Mérieux, International Center for Infectiology Research (CIRI), Inserm U1111, CNRS UMR5308, ENS de Lyon, Lyon, France*
[b] *Institute of Pharmaceutic and Biological Sciences, Public Health department, Biostatistics, University Claude-Bernard Lyon 1, Villeurbanne, France*
[c] *Faculty of Psychology, Swiss Distance Learning University, Brig 3900, Switzerland*
[d] *Pediatric Sleep Unit, Hospital for Women Mothers & Children, Lyon 1 University, France*
[e] *National Reference Centre for Orphan Diseases, Narcolepsy, Idiopathic hypersomnia and Kleine-Levin Syndrome (CNR narcolepsie-hypersomnie), France*
[f] *Lyon Neuroscience Research Centre, CNRS UMR 5292-INSERM U1028-Lyon 1 University, Bron 69005, France*
[g] *Centre pédiatrique des pathologies du sommeil, Hôpital Robert Debré, Paris, France*
[h] *Université de Paris Cité, Inserm, INRAE, CRESS, 75004 Paris, France*
[i] *Service de Psychiatrie de l'enfant et l'adolescent, Hôpital Couple enfant, CHU Grenoble Alpes CHU CS 10217, 38046 Grenoble, France*
[j] *University of Strasbourg; CNRS UPR 3212, Institute of Cellular and Integrative Neurosciences*
[k] *Réseau Morphée, Garches, France*

ARTICLE INFO

Article history:
Received 22 February 2022
Accepted 8 June 2022
Available online xxx

Keywords:
French Sleepiness Scale
Sleepiness
Adolescent
Narcolepsy

ABSTRACT

The objective of the present study was to validate the Short Version of French Sleepiness Scale for Adolescents (FSSA) with eight items (FSSA8).
Methods. – A total of 384 adolescents, aged between 12 and 18 years, completed the FSSA8. These included 269 nonclinical adolescents and 115 adolescents admitted for overnight polysomnography and Multiple Sleep Latency Test (MSLT) because of suspected hypersomnia (85 patients with narcolepsy and 30 with other sleep disorders). Item response theory (IRT) assumptions were tested and psychometric properties were analysed. Matching on sex ratio and age was conducted to estimate concurrent criterion, diagnostic validity and cut-offs.
Results. – IRT assumptions were validated confirming the one-dimensionality of the FSSA8. The latent continuum sleepiness for which the scale and its items are reliable encompassed most of the clinical subjects. FSSA8 is weakly correlated with MSLT. Distribution of scores for the nonclinical group and the clinical group differed significantly; the FSSA8 had very good screening validity in sleep disorders. The cut-off was seven points.
Conclusion. – The FSSA8 appeared to be more reliable for patients than for nonclinical participants and to be a good tool for screening excessive daytime sleepiness in sleep disorders.

Mots clés :
Echelle française de somnolence
Somnolence
Adolescent
Narcolepsie

RÉSUMÉ

La somnolence diurne est un indicateur important de la qualité et de la quantité du sommeil. Qu'il s'agisse de dépistage de pathologies rares comme l'hypersomnie ou de troubles plus fréquents comme le syndrome de retard de phase, la mesure de la somnolence est indispensable pour l'investigation clinique ou scientifique chez l'adolescent. A cette période de la vie, la somnolence diurne impacte les apprentissages, la scolarisation et le risque d'accidents sur la voie publique. L'objectif de cette étude est de tester les propriétés psychométriques de l'échelle française de la somnolence

* Corresponding author at: Pediatric Sleep Unit, Hôpital Femme Mère Enfant, Hospices Civils de Lyon, 59, boulevard Pinel, 69500 Lyon, France.
E-mail address: patricia.franco@univ-lyon1.fr (P. Franco).

https://doi.org/10.1016/j.encep.2022.06.004
0013-7006/© 2022 L'Encéphale, Paris.

pour les adolescents (*French Sleepiness Scale for Adolescents*, FSSA) en vue de discriminer la somnolence normale et pathologique.
Méthode. – La version française de la FSSA a été adaptée pour correspondre aux versions étrangères (8 items). 384 adolescents, âgés entre 12 et 18 ans, ont renseigné la FSSA8 dont 269 hors-consultations et 115 lors d'une consultation hospitalière de sommeil avec polysomnographie et des tests itératifs de latence à l'endormissement (TILE, mesure objective de la somnolence diurne) dans le cadre d'une suspicion d'hypersomnie. 85 patients présentaient une narcolepsie et 30 d'autres troubles du sommeil. Les hypothèses de la théorie de la réponse aux items (i.e. *Item Response Theory*, IRT) ont été testé et les qualités psychométriques ont été analysées. Un appariement genre-âge a été réalisé pour estimer la validité diagnostique de l'échelle et pour déterminer des scores seuils.
Résultats. – Les hypothèses de l'IRT ont été validées confirmant la structure uni-factorielle de la FSSA8. L'échelle présente une bonne sensibilité et spécificité surtout pour les sujets cliniques bien qu'elle soit faiblement corrélée avec les mesures objectives de la somnolence (TILE). La distribution des scores pour le groupe non-clinique et le groupe clinique était significativement différente; la FSSA8 a une très bonne validité diagnostique pour les troubles du sommeil. Le seuil pathologique est de 7.
Conclusion. – Le FSSA8 est un outil sensible dans le contexte de trouble du sommeil, et fiable pour le dépistage de la somnolence diurne excessive dans le cadre de pathologies du sommeil dans la population générale. Elle est cependant peu fiable chez les adolescents ne présentant aucun trouble du sommeil.

1. Introduction

In adolescents, excessive daytime sleepiness (EDS) is common and is associated with a decrease in academic achievement [1]. Insufficient duration of sleep is the main underlying cause of sleepiness at this age [2], which increases risk of drowsy driving, obesity, depression, and suicidal ideation [3]. Suicide and road accident are the main causes of mortality in adolescents. EDS is also a frequent symptom of several sleep disorders, including sleep-disordered breathing (SBD) [4], delayed sleep phase syndrome (DSPS) [5], but also central hypersomnia such as narcolepsy [6]. Some pathological conditions could be also intricated Indeed, obesity, a frequent comorbidity in narcoleptic patients, is also the most frequent cause of sleep disordered breathing during adolescence [7]. Precise evaluation of daytime sleepiness and screening for EDS are crucial in this age population.

