Downloaded 7/13/2017 from http://epworthsleepinessscale.com/about-the-ess/

About the ESS

Dr Johns first developed the ESS for adults in 1990 and subsequently modified it slightly in 1997. He developed it so he could assess the ‘daytime sleepiness’ of the patients in his own private practice of Sleep Medicine. He named the questionnaire after Epworth Hospital in Melbourne, where he established the Epworth Sleep Centre in 1988.

The ESS is a self-administered questionnaire with 8 questions. Respondents are asked to rate, on a 4-point scale (0-3), their usual chances of dozing off or falling asleep while engaged in eight different activities. Most people engage in those activities at least occasionally, although not necessarily every day. The ESS score (the sum of 8 item scores, 0-3) can range from 0 to 24. The higher the ESS score, the higher that person’s average sleep propensity in daily life (ASP), or their ‘daytime sleepiness’. The questionnaire takes no more than 2 or 3 minutes to answer. It is available in many different languages.

The 1997 version of the ESS is the standard version that can be used by most adults. A license is needed to use it, whether or not license fees are payable.

See sample copy below

The following topics are canvassed in the sections below

- What the ESS measures
- Questions included in the ESS
- The 1997 version of the ESS
- Recall period for the ESS
- Format of the ESS questionnaire
- How to score the ESS
- Reference range of normal ESS scores
- Psychometrics of the ESS
- Language of the ESS and its translation
- Limitations of the ESS

What the ESS Measures

The ESS asks the respondent to rate on a 4-point scale (0-3) their usual chances of having dozed off or fallen asleep while engaged in eight different activities that differ widely in their somnificity. These ESS item-scores provide estimates of eight different SSPs for that person (Johns, 1994; 2010). The total ESS score (the sum of 8 item-scores) gives an estimate of a more general characteristic, the person's 'average sleep propensity' or ASP, across a wide range of activities in their daily lives (Johns, 2002). There is no other measure of ASP available at present with which to compare ESS scores directly.

The ESS does not ask about the person's subjective feelings of alertness/drowsiness at some particular time, as measured by the Karolinska Sleepiness Scale. Nor does it measure how often, or for how long, the respondent sleeps during the day. The ESS is not a check-list for identifying those situations in which the respondent most frequently dozes during the day. Nor can it measure a person's level of alertness/drowsiness continuously, as Optalert technology does (Johns, 2008; Anderson et al, 2013).

The ESS specifically distinguishes reports of dozing behaviour (and estimates of SSPs) from feelings of fatigue and drowsiness/sleepiness, in the sense of 'weariness from exertion'. Fatigue and drowsiness/sleepiness are related concepts that are often confused (Johns 2000(a), 2003, 2009; Mairesse et al, 2016).

Questions included in the ESS

The particular questions included in the ESS were chosen on a priori grounds to represent activities with a wide range of different somnificities. Their relative somnificities were later confirmed by analysis of variance (Johns, 2010) and also by Rasch analysis (Hagell et al 2007; Izci et al, 2008; Sargento, et al, 2015). Item 5 ('lying down to rest in the afternoon when circumstances permit') is an activity with a much higher somnificity than Item-6 ('sitting and talking to someone').

The relative somnificities of ESS activities are similar in different diagnostic groups and populations, regardless of their levels of ASP and the presence or absence of sleep disorders (Johns, 2002, 2010).
The ESS items were not selected from a list of related questions by principal components analysis, as is commonly done in the development of other questionnaires.

The 1997 version of the ESS

With the initial (1990) version of the ESS some respondents did not answer all the questions, for whatever reasons. Even if one question was not answered, their ESS score was not valid because it was not possible to interpolate item-scores. Up to 5% of ESS scores were invalid in some groups who used the 1990 version.

In 1997 the instructions to respondents were changed, with the addition of an extra sentence, ‘It is important that you answer each question as best you can’. With this exhortation, nearly everyone answered all questions. The frequency of invalid ESS scores because of missed item-scores was reduced to less than 1%.

The 1997 version of the ESS is the standard version that can be used by most adults. It is available in many different languages as authorized translations.

Epworth Sleepiness Scale

Name: _____ Today's date: _____

Your age (Yrs): _____ Your sex (Male = M, Female = F): _____

How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired?

This refers to your usual way of life in recent times.

Even if you haven't done some of these things recently try to work out how they would have affected you.

Use the following scale to choose the **most appropriate number** for each situation:

0 = would **never** doze
1 = **slight chance** of dozing
2 = **moderate chance** of dozing
3 = **high chance** of dozing

It is important that you answer each question as best you can.

