Chapter 6

Human response to sound

Ian Flindell

6.1 INTRODUCTION

6.1.1 Objective standards

Human subjective and behavioural response to sound is related to physical sound levels, but can also be considerably affected by many other variables according to individual sensitivities, attitudes and opinions. This can be a problem for standards and regulations which are intended to reduce annoyance, sleep disturbance or effects on health, but which have to be based on physical quantities. It is not practicable to use actual annoyance, sleep disturbance or effects on health when setting standards and regulations or defining contracts. It is desirable that the physical metrics used for noise limits, targets and benchmarks should reflect human response, even if they are not directly connected. Some physical metrics are better than others in this respect, and this chapter sets out to explain why this is so and to outline where the major sources of uncertainty can be found. Human response to vibration is dealt with separately in Chapter 7.

6.1.2 Uncertainties

Acoustic metrics can be used to predict or estimate average subjective or behavioural response, but only by assuming homogeneous populations. Real people are not homogeneous. Individual response varies considerably above and below the mean, and this means that, depending on the characteristics of the population, actual average response varies above and below any overall central tendency. Generally speaking, around one-third of the total statistical variance in human response can be attributed to differences in measured physical sound levels (depending on the metrics used); another third of the variance can be attributed to many so-called non-acoustic variables such as the situation and context in which the sound occurs, and psychological and physiological differences in sensitivity, susceptibility and attitudes towards the source of the sound; and the remaining third of the variance is effectively random and therefore unpredictable.

Many of these problems arise because sound is a naturally occurring and mostly useful physical phenomenon. Sound only becomes harmful when it has some negative or adverse effect and can then be defined as noise. Sound level meters cannot, unfortunately, tell the difference between sound and noise. Only people can do this.

Standard sound level metrics such as L_{Aeq} and L_{AFmax} (see Section 6.3.4) are administratively convenient but do not reflect differences in the character of the sound, or sound quality, which can be as or more important than sound level alone. People pay selective attention to different features of a sound in different situations. There is no universally applicable method for predicting which features will be attended to.

6.2 AUDITORY ANATOMY AND FUNCTION

6.2.1 Research methods

Human listeners perceive and respond to different kinds of sound and noise in different ways. Human response depends on the physical structures and physiological functions of the auditory system and on the different ways in which central neural and endocrine processing systems are connected to, and stimulated by, the peripheral auditory systems. The anatomy is well understood. However, understanding the detailed function of each component part is more difficult, because investigation requiring dissection unavoidably interferes with normal function.

Psychophysical methods are used to observe the time-, frequency- and sound-level-resolving powers of the ear by comparing subjective impressions between different types of sound. Subjective descriptions or reports of perceived magnitudes have greater uncertainties than *just-detectable thresholds* and *just-noticeable differences*. Just-noticeable thresholds and differences can be measured by asking a listener to press a button when they can detect an audible difference or identify which of two similar sounds is louder or differs in some other defined way. Performance can be strongly affected by motivation and generally improves with practice and then deteriorates with increasing fatigue. Experimental design requires compromise between conflicting sources of potential bias. Subjective reports are affected by additional uncertainties because of possible differences in interpretation of instructions and explanations of tasks. The results of psychophysical tests can be related to the underlying structures by comparing auditory capabilities before and after damage caused by disease or trauma, or by comparing auditory capabilities against theoretical models based on the anatomy. Nondestructive scanning techniques are increasingly being used to observe normal functions under different conditions and seem likely to make an increasing contribution in the future, but are outside the scope of this chapter.

Additional and often unknown variance is contributed by differences in physiological functions that are dependent on cell biochemistry, which can be very sensitive to various chemicals and hormones circulating in the blood. However, it is not usually practical to be able to measure or control hormone concentrations, except in specialist research.

It would be useful if classical introspection could provide insights, because this research method requires no equipment or special facilities. Unfortunately, people rarely have much conscious awareness of how they actually hear things, and instead just perceive the meanings of sounds to which they have actually paid attention. It can be very difficult to 'hear' sounds as they really are, particularly if they are unfamiliar. However, thinking about how particular anatomical structures might have evolved and to what biological purpose can sometimes contribute valuable and important insights. No single research method can provide complete understanding and experimental design always requires compromise to balance conflicting sources of bias.

The next sections describe the anatomy and how this is related to the observable functions. A simplified diagrammatic cross-section of the ear is shown in Figure 6.1.

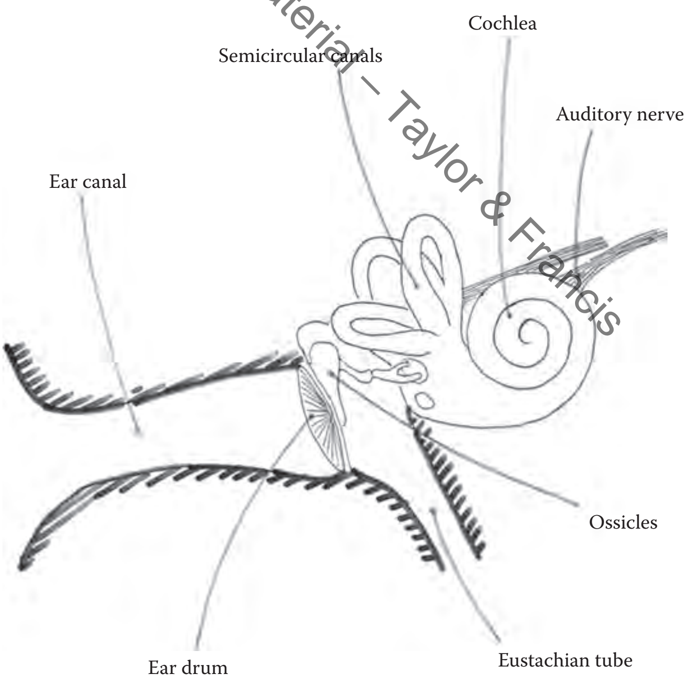

Figure 6.1 Simplified cross-section of the human ear.