In adults, the questionnaire most frequently used to evaluate the level of sleepiness is the Epworth Sleepiness Scale (ESS), developed by Johns [8]. The ESS proposes eight more and less soporific situations for which subjects are asked to rate how likely they would doze off or fall asleep on a four-tier scale. The ESS score is the sum of the eight item scores, and ranges from 0 to 24. The ESS is widely used in adults with sleep disorders associated with EDS [9]. Studies that have analysed its psychometric properties in adults are numerous [10–15]. The French version of the ESS in adults has been validated for internal consistency, test-retest reliability, construct, and criterion validity [16].

In adolescents or children, the most frequently used scale assessing EDS is also the ESS [17]; it is used for 20 years with [18–27] or without [1,28,29] modification of the original scale. Modified versions of ESS replaced some appropriate items for children and adolescents. For instance, item 8, "in a car while stopped for a few minutes in the traffic", was replaced with "during class at school" [26] or "doing homework or taking a test" [19] or "sitting and riding in a car or bus for about half an hour" [27].

An adapted version of ESS for children and adolescents from Snow et al. [26] was validated in French in a non-clinical population [30]: the French Sleepiness Scale for Adolescents (FSSA). Until now, there is no study evaluating in the same work nonclinical and narcoleptic adolescent patients to establish external validity and accurate cut-off points. Indeed,

an Australian (ESS-CHAD) and a Persian version have been validated in general population that included 297 and 1371 adolescents respectively [27,31]. Recently, the ESS-CHAD has been validated in treated narcoleptic patients with cataplexy [32]. The only study that computed a cut-off was conducted among children between 3 and 12 years of age with high apnea-hypopnea index [18].

Since 2009, in French pediatric sleep centers, we routinely administer the FSSA coupled with polysomnography and medical consultation to adolescents coming for potential sleep problems [33–35]. We excluded two items of the FSSA (FSSA8) in order to have the same number of items than the other versions (Appendix 1). We decided to exclude the two non-soporific conditions 1) playing outside with his/her friends or doing exercise and 2) playing alone on computer or video games. The aim of this work is to estimate psychometric properties of FSSA8, in particular its cut-off in pathological population.

2. Method

2.1. Study design

This study was designed as a cross-sectional survey for nonclinical participants and a retrospective cohort study for patients. Since the patients' data were collected retrospectively the analyses were performed without identification, approval by local ethics committee was not necessary. The parents have given their consent for their child's data to be used for research purposes. The methodology respected sound ethical practice (Data Protection Authorization: CNIL-n°19-087).

2.2. Participants

The cohort comprised two main groups of adolescents between 12 and 18 years old; i) the nonclinical group, was composed of adolescents attending partner schools provided by another study [36], and ii) the clinical group was composed of adolescents with EDS referred to the Narcobank study [37]. This clinical group was subdivided into two subgroups i) narcoleptic patients (Narcoleptic Group), and ii) patients suspected of hypersomnia with other sleep disorders (insomnia, phase delay syndrome, bad sleep hygiene, or obstructive sleep apnea (OSA) (Mixed Group). Then, we performed a matched screening on age and sex and obtained adolescents in the control group and a clinical matched group including patients with narcolepsy and with other sleep disorders. The patients did not receive any treatment. Body mass index (BMI) could not be

considered because it was missing in 72% of the whole dataset, essentially in the nonclinical group.

2.3. Diagnosis procedure

The clinical group underwent a systematic interview with a sleep specialist. Sleep and wake procedure were done and diagnoses were established according to the criteria of the third edition of the International Classification of Sleep Disorders [38] (technical details are provided in supplementary materials, page 1). No patients were treated at the time of the evaluation.

2.4. Measures

The FSSA (10 items) was derived from the English version of the ESS modified by Snow et al., 2002 [26] for adolescents with hypersomnolence and was already validated in French population in 2005 college students [30]. Transcultural adaptation (translation and back-translation) has already done in this pilot study. We excluded two items of the French version in order to have the same number of items than the other versions (Appendix 1). The FSSA8 (8 items) questionnaire was self-administered during consultation or school for the clinical or nonclinical group respectively. In this questionnaire, the chances of falling asleep in 8 different situations were estimated by a 4-points scale ranging from 0 to 3 points.

2.5. Procedure and statistical analysis

2.5.1. Pretesting of the FSSA8

In pretesting, understanding of the questions was analyzed according to the Vallerand cross-cultural adaptation procedure [39]. Before the questionnaire was used, 20 French-speaking adolescents were asked to assess their understanding for each item rated using a seven-point Likert-type scale of the retained French version. We then calculated averages and the first quartile of understanding scores for each item. Items which obtained average or upper quartile marks of less than four were modified to render them clearer.

2.5.2. Population

Patient characteristics and FSSA8 total scores were compared between the nonclinical and clinical groups. Comparisons were made using Fisher's exact test in case of categorical data and Kruskal-Wallis test in case of continuous data. The scalability of the scale, its factorial structure, its reliability and items properties were studied on the whole dataset, whereas the scale validity and Receiver Operating Characteristic (ROC) analyses were performed after frequency matching on age and sex .

2.5.3. Scalability of the scale: item response theory (IRT) assumptions

The FSSA8 ability to measure a latent trait (i.e. EDS in our case) should be verified. For this, the FSSA8 must verified the three IRT assumptions: 1) FSSA8 had to measure an only one latent trait, 2) there must be a dose-response relationship between the latent trait and the scoring of each item, and 3) the scoring of each items must be independent of the scoring of the others. The non-parametric IRT analysis [40] (Mokken scale analysis [41]) used in our study assumptions is detailed in supplementary materials page 2.

The psychometric properties of each item were estimated in case of non-rejection of IRT assumptions using confirmatory factorial analysis (CFA) by fitting an appropriate structural equation model that take into account the potential non-normality of items scores (see supplementary materials, page 3) [42]. A root mean square error of approximation (RMSEA) < 0.06 and comparative fit index (CFI) value > 0.95 were considered representative of a well-fitting [43]. In this model, the daytime sleepiness is represented by a latent variable (non-directly observable) that is considered to be continuous in nature. The model estimates the latent sleepiness for each individual.

2.5.4. Items

Floor and ceiling effects were examined for the 8 items. According to the criteria of Petrillo et al. [43], items have floor or ceiling effects when more than 40% of responders select category 0 or category 3 on a 4 point scale (supplementary materials, page 4). Skewness and kurtosis were reported to evaluate the non-normality of the item distribution. In this paper "easy items" refer to high-scored items with left long-tailed distribution and "difficult items" to low-scored items with right long-tailed distribution. Consequently, items with ceiling (*vs.* floor) effects are easy (*vs.* difficult) items. The validity of FSSA8 according to classical test theory was assessed using inter-item polychoric correlations and item-total polychoric correlations in case of non-normality of item scores.