Situation	**Chance of Dozing (0-3)**
Sitting and reading	
Watching TV	
Sitting, inactive in a public place (e.g. a theatre or a meeting)	
As a passenger in a car for an hour without a break	
Lying down to rest in the afternoon when circumstances permit	
Sitting and talking to someone	
Sitting quietly after a lunch without alcohol	
In a car, while stopped for a few minutes in the traffic	

THANK YOU FOR YOUR COOPERATION

© M.W. Johns 1990-97

SAMPLE COPY

Recall period for the ESS

Respondents to the ESS rate their chances of having dozed off or fallen asleep in particular situations ‘in recent times’. It was a deliberate decision not to specify this recall period more accurately. It was intended to be long enough for the respondent to have experienced at least most of the activities, so they could estimate in retrospect their chances of dozing in each. Thus, ‘in recent times’ was intended to mean a few weeks to a few months, not a few hours or days.

There may be circumstances in which this recall period needs to be specified more accurately. For example, clinicians may want to compare ESS scores before and after instigating some particular treatment for a sleep disorder, such as nasal continuous airway pressure treatment for obstructive sleep apnea. Under those circumstances it is possible, with permission, to use a specific version of the ESS which specifies, for example, an interval such as ‘over the last month’ (Broderick JE et al, 2013).

Format of answers to ESS questions

In the usual format, each ESS item-score is recorded as one number (0-3), written in a box. Alternatively 4 boxes, labelled 0 to 3, may be presented for each question, and the respondent then ticks the most appropriate box. Electronic versions are possible too, by arrangement. The ESS scores derived from interviews, whether by phone or personally, may be valid, but that needs confirmation.

How to score the ESS

All ESS item-scores are intended to be integers (0-3). However, some people cannot decide on one number and report half-values. It is recommended that these scores be taken at face value. If, after adding them up, the total ESS score includes a half, it should be rounded up to the next whole number. If one or more item-scores are missing, that ESS is invalid because it is not feasible to interpolate missing item-scores. The ESS score (the sum of 8 item-scores) is the only number required under most circumstances. Respondents should not be given an ‘interpretation’ of particular ESS scores when they answer the questionnaire because that may influence their responses.

Reference range of normal ESS scores

It was initially reported in 1991 that the ‘normal’ range of ESS scores was 2 to 10 (Johns, 1991). With more data, that proved to be incorrect. Adults in Australia who have no evidence of a chronic sleep disorder (including frequent snoring) had a mean ESS score of 4.6 (95% confidence intervals of 3.9-5.3) with a standard deviation of 2.8 and a range from zero to 10 (Johns and Hocking, 2004). By this criterion, the reference range of ‘normal’ ESS scores is zero to 10. That is the same as the range defined by the 2.5 and 97.5 percentiles.

A similar ‘normal’ range has been reported from the United Kingdom (mean = 4.5 +/- 3.3 (SD), n= 188) (Chen at al, 1995), Italy (4.4+/- 2.8, n=54) (Manni et al, 1999), and Turkey (3.6 +/- 3.0, n= 60) (Izci et al, 2008). However, more evidence is needed to be sure that a similar reference range applies to other populations.

Many ESS scores reported from the general community, disregarding the presence or absence of sleep disorders, are higher than this ‘normal’ range. That is presumably because sleep disorders that increase ASPs, especially the different forms of sleep disordered breathing, are common in the community (Mihăicută et al, 2013). However, it should not be assumed that sleep disordered breathing is the only factor affecting ESS scores. Other causes, including depression, are also important. Gender and age have little effect on ESS scores among adults, but ethnicity does. African-Americans have significantly higher ESS scores than Caucasian Americans (Mihăicută et al, 2013; Sanford et al, 2006).

ESS scores of 11-24 represent increasing levels of ‘excessive daytime sleepiness’ (EDS). The percentage of people with EDS varies widely between different groups, from about 10 to 40% or more (Johns and Hocking, 1997; Sanford et al, 2006). Almost all patients suffering from narcolepsy have severe or moderate EDS by these ESS criteria, as expected (Parkes et al, 1998; Johns, 2000; van der Heide et al, 2015).

In general ESS scores can be interpreted as follows:

0-5 Lower Normal Daytime Sleepiness

6-10 Higher Normal Daytime Sleepiness

11-12 Mild Excessive Daytime Sleepiness

13-15 Moderate Excessive Daytime Sleepiness

16-24 Severe Excessive Daytime Sleepiness

For practical purposes, the ESS is a unitary scale that is reliable and valid for measuring a person's ASP. It is very cheap and easy to use for individuals and large groups.