6.2.2 External ear

Why do humans have two ears, one on each side of the head? Providing more than two auditory sensors could improve spatial perception and redundancy against damage or disease. Having only two ears is probably an accident of evolution. However, two ears are better than one, because *binaural hearing* improves sampling of complex sound fields, helps to discriminate between sounds coming from either side of the head and still provides some redundancy. Differences in relative amplitude and time of arrival between sounds arriving from either side help the central nervous system to localise sources on the left or the right much faster and with greater precision than would be possible with one ear alone. In fish ancestors living underwater, this function could have evolved from directional sensitivity to vibration inputs coming from different directions. Directional sensitivity confers a survival advantage from being able to quickly identify and then move away from, rather than towards, vibration that generates threats. Underwater predators additionally benefit from being able to quickly calculate the difference between threats and prey, so as to be able move either away from the threat or towards the prey, and thereby acquiring additional auditory neural processing.

Binaural recordings require two microphones in positions representative of the ears on a real head. Microphones can be put on a headband to record exactly the same sounds as heard by the person wearing the microphones, or they can be put on a dummy head, or simply spaced apart with either a solid disc or nothing (i.e. air) between them. On a real head or on a representative dummy head, the microphones should be capable of recording the full range of interaural differences in amplitude and time of arrival which would have been heard by a listener at the same position in space and time. Simplified binaural systems using an intervening solid disc or air gap between the two microphones reflect many, but not all, of the *interaural differences*. The main technical defect of binaural systems is that they cannot easily record the differential effects of head rotation and other movements on the interaural differences which provide important cues in real-life listening.

In real-life listening, if the head is fixed in position, it is difficult to tell whether sounds are coming from the front or the back. In real life, rotating the head helps listeners to resolve front-to-back and other directional ambiguities. Rotating the head when listening to binaural recordings using headphones has no effect on interaural differences. This encourages perceived stereo images to collapse into the centre of the head. When listening to binaural recordings, some listeners are better than others at disregarding or ignoring conflicting head rotation cues, leading to a wide range of subjective impressions. Binaural recording systems have been advocated for noise assessment and regulation purposes, on the premise that they are more representative of human hearing than standard measurements using single omnidirectional microphones, although the impossibility of being able to represent head rotation effects retrospectively somewhat undermines this philosophy.

In humans, the *external ear* or *pinna* is small compared with the size of the head. The pinna is not acoustically very significant except at higher audio frequencies, where the wavelengths are comparable. In human evolution, having an increasingly large brain sticking up above the ears has presumably been more beneficial for survival than having bigger ears sticking up above the head. The various folds and convolutions in the external ear provide directionally dependent acoustic filtering in both azimuth and elevation at higher audio frequencies where the wavelength is small compared with the physical dimensions but have no effect at lower frequencies. There is some evidence that elevation-dependent filtering by the external ear contributes to the perception of height, but only for sounds with significant high-frequency content. Readers can easily demonstrate this for themselves by holding their own ears flat to their head using their hands.

Binaural hearing helps to focus attention on sounds coming from directly in front while simultaneously helping to discriminate against background noise coming from all other directions. By cupping one hand behind each ear, readers can easily demonstrate that bigger ears can assist forward discrimination and can also 'improve' perceived stereo images when listening to two-channel loudspeaker stereo hi-fi systems. However, stick-on Mickey Mouse ears have not caught on as upgrades for stereo hi-fi systems, because it would be difficult to make them expensive enough.

It is interesting to speculate on what might have happened if humans had evolved bigger ears in proportion to their increasingly large brains. The very small contribution made by the current design of external ear to source height localisation at higher audio frequencies would have been extended lower down the audio-frequency scale, but the inconveniences of getting caught up on overhanging branches or being more visible to predators from greater distances might have offset these advantages. It is also possible that the shape and size of the external ear in humans could have been influenced by sexual selection without having very much to do with auditory function at all. The limited information available from the fossil record provides no definitive answers.

6.2.3 Middle ear

The middle- and inner-ear systems are protected within the lower part of the skull. Incident sound waves produce pressure disturbances at the entrance to the *ear canal* which are transmitted by wave motion to the sensitive *eardrum* which terminates the canal. The ear canal is acoustically resonant in the upper midrange of audio frequencies from around 3 to 4 kHz, depending on precise dimensions which vary between different individuals. Some authorities suggest that this resonance is the primary function of the ear canal, but a more plausible explanation is that the dimensions evolved simply to provide mechanical protection. Human infants with ear canals wider than their little finger would soon suffer damaged eardrums. In theory, fatter fingers could have provided an alternative evolutionary solution, but would have compromised operating mobile phones or picking one's nose.

The eardrum vibrates in response to rapid variations of sound pressure in the ear canal. The real-ear, head-related *transfer function* describes the relationship between the complex frequency spectrum of sound pressure at the position of the eardrum with no listener present and the spectrum at the eardrum. This transfer function varies for different angles of azimuth and elevation relative to the extended axis between the two ears, and between different individuals. This means that the frequency spectrum at the eardrum (which is what a person hears objectively) will be different from that recorded using an omnidirectional instrumentation-grade microphone as specified in most measurement standards. Most standards are intended to represent the preexisting sound field at the position in space where the listener's eardrum(s) would have been if the listener had been present and without any filtering due to the head-related transfer function. This difference is not a problem in practice, because human listeners each have a lifetime's experience of their own real-ear, head-related transfer functions, for which they make subconscious allowances when forming subjective impressions of the external acoustic environment.