The parameters of the normal ogive two-parameter model give the discriminative power or loading factor (slope) of each item. The global difficulty of each item was computed as being the point on the latent sleepiness at which the highest and lowest categories have equal probability of being observed (see supplementary materials, pages 4–5).

The item characteristic curves (ICCs) were computed for each item. ICCs give the probability of obtaining the different response category as a function of the latent continuum sleepiness (see supplementary materials, page 5). These curves allow knowing the range of latent sleepiness for which a particular category is the most likely scored. The calibration curves summarize the previous information by representing each ICC by a band per item.

The item information curves gave the reliability of each item according to the latent sleepiness. As a rule of thumb item reliability is acceptable when item information is greater than 3.3 [44].

2.5.5. Scale reliability

Internal reliability of the FSSA8 was assessed by usual Cronbach's alpha [45] that estimates the percentage of the variability in the total sum of scores explained by the underlying latent sleepiness. This coefficient indicates an acceptable reliability when greater than 0.70 [46]. The total information of the scale that reflects the variation of its reliability according to the latent sleepiness was also reported. The reliability of the scale is considered acceptable for patients with a latent sleepiness for which the total information is greater than 3.3.

2.5.6. Scale validity (using matched subgroups on age and sex)

2.5.6.1. Concurrent criterion validity. The traditional definition of concurrent criterion validation is the correlation between the scale and some other measure of disorder under study measured at the same time. In the context of the present study, the criterion validity was assessed by the Spearman correlation between FSSA8 total score and the propensity to fall asleep objectively measured by the MSLT.

2.5.6.2. Divergent construct validit. This was assessed using the difference between the FSSA8 total scores in nonclinical and clinical (mixed and narcoleptic) matched groups.

2.5.6.3. Diagnostic validity. ROC analysis was performed to determine, using Youden criteria, the best cut-off that allowed identification of clinical or narcoleptic subjects among nonclinical subjects. Area under ROC curve (AUC), Youden threshold,

Table 1
Characteristics of participants for the whole dataset and for the matched groups.

		Two main groups						
		Nonclinical group ($n=269$)	Clinical group ($n=115$)					
			Mixed Subgroup ($n=30$)	Narcoleptic Subgroup ($n=85$)	p[b]	Narcoleptic subgroup ($n=85$)		
						type 1 ($n=71$)	type 2 ($n=14$)	p[b]
All participants	n (%)	269 (70)	30 (~~7.8~~)	85(22)		71(84)	14(16)	
	Female sex: n (%)	137 (51)	13(43)	45(53)	0.67	36(51)	9(64)	0.39
	Age (years)[a]	13.7 ± 1.1	15.2 ± 1.4	14.7 ± 1.5	$<10^{-3}$	14.8 ± 1.5	14.5 ± 1.2	0.48
	MSLT: median [IQR]		13.4 [9.3–17.2]	3.4[2.0–5.0]	$<10^{-3}$	3.3[2.0–4.0]	4.5[1.9–6.9]	0.19
	FSSA8 total score[a]	2.7 ± 2.9	12.0 ± 5.1	15.5 ± 4.4	$<10^{-3}$	15.9 ± 4.2	13.5 ± 4.8	0.086
	FSSA8 mean score[a]	0.34 ± 0.37	1.50 ± 0.63	1.94 ± 0.55		1.99 ± 0.53	1.69 ± 0.60	
Matched groups	n (%)	73 (50.3)	23 (15.6)	50(34)		40(80)	10(20)	
	Female sex: n(%)	35 (47.3)	9 (39.1)	25(50)	0.73	19(47.5)	6(60)	0.73
	Age (years)[a]	14.2 ± 1.2	14.7 ± 1.2	14.0 ± 1.1	0.13	14.0 ± 1.2	14.0 ± 0.8	0.95
	MSLT median [IRQ]		13.7 [10.2–17.2]	2.4[1.4–4.0]	$<10^{-3}$	2.3[1.4–3.9]	4.3[1.4–6.1]	0.21
	FSSA8 total score[a]	3.3 ± 3.3	11.7 ± 4.8	16.1 ± 4.3	$<10^{-3}$	16.7 ± 3.9	13.5 ± 5.1	0.056
	FSSA8 mean score[a]	0.41 ± 0.42	1.46 ± 0.59	2.01 ± 0.54		2.09 ± 0.48	1.69 ± 0.64	

Type 1 (Type 2): narcoleptic patient with (without) cataplexy. IQR: [first quartile–third quartile]. MSLT: Multiple Sleep Latency Tests, FSS8: French Sleepiness Scale 8 items
[a] mean ± Standard Deviation
[b] Exact fisher test for categorical variable; Kruskal-Wallis test (>2 groups) or Wilcoxon rank sum test (2 groups) for quantitative variable

Table 2
Item description for the 8-items French Sleepiness Scale for Adolescents according to main groups.

		n	Missing	Mean ± SD	Skewness	Kurtosis	Minimum	Maximum	Extreme	
Main groups	Item								Skewness	Kurtosis
All	1	383	1	0.75 ± 1.03	1.06	−0.26	0	3		
	2	383	1	0.95 ± 1.02	0.68	−0.78	0	3		
	3	383	1	0.60 ± 0.96	1.4	0.65	0	3		
	4	384	0	1.41 ± 1.16	0.18	−1.44	0	3		
	5	383	1	1.11 ± 1.24	0.55	−1.37	0	3		
	6	382	2	0.22 ± 0.56	2.72	7.11	0	3	>2	>2
	7	382	2	0.54 ± 0.97	1.54	0.90	0	3		
	8	382	2	0.68 ± 1.05	1.18	−0.13	0	3		
Non Clinical	1	269	0	0.32 ± 0.65	2.2	4.61	0	3	>2	>2
	2	269	0	0.58 ± 0.82	1.31	0.92	0	3		
	3	268	1	0.16 ± 0.41	2.48	5.66	0	2	>2	>2
	4	269	0	0.87 ± 0.89	0.77	−0.22	0	3		
	5	268	1	0.50 ± 0.82	1.64	1.88	0	3		
	6	267	2	0.04 ± 0.28	7.75	65.71	0	3	>2	>2
	7	267	2	0.06 ± 0.31	6.38	44.52	0	3	>2	>2
	8	267	2	0.16 ± 0.50	3.33	11.35	0	3	>2	>2
Clinical	1	114	1	1.76 ± 1.04	−0.36	−1.07	0	3		
	2	114	1	1.82 ± 0.92	−0.32	−0.81	0	3		
	3	115	0	1.62 ± 1.10	−0.2	−1.29	0	3		
	4	115	0	2.68 ± 0.63	−1.94	3.23	0	3		>2
	5	115	0	2.53 ± 0.82	−1.71	2.00	0	3		
	6	115	0	0.63 ± 0.79	0.97	−0.06	0	3		
	7	115	0	1.67 ± 1.03	−0.22	−1.12	0	3		
	8	115	0	1.88 ± 1.00	−0.54	−0.79	0	3		

sensitivity, and specificity and their 95% confidence intervals were estimated using bootstrap.