Psychometrics of the ESS

The psychometric properties of the ESS have been investigated widely. The internal consistency of responses to the eight questions has been tested by Cronbach's alpha, which has varied between 0.73 and 0.90 (mean = 0.82) in ten separate investigations (eg. Johns,1992; Hagell et al, 2007). The test-retest reliability of ESS scores (measured over a few weeks to a few months) has been tested by the intraclass correlation coefficient which has varied between 0.81 and 0.93 in five separate investigations (eg. Gibson et al. 2006; Izci et al. 2007; Cho et al 2011; van der Heide et al. 2015).

Principal components analysis of ESS item-scores has yielded variable results, with a single factor in some investigations, but more than one factor in others. In the latter investigations there has been one dominant factor, as well as one or two minor factors with Eigenvalues not much above 1.0, the assumed cut-off point (eg. Johns, 1992; Sargento et al, 2013). We might conclude that there is one dominant factor, with high loadings on all scales, but sometimes there are additional minor factors that vary between groups.

Rasch analysis of ESS item-scores has enabled differences between the items to be assessed at the same time as differences between people, based on Item Response Theory. This analysis has confirmed that the ESS involves an ordinal sequence of Items, from Item 5 (the least 'difficult') to Item 6 (the most 'difficult'), which can be interpreted in this context as differences of somnificity. The evidence from several different Rasch analyses of the ESS

indicates that it has a unitary structure (Hagell, et al, 2007; Izci et al, 2008; Sargento et al, 2015).

External Criterion Validity of the ESS

Strong evidence for the external criterion validity of the ESS has come from investigations of the sensitivity and specificity of ESS scores for distinguishing narcoleptic patients from normal controls, who have very different ASPs by definition (Parkes et al, 1998; Johns, 2000(b).

A functional MRI study of ‘normal’ adults has shown that those with higher ESS scores (even within the ‘normal’ range) have lower connectivity between the bilateral thalamus and cortical regions involved in somatosensory and motor functions in the resting awake state (Kilgore et al, 2015).

The external criterion validity of the ESS has also been tested, less conclusively, by the correlation between ESS scores and mean sleep latencies in the Multiple Sleep Latency Test (MSLT). It has been shown repeatedly that this is not a close relationship, statistically significant in some but not all reports (eg. Johns, 1992; Sangal et al, 1997; Chervin, et al, 1999).

The external criterion validity of the ESS has also been tested by examining the relationship between ESS scores and the severity of obstructive sleep apnea, measured by the apnea-hypopnea index (AHI). That too is not a very close relationship, usually but not always statistically significant (Manni et al 1999; MihÄƒicutÄƒ et al, 2013). That is also true for the relationship between the severity of OSA and measures of daytime sleepiness other than the ESS, such as the MSLT (Guilleminault et al, 1988). We might conclude that such relationships are of limited use for testing the validity of any method for measuring daytime sleepiness, whether subjective or objective.

The responsiveness of ESS scores to treatment effects has been demonstrated by their reduction after nasal continuous positive airway pressure treatment for obstructive sleep apnea (Standard Response Mean >0.8) (Chen et al, 2002; Hardinge et al, 1995), and also after the treatment of narcolepsy with stimulants (Broughton RJ, et al, 1997; van der Heide et al. 2015).

Translation of the ESS

The ESS was first developed in English for Australia, but has been translated into many other languages, especially by Mapi Research Trust who have used standardised procedures. For the ESS to remain useful internationally it is important that it is standardised and not modified. In languages other than English, it is important that the meaning of the original (English) words be retained. The copyright prohibits any changes to the ESS, except under special circumstances and with written permission.

Limitations of the ESS

Because ESS item-scores are based on subjective reports, they can be influenced by the same sources of bias and inaccuracy as any other such reports. The ESS should not be used in isolation in circumstances where the scores could determine outcomes with potential legal implications, such as granting or withholding a driver's license. Confirmatory evidence of 'excessive daytime sleepiness' or an increased risk of a drowsy road-crash should be sought from other sources too.

The ESS does not usually enable accurate predictions to be made of a person's level of drowsiness, and hence their crash-risk, when driving a vehicle at some particular time. However, there may be an exception to this among people with very high ESS scores (>15), whose ASP is very high under most circumstances.

The ESS does not distinguish which factors, or which sleep disorders, have caused any particular level of ASP. The ESS is not a diagnostic tool by itself. Nor does it assess other aspects of a person's sleep habits, for which other methods are available.

The ESS is not suitable for use among people with serious cognitive impairment. Nor is it suitable for measuring rapid changes in sleep propensity over periods of hours, eg. to demonstrate the short-term sedative effects of a drug, or to assess the circadian rhythm of sleep propensity.