There are special cases such as *bone conduction* or transmission via the *Eustachian tubes* at the back of the nasal cavity, or where hearing aids or headphones are used where a measurement in the ear canal may be more appropriate; but even in these special cases, it is often useful to be able to relate the measurements back to an equivalent free-field (listener not present) sound field for comparison purposes.

The air-filled *middle-ear cavity* behind the eardrum allows the eardrum to vibrate and to drive a chain of three small bones (the *ossicles*) to deliver vibration via a membrane (the *oval window*) to the fluid-filled cavities of the inner ear (the *cochlea*). The obvious function of the ossicles is to transmit vibrational motion from the eardrum across the air-filled middle-ear cavity, but they also provide leverage to convert motion from the relatively low *mechanical impedance** of the air outside to the relatively high mechanical impedance of the fluid-filled inner-ear cavities. The ossicles are thought to have evolved from jawbones in ancestor lizards which then became separated from the jaw to avoid interference when chewing food.

* For the purposes of this chapter, the mechanical impedance of fluids is considered as pressure over velocity rather than force over velocity in solids, as defined in Chapter 3.

Small muscles between the ossicles are able to tighten up the connections to provide limited protection against high-level transient sounds. This is known as the *acoustic reflex*, but it cannot protect against high-level continuous sounds. A fit person shouting can easily generate sound levels of 120 dBA or more at a distance of 1 m for up to a few seconds at a time. Close up to another person's ear canal, the sound levels are much higher and potentially damaging. The acoustic reflex provides some protection against this type of sound, but not against the type of continuous high-level sounds produced by machinery and heavy industry.

The middle-ear cavities are connected to the back of the throat and thence to the outside world via the Eustachian tubes, to allow drainage and equalisation of atmospheric pressure on both sides of the eardrums. The tubes can become blocked by inflammation and swelling of the lining, causing discomfort and reduced hearing sensitivity, which then recovers if the inflammation and swelling is temporary. Blockage is a common source of discomfort when the cabin pressure increases rapidly in a landing aircraft.

6.2.4 Inner ear

There are two flexible membranes in the dividing wall between the air-filled middle-ear cavity and the fluid-filled inner ear. The upper membrane, known as the *oval window*, accepts movement from the middle-ear ossicles. The lower membrane, known as the *round window*, vibrates in sympathy to release the internal pressure in the almost incompressible inner-ear fluid, which would otherwise prevent the oval window from moving. The inner ear comprises the parts concerned with hearing, the cochlea, and additional parts concerned with sensing motion, the semicircular canals. The cochlea is coiled through approximately two-and-a-half turns in a spiral. The *basilar membrane* divides the upper and lower chambers of the cochlea throughout its length, as shown in cross-section in Figure 6.2. It supports inner and outer rows of sensitive *hair cells* which are stimulated by the shearing action of the *tectorial membrane* across their tips whenever the basilar membrane as a whole is deflected up or down by vibrational waves in the cochlea fluid. The pattern of vibration along the basilar membrane represents the time-varying frequency content of incoming acoustic signals in the following way:

1. Incoming positive sound pressures deflect the eardrum, middle-ear ossicles and oval window inwards, creating a positive pressure in the upper chamber of the cochlea. The transmission of vibration from the ear canal to the cochlea involves a finite time delay, leading to frequency-dependent phase lag.
2. The positive pressure in the upper chamber causes downward deflection of the basilar membrane, with corresponding outward deflection of the round window directly beneath the oval window.

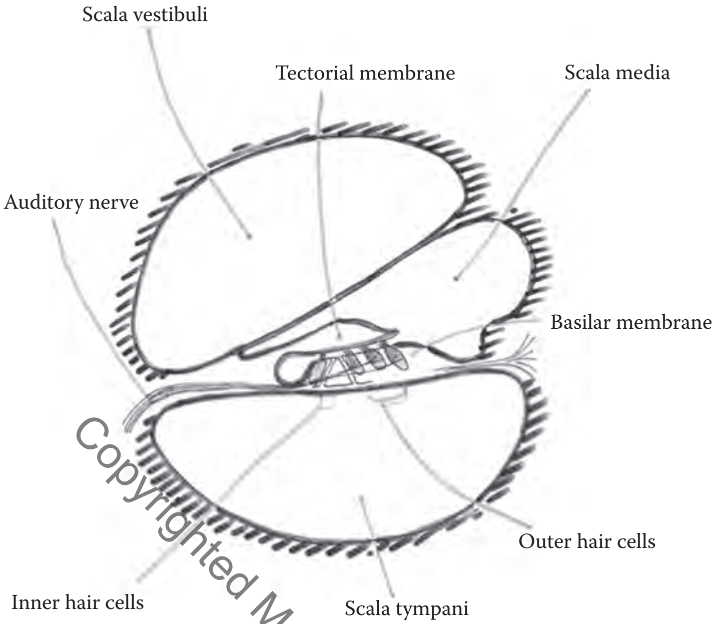

Figure 6.2 Cross-section of the basilar membrane.

3. The initial downward deflection of the basilar membrane progresses from the base end adjacent to the oval window towards the apex at the other end of the spiral as a travelling wave.
4. The travelling wave reaches a maximum amplitude at a position on the basilar membrane, which depends on the frequency content of the acoustic stimulus. The travelling wave is rapidly attenuated further along.
5. Low audio frequencies penetrate along to the apex of the cochlea spiral. The point of maximum amplitude moves further along from the base end (near to the oval window) as the frequency decreases. There is a physical connection between the upper and lower chambers at the apex known as the helicotrema which contributes to the low-frequency cut-off. Higher audio frequencies penetrate only a short distance from the base end along the basilar membrane. There is a frequency-dependent phase shift of the travelling wave along the basilar membrane.