Statistical analyses were performed using R language version 3.5.2 available at http://cran.r-project.org/and Mplus 7.11 for factorial analyses and item properties. The R packages used were mokken for the scalability of the scale [47], psy and boot for Cronbach coefficient, mplusAutomation [48], and pROC.

3. Results

3.1. Pretesting of the FSSA8

The participants for the pretesting were 11 boys and 9 girls with a mean age of 15.9 years (SD = 1.99). None of the items received a score lower than six out of seven. All items were rated in the first quartile with a score of seven out of seven. Therefore, all items of the FSSA8 are concluded to be well understood.

3.2. Population

A total of 384 adolescents were included; 269 in the nonclinical group and 115 in the clinical group: 85 narcoleptic patients (71 with cataplexy (type 1) and 14 without (type 2)) and 30 patients with ~~had~~ other sleep disorders (10 had insomnia, 9 had phase delay syndrome, 7 had bad sleep hygiene, 5 with upper airway resistance syndrome and 5 with OSA (AHI = 7.5 (3–14). In the nonclinical group, 4 participants forgot to fill one item and 2 participants 2 items (0.26% missing items responses), in the clinical group only 2 patients forgot to fill one item. We replaced these missing responses by the item's mean of the participants for Epworth score computation. No imputation was performed for the factorial analyses.

There was a significant difference in the mean age between groups. We performed a matched screening on age and sex and obtained 73 adolescents in the control group and a clinical matched group that included 50 narcoleptic patients and 23 with other

Table 3
Psychometric properties of French Sleepiness Scale for Adolescents (FSSA8) in item level.

Item	Item-total Corrélation	Loading factor (discrimination)	Item difficulty
1	0.83	1.74	0.87
2	0.71	1.21	0.72
3	0.84	1.91	1.04
4	0.77	1.58	0.03
5	0.85	2.32	0.36
6	0.79	1.53	2.05
7	0.89	2.73	1.07
8	0.86	2.18	0.93

sleep disorders. There was no significant difference in terms of age and sex between matched groups (Table 1). Adolescents with narcolepsy had lower sleep latency at MSLT than those with other sleep disorders (with median 2.4 vs. 13.7, $p < 0.001$), and there was a trend towards a lower sleep latency at MSLT in those with type 1 narcolepsy than in those with type 2 narcolepsy (with median 2.3 vs. 4.3, $p = 0.21$).

3.3. Preliminary analysis of items

The discrepancy with normality was more frequent in the nonclinical group (5 items with skewness > 2) whereas items in the clinical group showed negative skewness, except item 6 (Table 2). All items had a floor effect in the nonclinical group. In the clinical group, only item 6 had a floor effect and items 4 and 5 had a ceiling effect (supplementary materials data: Figure S1). These results suggested that items 4 and 5 were too difficult for nonclinical participants but too easy for clinical group whereas item 6 was difficult for both groups (nonclinical and clinical).

Item-total polychoric correlations between items score and the total FSSA8 score for all 384 participants showed strong positive correlation ranging from 0.71 to 0.89 (Table 3). The average of polychoric inter-item correlation was 0.76 (range 0.61–0.87).

3.3.1. Scalability of the scale: item response theory (IRT) assumptions

The IRT assumptions were verified in the whole dataset (supplementary materials: Tables S1-S2). The overall scalability of the scale was 0.71 (95%CI: 0.67–0.75) denoting a strong scale (> 0.5) according to Mokken [41]. The normal ogive two-parameter model (CFA) presented good fitting indexes (RSMEA = 0.049 and CFI = 0.998). FSSA8 was unidimensional measuring only one latent trait.

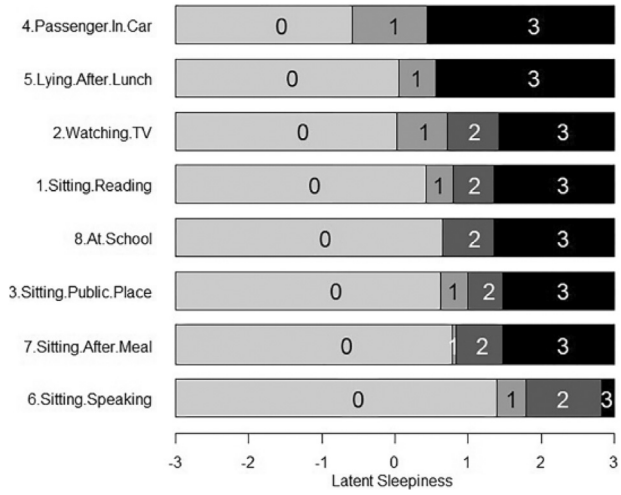

Fig. 2. Calibration curve is represented by a horizontal bar for each item of the 8-items French Sleepiness Scale for Adolescents (n = 384). Items are sorted by increasing difficulty from top to bottom. The different grey levels indicate the region of latent sleepiness in which a certain category is most likely: light grey for category 0 (non-existent risk), medium grey for category 1 (negligible risk), dark grey for category 2 (moderate risk) and black for 3 (important risk). Example: the probability of scoring 3 (instead of 0, 1 or 2) on item 6 is the highest for patients with latent sleepiness more than 2.82.

3.3.2. Psychometric properties of items according to IRT

Item loading factors (slope) ranged from 1.21 to 2.73 and item difficulties from 0.03 to 2.05 on the latent sleepiness scale. The Item-Person map located the items by their difficulty and the participants by their latent sleepiness on the same latent continuum sleepiness (Fig. 1). Participants were most likely to fall asleep when sitting "in a car for one-hour drive" (item 4) and when "lying down after lunch (when circumstances permit)" (item 5). They were least likely to fall asleep when "sitting and talking to someone (item 6)". The 5 items (1,2,3,7,8) had close difficulties (range: 0.72 to 1.07).