The different vibrational patterns of the travelling wave at different frequencies act as a mechanical frequency analyser. The sensitive hair cells along the basilar membrane are organised such that auditory nerve impulses originating from different parts of the basilar membrane are associated

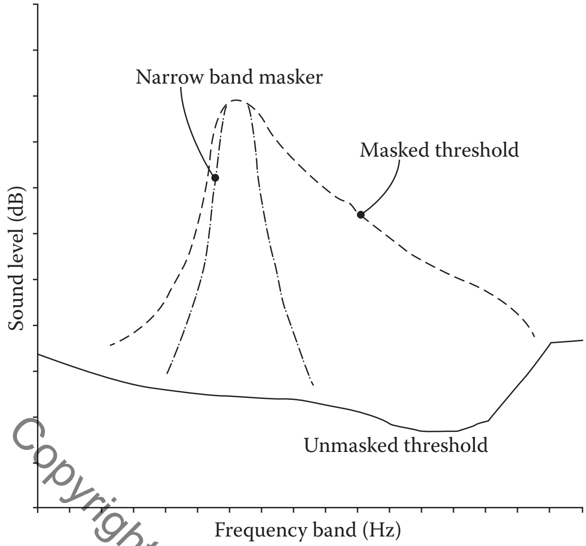

Figure 6.3 Upward spread of masking.

with different frequencies in the incident sound. The travelling wave passes through higher-frequency regions before reaching the point of maximum amplitude, and then decays rapidly. Lower-frequency sounds make it harder to hear higher-frequency sounds (this is known as the *upward spread of masking* – see Figure 6.3) but not the other way around. The travelling wave from a pure-tone test sound at a higher frequency than a masking noise reaches its maximum amplitude at a position along the basilar membrane, which is already stimulated by the lower-frequency masking noise, and this can make it harder to detect. The travelling wave from a pure-tone test sound at a lower frequency than a masking noise reaches its maximum amplitude at a position further along the basilar membrane beyond the area stimulated by the masking noise and can thereby be more easily detected or identified. One practical consequence of this design fault is that if the bass in a rock band is too loud, it can be difficult to hear the singer at higher audio frequencies.

The inner rows of hair cells appear to be mainly responsible for sending nerve impulses to the brain. Each nerve impulse is a single electrochemical discharge which travels from the cell body along the associated nerve fibres to the next connected neural interchange point. Every nerve impulse consumes biochemical resources which must be replenished before the cell can fire again. Consequently, there is a maximum firing rate for each nerve cell, and an associated and much lower resting firing rate, when the nerve cell fires spontaneously even when there is no input stimulation. The precise time at which each nerve cell is most likely to fire is uncertain, because it depends on the elapsed time since the last firing, the magnitude of input stimulation and the rate of biochemical recovery since the last impulse. Large-amplitude input signals generate stronger vibratory stimulation of hair cells and thus a higher firing rate for each associated nerve fibre. The dynamic range of individual hair cells is limited and the most sensitive connected nerve fibres become saturated at higher input sound levels. Other less sensitive combinations of hair cells and nerve fibres take over at higher sound levels to overcome the problems of the limited dynamic range of each hair cell and nerve fibre combination considered separately.

Wideband transient signals excite the basilar membrane throughout its length, causing complex patterns of nerve impulses to be sent along the auditory nerve, with the high-frequency-connected nerve fibres excited first. Continuous pure tones and narrowband frequency components stimulate continuous sequences of nerve impulses in the nerve fibres most closely associated with the part of the basilar membrane which is most strongly excited. The nerve-firing rate diminishes over time, as the biochemical resources of each hair cell and associated nerve fibres become depleted. The capabilities of the overall system with respect to relative phase information are limited by the general time uncertainty of separate nerve impulses, which exceeds the period of incident acoustic signals at middle and higher audio frequencies. There is some evidence that volleys of nerve impulses are phase-locked to the incident acoustic signal at lower audio frequencies, but it is not clear in what ways, if any, this information might be used.

For many practical purposes, the peripheral auditory system can be modelled as a bank of overlapping band-pass filters operating in real time, but it should be noted that there is no point in the real system at which the separate output of each modelled band-pass filter could be observed. The effective frequency selectivity of each modelled band-pass filter is of interest, and while this is generally assumed to be approximately equivalent to one-third of an octave over most of the audio-frequency range, the measured equivalent bandwidths vary depending on the method and conditions of measurement.

There is evidence that the outer rows of hair cells assist in fine frequency tuning. The outer rows are believed to vibrate in sympathy with incoming acoustic signals and thereby assist in tuning the system to narrowband and discrete frequency input signals. This enhances both frequency selectivity and sensitivity to weak sounds. It has been difficult to study this hypothesised active process *in vivo*, because access to the systems involved cannot be obtained without drastically interfering with or destroying them. The active process assists the ear to achieve the best compromise solution between time and frequency resolution that can be achieved. Because of *time–frequency uncertainty* (see Section 4.3.4), it is not possible to have both fine frequency resolution and fine temporal resolution in the same physical system at the same time. Fine frequency resolution requires that a sufficiently large number of cycles of a continuous waveform are observed to allow discrimination from slightly higher and slightly lower frequencies. For example, and in generic terms, it is necessary to observe 100 cycles of a cyclically repeating waveform within a fixed time period to be able to determine the frequency to within ±1%. Resolving the frequency of a 200 Hz tone to within 1% requires an observation time of around half a second, which could be too long to wait if an immediate reflex action is required.

In survival-critical situations, a rapid time response may be more important than fine frequency resolution. However, if a detected incoming acoustic signal is still present after the first few tens of milliseconds, an increasing capability to be able to resolve frequency content then becomes important for cognitive appraisal of the sound. Varying auditory capabilities in terms of frequency resolution increasing over short timescales from the initial transient appear to be consistent with the temporal and spectral variation of normal speech signals, suggesting synergies between the evolution of capabilities to produce speech sounds and the evolution of auditory systems intended for their resolution.

6.2.5 Auditory nerves

The final link in the chain is the system of auditory nerves which transmit nerve impulses from the hair cells in the basilar membrane to the brain. The auditory nerves pass through a number of interconnection points where nerves branch off to different parts of the brain and eventually reach the auditory cortex, where most of the complex processing involved in selective attention and cognitive appraisal is believed to take place. There is evidence of the presence of specialised nerve cells which are sensitive to specific acoustic features present in incoming acoustic signals, but it is not clear to what extent these systems are determined by generic coding or develop naturally as a result of learning and experience. Multiple cross-connections between the left and right auditory nerves facilitate binaural and higher-order neural processing, which support source localisation and other complex perceptual functions. It seems unlikely that any simple model would be capable of representing the full complexity of the overall system.

6.2.6 Overview

The auditory system does not operate in the same way as a microphone. Incident sounds are represented by time-varying patterns of nerve impulses across parallel auditory nerve fibres, which are then interpreted by higher-order neural processing in the brain. People do not perceive incoming sounds as they really are. Instead, the brain constructs internal perceptual images of the external environment based on the overall pattern of incoming neural signals interpreted in the light of expectation and experience.

People are not always consciously aware of all sounds that are technically capable of being heard. Unfamiliar sounds might be completely unrecognisable. Conversely, it is also possible to perceive sounds that are not physically present, or to hear sounds not as they really are. Auditory consciousness remains a mystery.

6.3 AUDITORY CAPABILITIES AND ACOUSTIC METRICS

6.3.1 Auditory thresholds

The normal range of hearing is shown in Figure 6.4. The normal threshold-of-hearing curve at the bottom of the figure shows maximum sensitivity between 3 and 4 kHz, corresponding to the first quarter-wave resonance of the open-ended ear canal. In this frequency range, for healthy young adults with no hearing loss from any cause (defined as normal hearing), the minimum *r.m.s. sound pressure* (see Chapter 2) that can be just detected when listening in a diffuse field with both ears (defined as the *minimum audible field*) is around 10–20 μPa, which is equivalent to −6 to 0 dB sound pressure level (SPL or L_p) relative to 20 μPa when expressed in decibel terms. The corresponding *minimum audible pressure* measured in the ear canal is generally a bit higher than the minimum audible field and is similarly variable between different individuals. The differences between minimum audible field and minimum audible pressure may need to be taken into account

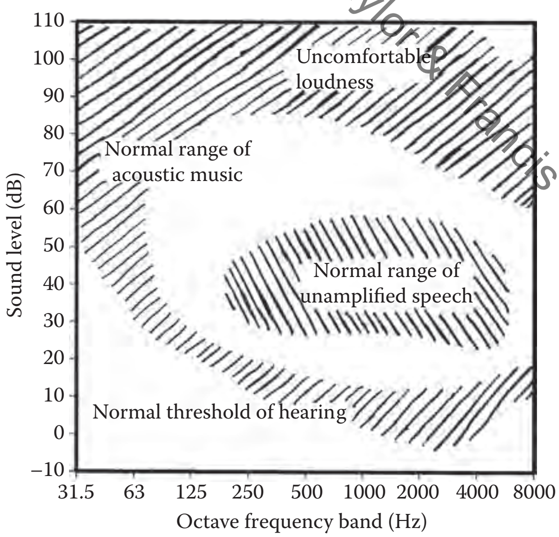

Figure 6.4 Normal range of hearing.

by audiometric equipment designers, but are not otherwise of any great concern.

'Just detected' means that the presence of the sound can be correctly identified at slightly better than chance level. Low-level sounds at the just-detected threshold would not necessarily be of any practical significance for everyday conditions, unless the listener had become particularly sensitised to those sounds for any reason. Individual thresholds vary above and below the mean at different frequencies, even within a sample of people where everyone has so-called normal hearing. It is not clear how much of this variability represents fundamental differences in hearing sensitivity or how much can be attributed to measurement uncertainty. Some people tend to be rather cautious when asked to press a button to show they can hear a particular sound, while others tend to press the button even when they are not quite sure if they can really hear the sound or not.

The normal hearing thresholds increase significantly below 300 Hz and above 8 kHz. There are upper and lower limits to the range of frequencies that can be heard by people with normal hearing. Children may be able to hear higher-frequency sounds up to around 20 kHz, and in some cases even higher, but a more practical upper limit for most adults would be around 15 kHz or less. It is not clear what survival advantages, if any, would accrue from being able to hear very high audio frequencies, which are mostly highly directional and potentially more useful for echolocation than general perception.

Most people are equally insensitive to very low audio-frequency sounds, particularly below 50 Hz. There is very little information content in speech sounds below about 300 Hz. Most natural low-frequency sounds are generated by distant sources which are less important for individuals than nearby sources. If there had been any significant survival advantages from being able to hear very low-frequency sounds generated by distant thunderstorms or by large predators stamping their feet, then this would have required larger ear structures to have evolved, which might in turn have had other disadvantages, such being more likely to be seen and eaten by those same predators.