The calibration curve (Fig. 2) showed that for all items, the model predicted that subjects who had a latent sleepiness less than −0.58

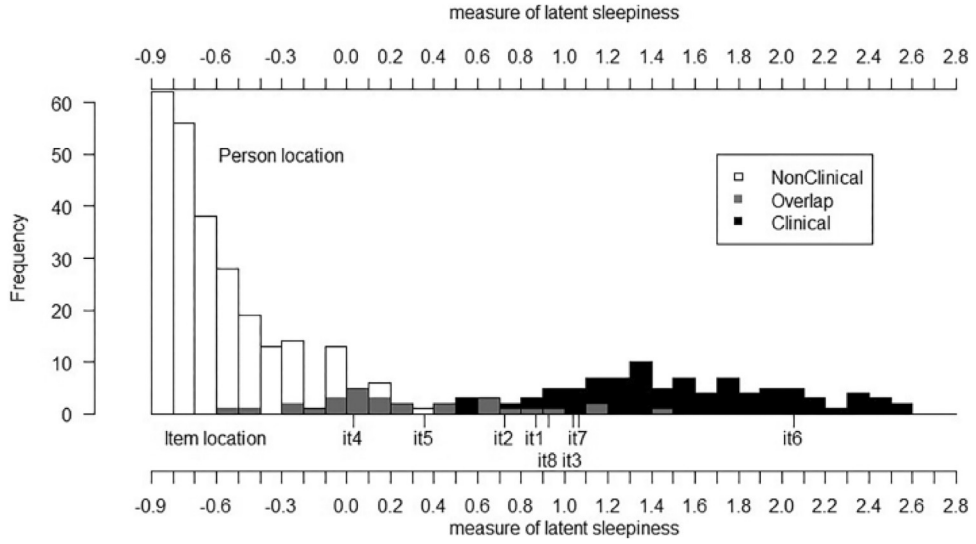

Fig. 1. Item-Person map for the 8-items French Sleepiness Scale for Adolescents (n = 384). The nonclinical (n = 269) and clinical (n = 115) groups are represented in white and black respectively with overlapping of both groups in dark grey. Item location correspond to the item difficulty. The easiest item was item 4 and the most difficult item 6.

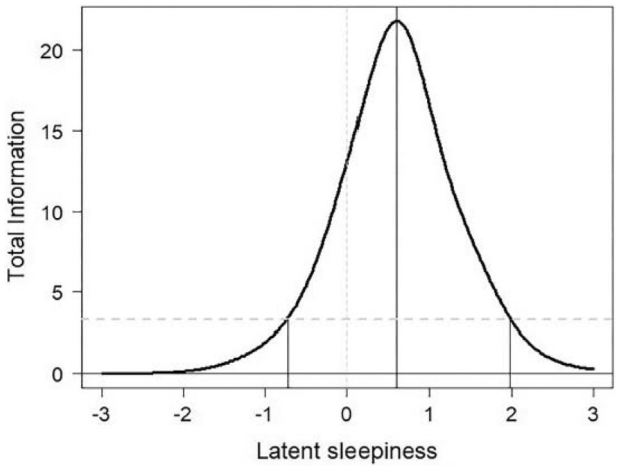

Fig. 3. Total Information curve of the 8-items French Sleepiness Scale for Adolescents ($n = 384$). The horizontal dashed grey line indicates the information threshold above whom (>3.3) the scale is sufficient reliable to place subject on the latent scale with precision. The vertical black lines indicate the latent sleepiness for which the information is maximum or above 3.3. The scale is the most precise for measuring individuals with latent sleepiness around 0.6 and had reliable estimates for subject between -0.7 and 2.0 (the two extreme vertical black lines).

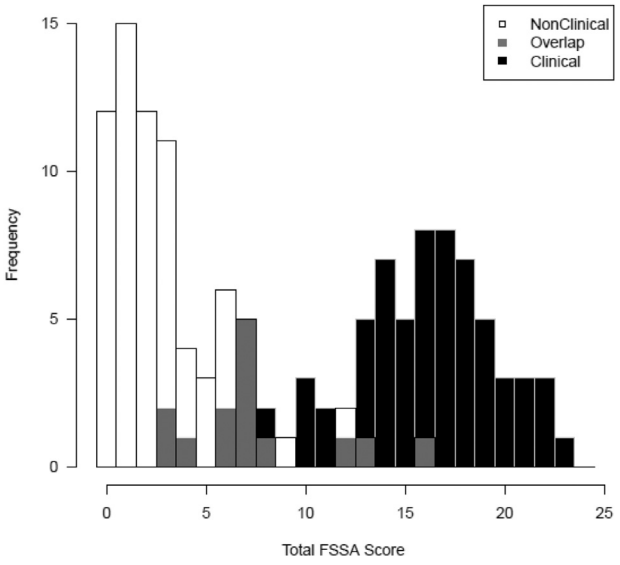

Fig. 4. Distribution of the total score of the 8-items French Sleepiness Scale for Adolescents (FSSA8) in the two matched nonclinical ($n = 73$) and clinical ($n = 73$) groups fully described in Table 1. The dark grey area within the histogram reflects overlapping of both groups.

or greater than 2.82, scored most likely the response category, 0 or 3 respectively (for details see supplementary materials: Figure S2 and Table S3). Only 163 subjects had a latent sleepiness less than −0.58 and all were in the nonclinical group representing 61% of this group. This suggests a weak discriminative power of all items for nonclinical subjects. Items 4, 5, and 8 functioned like trichotomous items because only 3 categories are most likely selected. The 5 items (1,2,3,7,8) changed their most likely scored category from 2 to 3 at close latent sleepiness between 1.35 and 1.47.

Response category 0 tended to be scored only by nonclinical participants for items 4, 5, and 2 (supplementary materials Figure S3). Response categories above 0 were never scored by the nonclinical participants for item 6. Response category 3 tended to be scored by patients for items 2, 1, 8, 3 and 7.

The item information according to latent sleepiness was acceptable (> 3.3) only for the 3 items (items 7, 5, 8) having as expected the highest loading factor and for participants with latent sleepiness between −0.14 and 1.10 i.e. for 32% of patients and 13% of nonclinical participants (supplementary materials: Figure S4).

3.4. Scale reliability

The FSSA8 had a Cronbach of 0.92 (95% CI: 0.91–0.93, $n = 384$); this was greater than 0.7, confirming the good reliability of this scale. The scale showed an acceptable reliability (information > 3.3) for latent continuum sleepiness between −0.7 and 2 (Fig. 3). This range of sleepiness applied to 56% (151/269) of nonclinical participants and 84% (97/115) of patients. The scale appeared to be most reliable for the clinical group.