The upper limit of normal hearing is often quoted as being somewhere around 120 dB L_p, although there is no clear definition of what the upper limit actually means. In the midfrequency range from around 1 to 5 kHz, a pure tone at 120 dB L_p can be extremely unpleasant and would certainly become strongly objectionable very rapidly if continued. Midfrequency pure tones could be increased to 140 dB and still be 'heard', that is, within the range of audibility, although would nevertheless be highly uncomfortable or even painful to listen to and would certainly lead to significant temporary threshold shift (TTS) (see Section 6.4.2). A continuous tone at 140 dB L_p and more could cause permanent damage in the most susceptible individuals, even after only relatively short exposure. However, a 50 Hz continuous tone at 140 dB would not be so uncomfortable and might even be inaudible if the frequency was below 20 Hz. It should be noted that the measurement of auditory thresholds at very low audio frequencies can be complicated by the need to avoid higher-frequency harmonic distortion products or nonlinear acoustic resonances in the testing loudspeakers or in nearby objects which might be relatively more audible than the original very low audio-frequency tones.

At higher sound levels, broadband sounds tend to be less uncomfortable than pure tones. This could be because the vibratory excitation pattern on the basilar membrane is not concentrated in one place, as it is for pure tones. High-level pure tones tend to increase in unpleasantness with increasing frequency, until the frequency eventually goes above the normal audible range and the sound becomes inconsequential because it cannot be heard. Similarly, the harmonic distortion products of higher-frequency tones reproduced through imperfect loudspeakers become inconsequential above the audible frequency range.

6.3.2 Just-noticeable differences

The smallest just-noticeable differences between otherwise similar pure tones or narrowband sounds occur in the middle range of sound levels from 40 to 80 dB L_p and frequencies from 300 Hz to 5 kHz. This usefully corresponds to the sound-level and frequency ranges which provide the most differentiation between different speech sounds. Under laboratory listening conditions, it is quite possible within this most sensitive range to be able to correctly identify at better than chance level the louder or higher-pitched of two otherwise similar sounds down to level and frequency differences of <1 dB or <1% respectively. Any significant time gap between the two sounds results in reduced discrimination performance, which might more properly be described as a failure of auditory memory. The just-noticeable differences for sound level and intensity become progressively wider outside the middle range of sound levels and frequencies.

There is a weak correlation between subjective loudness and objective sound level. People are usually much better at making relative or comparative judgements than at judging absolute magnitudes. The average human listener can judge which is the louder or higher-pitched of two sounds without necessarily being able to judge the absolute magnitude or frequency of either sound on its own. Relative loudness is similar to relative pitch, in that subjective magnitude tends to be proportional to the ratio of objective physical magnitudes rather than the linear-scale values. This is one of the main justifications for using a decibel scale for measuring sound level. A difference between two sounds of x decibels always represents the same ratio of physical magnitudes irrespective of absolute sound level. A just-detectable difference in subjective magnitude between two sounds of 1 dB could represent a very small linear-scale difference of 0.001 Nm^{-2} at low sound levels or a much larger linear-scale difference of 1 or 10 Nm^{-2} at higher SPLs (Morfey 2001).

Unfortunately, or perhaps unsurprisingly, the general public has almost no understanding of how decibel scales actually work when applied to sound levels. The public mostly understand the word 'decibel' simply as a synonym for 'loud'. If acoustical engineers and physicists use the word 'decibel' when describing or explaining noise management to the general public, not many people will actually understand what they are talking about. This can lead to outcomes not meeting expectations and can contribute to misunderstandings and even mistrust.

Just-noticeable differences are wider under real-life conditions than the typically 1 dB differences in sound level and 1% differences in frequency observable under controlled-listening laboratory test conditions. Under real-life conditions, it is unlikely that the average listener would be able to correctly identify at better than chance level the louder of two otherwise similar aircraft flyover or road-vehicle pass-by events which differed in maximum sound level by < 3 dB. If the second of the two events was delayed by more than a few minutes, an even bigger difference would be required to achieve reliable discrimination. The relative probability of judging the most recent of two sounds as being louder increases with time delay, simply because the auditory memory of the earlier sound fades. This could be one of the reasons why the significant reductions in aircraft-noise sound levels which have been achieved over the past 30–40 years have not led to expected increases in public satisfaction.

At planning inquiries where noise is an issue, the parties often disagree about the significance of small changes in sound levels. Residents are often more aware of other changes in their own right than any small resulting changes in sound levels. For example, residents are more likely to notice a doubling of road traffic than the 3 dB increase in long-time average sound levels caused by the increase in traffic. Alternatively, increasing the proportion of heavy vehicles with no change in overall traffic flow could increase long-time average sound levels by similar amounts, but in this case, it would be the noisier heavy vehicles which would be noticed rather than the difference in long-time average sound levels.

The smallest just-noticeable differences for frequency are concentrated in the frequency range where the ear has its highest sensitivity, from about 300 Hz up to about 5 kHz. It is presumably just as beneficial to be able to discriminate between similar frequencies in this range as it is to be able to hear them in the first place. The smallest just-noticeable differences for sound levels are concentrated in the middle region between quieter and louder sounds from about 40 to about 80 dB L_p. The human ear appears to be well adapted to be able to discriminate between the different sounds involved in human speech. Which evolved first, the auditory discrimination capabilities of the human ear, or the sound-level and frequency differences produced by the vocal tract when generating and differentiating different speech sounds? Either both systems evolved in tandem, or there was an intelligent designer at work, and for the purposes of this chapter, it does not much matter which.

According to evolutionary theory, it might be naively assumed that all features of the human auditory system have been optimised for biological survival. However, this is not necessarily true. One aspect of hearing that suggests that evolution does not always come up with the best design solution is the *localisation gap* in the middle of the audio-frequency range, where auditory localisation does not work as well as at higher and lower audio frequencies. Interaural phase differences become increasingly ambiguous at frequencies above about 750 Hz, where they can exceed 180°, because the wavelength is small compared with the angle-dependent differences in propagation distance. Interaural intensity differences are too small to be detectable at frequencies below about 2 kHz where the head is too small compared with the wavelength of the sound to have any significant screening effect. In the middle audio-frequency range from about 750 Hz up to about 2 kHz, neither interaural phase nor interaural intensity difference make much contribution to auditory localisation of pure tones. With the standard ear spacing, it is harder to localise the position in space of a nearby talker than it would be if the ears had been much further apart. Perhaps being able to localise the position in space of a nearby talker was less important for biological survival than the presumably negative effects of having wide-spaced ears on stalks on either side of the head.