3.5. Scale validity

3.5.1. Concurrent criterion validity

Correlation between FSSA8 score and MSLT was −0.42 ($p = 0.0003$, $n = 73$) for matched patients (mixed and narcoleptic).

3.5.2. Divergent construct validity

FSSA8 total score was significantly different between nonclinical and clinical matched groups (Mann and Whitney test; median 2 for nonclinical and 16 for clinicals) with non-normal distribution within groups (Fig. 4). The FSSA8 total score distribution differed between groups. It was highly skewed with a right tail for nonclinical meaning that the scale was quite difficult for nonclinicals (with low or weak scores) and it was bimodal for clinical groups (Fig. 4).

Table 4
Results of the ROC analysis with matched groups.

	Nonclinical vs Clinical[a]	Nonclinical vs Narcoleptic[a]
AUC	0.96(0.93–0.98)	0.98(0.96–0.99)
Threshold (>)	7(5.5–10)	11 (6.50–13.5)
Sensitivity	0.90(0.79–0.99)	0.92(0.82–1)
Specificity	0.90(0.80–0.99)	0.96(0.85–1)

AUC: Area Under the Curve; Threshold: Youden criteria.
[a] Estimate and (95% confident interval obtained by bootstrap).

3.5.3. Diagnostic validity

The ROC analysis of data was performed on the matched groups (Table 4). Subjects with FSSA8 score above (>) 7 may present underlying EDS with 90% sensitivity and 90% specificity. Subjects with FSSA8 score above (>) 11 may present underlying narcolepsy with 92% sensitivity and 96% specificity.

4. Discussion

4.1. Statements of principal findings

The one-dimensionality of the FSSA8 was confirmed in a heterogeneous population (general and clinical samples). The scale appeared to be more reliable for patients than nonclinical adolescents, and to be a good tool for screening EDS in the French general population with a cut-off of 7 of the total score (>11 for suspicion of narcolepsy).

4.2. Strengths and weaknesses of the study

The present study is the first to have included patients with narcolepsy in order to evaluate the psychometric property of a sleepiness scale for adolescents. Given that narcolepsy is the sleep disorder with the highest level of sleepiness, to study this population is essential to estimate scale validity. The prevalence of narcolepsy in general population is low, our sample was very large

and allowed to compare with age and gender equivalence sleepiness of general and clinical population.

The study has several limitations. For instance, as sleep deprivation is the most frequent cause of EDS in adolescents, it should be necessary to estimate the test-retest fidelity between school term which increased sleep deprivation risk and holiday period. We did not assess measurement invariance according to clinical status or gender at this stage of development. In addition, the retrospective nature of the study did not allow us collecting the weight and the height for all adolescents. It would have been interesting to adjust on BMI when comparing global score between groups. This study focused only on adolescents and the validation for the less than 12 is not performed yet. Moreover, in our reference centre for narcolepsy, we had very few children with idiopathic hypersomnia ($n = 2$) and we did not include them in our study.

4.3. Strengths and weaknesses in relation to other studies

As expected, the FSSA8 scale appeared to be more reliable in patients than in nonclinical individuals. The FSSA8 items were the most precise in a small band of positive latent sleepiness corresponding mainly to the patients. This scale seemed to be less reliable in nonclinical than in patients.

Some FSSA8 items had extreme difficulty. Item 6 was the most difficult and item 5 one the two easiest items. These results were consistent with other versions of ESS for adolescents [27,31]. Despite an acceptable internal consistency, the analysis of the total information suggests that global scale was more reliable in range of latent sleepiness corresponding mainly to clinical participants patients in line with what was observed per item.

As others studies [9,49,50], the concurrent criterion validity related to MSLT was low ($r = -.4$). It might be due to the fact that FSSA8 evaluates the subjective propensity to fall asleep according to various postures and activities whereas MSLT objectively measures sleepiness in the same sleep lab conditions. ~~as~~ Several authors criticized the scales measuring sleepiness [51,52]. Nevertheless, correlation with MSLT was not relevant for all items. Situations described in the FSSA8 diverge on the subject's need to stay awake; for instance, the motivation to stay awake in class is not comparable to that elongated after the meal. The motivation to stay awake is more comparative to maintenance of wakefulness test rather than MSLT.

The distribution of global scores for the nonclinical group and the clinical group was significantly different, even considering that certain adolescents in the nonclinical group might have experienced sleep disturbances.

We retained the cut-off of 7 to screen EDS in the general population. FSSA8 correctly detected 90% pathological sleepiness in the clinical group. Nonclinical and clinical groups distribution did not much overlap and this confirmed the ROC analysis results: the FSSA8 is a robust screening tool for EDS in adolescents. The Pediatric Daytime Sleepiness Scale (PDSS) differs from FSSA8 and the cut offs proposed by Meyer et al. [53] in general population may not be comparable with ours. In 2011, Lecendreux et al. decided to use a scale derived to ESS and not PDSS [30]. Indeed, during this period in NARCOBANK study, we showed that PDSS was not sensible for EDS in narcoleptic children [37]. In the other hand, the only study that reported a cut off for clinical diagnostic was conducted among children between 3 and 12 years old, suspected of sleep disordered breathings and using a modified ESS for children in 2009 [18]. The cut off reported in this study may not be compared as well.

4.3.1. Meaning of the study: implications for clinicians

The analysis of the group distribution for each response category in the calibration curve can guide the clinician in detecting EDS linked to potential sleep pathologies. Adolescents who scored 0 for items 4, 5 and 2 had a weak probability to be clinical subject. Subjects who scored 3 for items 2, 1, 8, 3 and 7 are likely narcoleptic and the ones who scored at least 1 for item 6 are likely clinical subjects.

4.4. Unanswered questions and future work

Our data were collected retrospectively, in the meantime other cultural version of the ESS adapted to adolescents has been published [27,31,54]. It differs just a little from FSSA8: item 4 also mentions afternoon and not only morning; and item 8 mentions a car trip of one hour and not of 30 minutes. Meanwhile, the ESS-CHAD scale developed by Johns has been validated in clinical group in adolescents [32]. A comparison between ESS-CHAD and FSSA should be made. It might be relevant to have teenagers who complete both versions at the same time in a counterbalanced order.

Financial Disclosure

The authors have no financial relationships relevant to this article to disclose

Disclosure of interests

The authors declare that they have no competing interests.

Acknowledgements

This study was run with the support of the child and adolescent sleep working group (GTSEA) of the French Society for Sleep Research and Medicine (SFRMS). We thank all health care professionals who participated in the data collection.

Online Supplement. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.encep.2022.06.004.

References

[1] Gibson ES, Powles ACP, Thabane L, et al. Sleepiness" is serious in adolescence: two surveys of 3235 Canadian students. BMC Public Health 2006;6:116, http://dx.doi.org/10.1186/1471-2458-6-116.
[2] Merdad RA, Akil H, Wali SO. Sleepiness in Adolescents. Sleep Med Clin 2017;12:415–28, http://dx.doi.org/10.1016/j.jsmc.2017.03.014.
[3] Owens JA, Weiss MR. Insufficient sleep in adolescents: causes and consequences. Minerva Pediatrica 2017;69:326–36, http://dx.doi.org/10.23736/S0026-4946.17.04914-3.
[4] Verhulst SL, Van Gaal L, De Backer W, et al. The prevalence, anatomical correlates and treatment of sleep-disordered breathing in obese children and adolescents. Sleep Med Rev 2008;12:339–46, http://dx.doi.org/10.1016/j.smrv.2007.11.002.
[5] Sivertsen B, Pallesen S, Stormark KM, et al. Delayed sleep phase syndrome in adolescents: prevalence and correlates in a large population based study. BMC Public Health 2013;13:1163, http://dx.doi.org/10.1186/1471-2458-13-1163.
[6] Kothare SV, Kaleyias J. Narcolepsy and other hypersomnias in children. Curr Opin Pediatr 2008;20:666–75, http://dx.doi.org/10.1097/MOP.0b013e328316bd85.
[7] Sinha A, Kling S. A review of adolescent obesity: prevalence, etiology, and treatment. Obes Surg 2009;19:113–20, http://dx.doi.org/10.1007/s11695-008-9650-4.
[8] Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991;14:540–5, http://dx.doi.org/10.1093/sleep/14.6.540.
[9] Chervin RD. Epworth sleepiness scale? Sleep Med 2003;4:175–6, http://dx.doi.org/10.1016/s1389-9457(03)00030-3.
[10] Hagell P, Broman J-E. Measurement properties and hierarchical item structure of the Epworth Sleepiness Scale in Parkinson's disease. J Sleep Res 2007;16:102–9, http://dx.doi.org/10.1111/j.1365-2869.2007.00570.x.
[11] Ulander M, Årestedt K, Svanborg E, et al. The fairness of the Epworth Sleepiness Scale: two approaches to differential item functioning. Sleep and Breathing 2013;17:157–65, http://dx.doi.org/10.1007/s11325-012-0664-8.