6.3.3 Equal-loudness contours

The internationally standardised equal-loudness contours (BSI 2003b), plotted in Figure 6.5, show how the relative loudness of pure tones varies across the audio-frequency range and across the range of sound levels from 0 to 120 dB L_p. Listeners are asked to adjust the level of a pure tone at a given frequency until it matches the subjective loudness of a reference pure tone at a fixed reference frequency (normally 1 kHz). Different listeners give different results, so composite loudness contours are averaged across statistically representative groups of listeners. Alternative psychometric methods used in different listening laboratories also give different results, suggesting that the precise shape of the current contours is generically representative of, but not completely determined by, underlying auditory capabilities.

One of the methodological problems is that comparing the relative loudness of pure tones at different frequencies is more difficult than it seems. Different listeners probably adopt different strategies when attempting this task. Comparative loudness judgements are much more subjective than observing just-detectable thresholds at different frequencies. In addition,

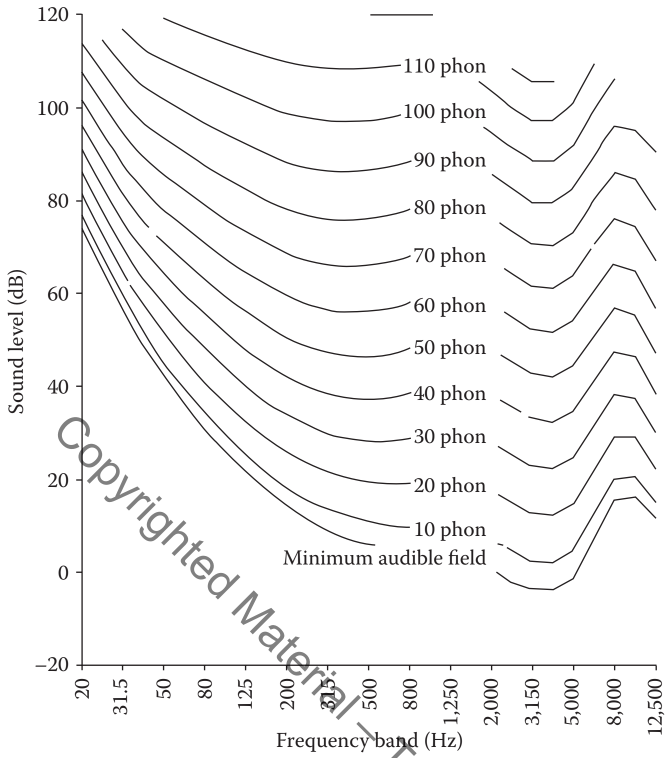

Figure 6.5 Standardised equal-loudness contours.

pure tones are rare in nature. Nobody has yet devised a method capable of deriving equivalent equal-loudness contours for more representative wide-band complex sounds.

Notwithstanding these uncertainties, equal-loudness contours can be assumed to represent fundamental properties of the peripheral auditory system as follows:

1. The higher hearing thresholds at the upper and lower extremes of the auditory frequency range (which are not in doubt because they can be measured objectively) appear to be reflected in reduced ‘loudness’ of very low and very high audio-frequency pure tones as compared with middle audio-frequency pure tones.
2. At the lower audio frequencies, the normal auditory threshold curve falls rapidly with increasing frequency and the corresponding equal-loudness contours above it flatten out at higher sound levels. This suggests that the relative or comparative loudness of low-frequency

sounds is proportionately greater at higher sound levels. However, it should be noted that the changing shape of the equal-loudness contours at higher sound levels could alternatively be an artefact of listeners preferring to avoid even higher sound levels when given experimental control.

3. Subject to the caveat noted in 2, the growth of subjective loudness with increasing sound levels appears to be much steeper at low and very high audio frequencies than in the middle frequency range. This is also associated with reduced sound-level and frequency discrimination outside the middle range of frequencies and sound levels.
4. Equal-loudness contours suggest that at 1 kHz, an average listener should be able to discriminate between 100 or more different sound levels between the auditory threshold at around 0 dB L_p and the so-called uncomfortable loudness level at around 120 dB L_p. At 50 Hz, the dynamic range between the auditory threshold and the uncomfortable loudness level is much less, and the just-noticeable difference for loudness is also much greater, suggesting that a far smaller number of different sound levels (and frequencies) should be separately distinguishable. This implies that bass players, for example, should be able to make more mistakes than violin players without the audience noticing.

The A-, B- and C-frequency weightings which were provided in traditional precision-grade sound level meters were intended to reflect the implied differences in the relative frequency response of the ear at different sound levels. The main function of a sound level meter is to measure sound levels as accurately and as repeatably as possible. These requirements are met by providing a precision condenser microphone, accurate and stable electronic systems, and appropriate and regular calibration. Problems arise when deciding which of the many possible acoustic metrics should be used for the measurement, each of which will give different results under different circumstances. A great deal of research effort has been expended over the past 50–60 years in various attempts to achieve higher correlations between sound level meter readings and subjective judgements to find the most appropriate acoustic metric. Hindsight has shown that this goal is unrealistic. It is not feasible for any single acoustic metric to represent all of the many variables which contribute to or determine subjective responses to noise, many of which are not even acoustic. On the other hand, the A-frequency weighting and the F- and S-time weightings are in general use for mainly practical reasons, and it is important that acoustical engineers and physicists should understand their advantages and limitations in everyday use.