[12] Smith SS, Oei TPS, Douglas JA, et al. Confirmatory factor analysis of the Epworth Sleepiness Scale (ESS) in patients with obstructive sleep apnoea. Sleep Med 2008;9:739–44, http://dx.doi.org/10.1016/j.sleep.2007.08.004.
[13] Miletin MS, Hanly PJ. Measurement properties of the Epworth sleepiness scale. Sleep Med 2003;4:195–9, http://dx.doi.org/10.1016/s1389-9457(03)00031-5.
[14] Violani C, Lucidi F, Robusto E, et al. The assessment of daytime sleep propensity: a comparison between the Epworth Sleepiness Scale and a newly developed Resistance to Sleepiness Scale. Clin Neurophysiol 2003;114:1027–33, http://dx.doi.org/10.1016/s1388-2457(03)00061-0.
[15] Takegami M, Suzukamo Y, Wakita T, et al. Development of a Japanese version of the Epworth Sleepiness Scale (JESS) based on Item Response Theory. Sleep Med 2009;10:556–65, http://dx.doi.org/10.1016/j.sleep.2008.04.015.
[16] Kaminska M, Jobin V, Mayer P, et al. The Epworth sleepiness scale: self-administration versus administration by the physician, and validation of a french version. Can Respirat J 2010;17:e27–34, http://dx.doi.org/10.1155/2010/438676.
[17] Benmedjahed K, Wang YG, Lambert J, et al. Assessing sleepiness and cataplexy in children and adolescents with narcolepsy: a review of current patient-reported measures. Sleep Med 2017;32:143–9, http://dx.doi.org/10.1016/J.SLEEP.2016.12.020.
[18] Chan EYT, Ng DK, Chan C, et al. Modified Epworth Sleepiness Scale in Chinese children with obstructive sleep apnea: a retrospective study. Sleep and Breathing 2009;13:59–63, http://dx.doi.org/10.1007/s11325-008-0205-7.
[19] Storfer-Isser A, LeBourgeois M, Harsh J, et al. Psychometric properties of the adolescent sleep hygiene scale (ASHS). J Sleep Res 2013:22, http://dx.doi.org/10.1111/JSR.12059.
[20] Anderson B, Storfer-Isser A, Taylor HG, et al. Associations of executive function with sleepiness and sleep duration in adolescents. Pediatr 2009;123:e701–7, http://dx.doi.org/10.1542/peds.2008-1182.
[21] Moore M, Kirchner HL, Drotar D, et al. Relationships among sleepiness, sleep time, and psychological functioning in adolescents. J Pediatr Psychol 2009;34:1175–83, http://dx.doi.org/10.1093/jpepsy/jsp039.
[22] Lee YJ, Cho S-J, Cho IH, et al. Insufficient sleep and suicidality in adolescents. Sleep 2012;35:455–60, http://dx.doi.org/10.5665/sleep.1722.
[23] Evans DS, Snitker S, Wu S-H, et al. Habitual sleep/wake patterns in the Old Order Amish: heritability and association with non-genetic factors. Sleep 2011;34:661–9, http://dx.doi.org/10.1093/sleep/34.5.661.
[24] Campbell IG, Higgins LM, Trinidad JM, et al. The increase in longitudinally measured sleepiness across adolescence is related to the maturational decline in low-frequency EEG power. Sleep 2007;30:1677–87, http://dx.doi.org/10.1093/sleep/30.12.1677.
[25] Melendres CS, Lutz JM, Rubin ED, et al. Daytime sleepiness and hyperactivity in children with suspected sleep-disordered breathing. Pediatrics 2004;114:768–75, http://dx.doi.org/10.1542/peds.2004-0730.
[26] Snow A, Gozal E, Malhotra A, et al. Severe hypersomnolence after pituitary/hypothalamic surgery in adolescents: clinical characteristics and potential mechanisms. Pediatrics 2002;110:e74, http://dx.doi.org/10.1542/peds.110.6.e74.
[27] Janssen KC, Phillipson S, O'Connor J, et al. Validation of the Epworth sleepiness scale for children and adolescents using Rasch analysis. Sleep Med 2017;33:30–5, http://dx.doi.org/10.1016/j.sleep.2017.01.014.
[28] Bootzin RR, Stevens SJ. Adolescents, substance abuse, and the treatment of insomnia and daytime sleepiness. Clin Psychol Rev 2005;25:629–44, http://dx.doi.org/10.1016/j.cpr.2005.04.007.
[29] Shin C, Kim J, Lee S, et al. Sleep habits, excessive daytime sleepiness and school performance in high school students. Psychiatr Clin Neurosci 2003;57:451–3, http://dx.doi.org/10.1046/j.1440-1819.2003.01146.x.
[30] Lecendreux M, Gamble H, Sanchez-Garrido L, et al. An evaluation study of the Epworth sleepiness scale adapted for children and adolescents: a tool for case finding of pediatric excessive daytime sleepiness (EDS). Eur Psychiatry 2011;26:1559, http://dx.doi.org/10.1016/s0924-9338(11)73263-1.
[31] Imani V, Lin C-Y, Jalilolgadr S, et al. Factor structure and psychometric properties of a Persian translation of the Epworth Sleepiness Scale for Children and Adolescents. Health Promotion Persp 2018;8:200, http://dx.doi.org/10.15171/HPP.2018.27.
[32] Wang YG, Menno D, Chen A, et al. Validation of the Epworth Sleepiness Scale for Children and Adolescents (ESS-CHAD) questionnaire in pediatric patients with narcolepsy with cataplexy aged 7-16 years. Sleep Med 2022;89:78–84, http://dx.doi.org/10.1016/J.SLEEP.2021.11.003.
[33] Zhang M, Inocente CO, Villanueva C, et al. Narcolepsy with cataplexy: does age at diagnosis change the clinical picture? CNS Neurosci Therapeut 2020;26:1092–102, http://dx.doi.org/10.1111/cns.13438.
[34] Barateau L, Lopez R, Arnulf I, et al. Comorbidity between central disorders of hypersomnolence and immune-based disorders. Neurology 2017;88:93–100, http://dx.doi.org/10.1212/WNL.0000000000003432.
[35] Lecendreux M, Lavault S, Lopez R, et al. Attention-deficit/hyperactivity disorder (ADHD) symptoms in pediatric narcolepsy: A cross-sectional study. Sleep 2015;38:1285–95, http://dx.doi.org/10.5665/sleep.4910.
[36] Spruyt K, Herbillon V, Putois B, et al. Mind-wandering, or the allocation of attentional resources, is sleep-driven across childhood. Sci Rep 2019:9, http://dx.doi.org/10.1038/s41598-018-37434-5.
[37] Inocente CO, Lavault S, Lecendreux M, et al. Impact of obesity in children with narcolepsy. CNS Neuroscience & Therapeutics 2013;19:521–8, http://dx.doi.org/10.1111/cns.12105.
[38] American academy of sleep medicine. The International Classification of Sleep Disorders 3rd ed. Darien, IL: American academy of sleep medicine; 2014.
[39] Vallerand RJ. Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: Implications pour la recherche en langue française. Can Psychol 1989;30:662–80, http://dx.doi.org/10.1037/h0079856.
[40] Sijtsma K, Molenaar I. Introduction to Nonparametric Item Response Theory. 2455 Teller Road. In: Thousand Oaks California 91320 United States of America. SAGE Publications, Inc; 2002, http://dx.doi.org/10.4135/9781412984676.
[41] Mokken RJ. A theory and procedure of scale analysis. Mouton/De: The Hague/Berlin; 1971.
[42] Muthén B, du Toit SHC, Spisic D. Robust inference using weighted least squares and quadratic estimating equation in latent variable modeling with categorical and continuous outcomes. Unpublished Manuscript 1997.
[43] Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling. Multidiscipl J 1999;6:1–55, http://dx.doi.org/10.1080/10705519909540118.
[44] Petrillo J, Cano SJ, McLeod LD, et al. Using classical test theory, item response theory, and rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. Value in Health 2015;18:25–34, http://dx.doi.org/10.1016/j.jval.2014.10.005.
[45] Cronbach IJ. Coefficient alpha and the internal structure of test. Psychometrika 1951:297–334, http://dx.doi.org/10.1007/BF02310555.
[46] Bland JM, Altman DG. Cronbach's alpha. BMJ (Clinical Research Ed) 1997;314:572, http://dx.doi.org/10.1136/bmj.314.7080.572.
[47] Ark LA, van der. New Developments in Mokken Scale Analysis in R. J Statisti-Software 2012;48:1–27, http://dx.doi.org/10.18637/jss.v048.i05.
[48] Hallquist M, Wiley JF, MplusAutomation: An R. Package for facilitating large-scale latent variable analyses in Mplus. Structural Equation Modeling. Multidiscipl J 2018;25:621–38, http://dx.doi.org/10.1080/10705511.2017.1402334.
[49] Sangal RB, Mitler MM, Sangal JM. Subjective sleepiness ratings (Epworth sleepiness scale) do not reflect the same parameter of sleepiness as objective sleepiness (maintenance of wakefulness test) in patients with narcolepsy. Clin Neurophysiol 1999;110:2131–5, http://dx.doi.org/10.1016/s1388-2457(99)00167-4.
[50] Chervin RD, Aldrich MS. The Epworth Sleepiness Scale may not reflect objective measures of sleepiness or sleep apnea. Neurology 1999;52:125–31, http://dx.doi.org/10.1212/wnl.52.1.125.
[51] Bliwise DL. Is the measurement of sleepiness the Holy Grail of sleep medicine? Am J Respirat Crit Care Med 2001;163:1517–9, http://dx.doi.org/10.1164/ajrccm.163.7.2103055e.
[52] Tachibana N, Taniguchi M. Why do we continue to use Epworth sleepiness scale? Sleep Med 2007;8:541–2, http://dx.doi.org/10.1016/j.sleep.2006.08.012.
[53] Meyer C, Barbosa D, Junior G, et al. Proposal of cutoff points for pediatric daytime sleepiness scale to identify excessive daytime sleepiness. Chronobiol Int 2018;35:1–9, http://dx.doi.org/10.1080/07420528.2017.1400980.
[54] Grace Wang Y, Benmedjahed K, Lambert J, et al. Assessing narcolepsy with cataplexy in children and adolescents: development of a cataplexy diary and the ESS-CHAD. Nature and Science of Sleep 2017;9:201–11, http://dx.doi.org/10.2147/NSS.S140143.