For maximum objectivity, it is clearly desirable that the sound level meter should not impose any unnecessary time or frequency weighting or other nonlinearity on the measurement. Defined upper- and lower-frequency limits are desirable, so that inaudible infrasonic and ultrasonic frequency components can be excluded. Similarly, some form of averaging over short time periods is necessary, and it is desirable that the time weighting should not be unrepresentative of the equivalent time-averaging process in the human ear. In addition, any practical microphone imposes its own frequency response and high-frequency directivity onto the measurement. Precision condenser microphones can be set up in different ways to account for some, but not all, of these differences, but it is impossible to precisely match the frequency response and directivity of the human ear, particularly when binaural hearing is taken into account.

Most sound level meters available today are provided with a standardised *A-frequency weighting* setting in addition to an *unweighted* or *Lin* frequency response setting. Users are cautioned that unweighted or Lin frequency response settings in older equipment might not comply with current standards. Measurements using sound level meters with different unweighted or Lin frequency responses could give different results.

Most outdoor community noise standards specify the A-weighting. This reduces the sensitivity of the sound level meter at low and high frequencies compared with the upper-middle band of frequencies from around 1 kHz to around 4 kHz. The A-frequency weighting approximately follows the shape (inverted) of the 40 dB at 1 kHz equal-loudness contour. B- and C-frequency weightings may also be provided, particular on older instruments. The B- and C-weightings approximately follow the shape (inverted) of the 70 dB and 100 dB (respectively) at 1 kHz equal-loudness contours. The A-, B- and C-frequency-weighting curves are shown at Figure 6.6.

The original intention was that the A-, B- and C-frequency weightings should be used for the measurement of quiet, medium and loud sounds at around 40, 70 and 100 dB respectively. Three different frequency weightings were considered necessary, because the equal-loudness contours are steeply curved at the auditory threshold and become flatter at increasing sound levels. This well-intentioned scheme had two major defects. Approximating the assumed sound-level-dependent frequency response of the ear in this way does not account for many of the other acoustic (and non-acoustic) features that are likely to affect subjective impressions of the sound being measured. In addition, for sounds with prominent low- or high-frequency content, switching from one frequency weighting to another can have a considerable effect on the reading. To achieve consistency, selecting the appropriate frequency weighting for standardised measurements at different sound levels would have required arbitrary and precisely specified rules which have never been agreed.

Practical experience shows that the A-frequency weighting discriminates against predominately low-frequency wind noise generated at the microphone windshield when taking measurements outdoors and also

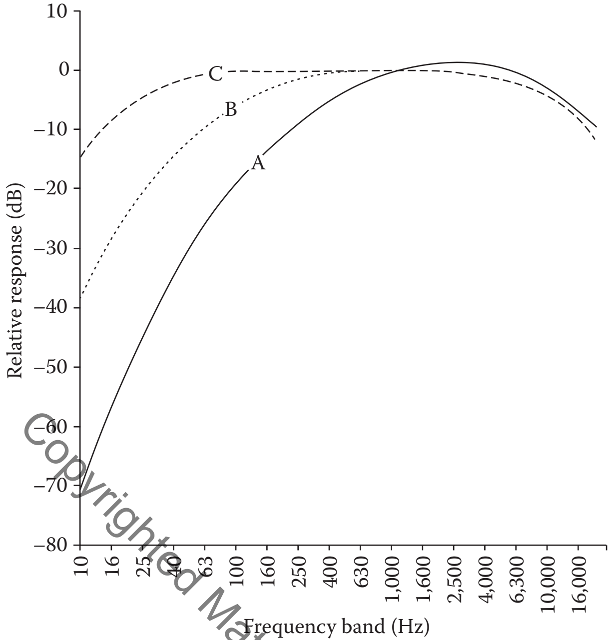

Figure 6.6 A-, B- and C-frequency weightings.

discriminates to some extent against low-frequency background noise from distant sources. These features alone are probably enough to explain its continued use. In addition, most field measurements carried out for research studies, and probably an even greater proportion of outdoor sound-level measurements carried out for many other purposes over the last 50–60 years, have used the A-weighting. Modest correlations have been obtained between A-weighted measurements of outdoor sound levels and subjective impressions of annoyance and disturbance across numerous noise-effects surveys. These correlations have often been wrongly accepted by many authorities as sufficient evidence that the A-weighting is the best weighting curve that could have been devised. In fact, for many field studies no other frequency weighting was tested, and for those studies where more than one frequency weighting was investigated, the two (or more) metrics are often so highly correlated with one another that no statistical distinctions could be drawn between them.

The B-frequency weighting has largely fallen out of use, mainly because it has not generally been found useful except for certain specialised applications. The C-frequency weighting is sometimes specified where there are low-frequency components present and there is concern that the

A-frequency weighting does not account for them properly. It should be noted that unweighted or 'flat' measurements are relatively more sensitive to extreme low- and high-frequency components (where present) than the human ear.

The A-frequency weighting has been compared against other frequency-weighting schemes in many laboratory-type listening studies. In theory, the derivation of the A-weighting from the 40 dB at 1 kHz equal-loudness contour means that it should become increasingly incorrect at higher sound levels, where it is most often used. It has been criticised for its indoor measurements, where typical outdoor-to-indoor building facade attenuation suppresses mid- and high-frequency components to a much greater extent than low-frequency components. A-weighted sound-level benchmarks and limit values for the interiors of houses might not, therefore, provide sufficient protection against low-frequency noise indoors for typical houses. The A-weighting is similarly inappropriate for measuring internal vehicle noise, because of the predominance of low frequencies inside the vehicle cabin, although it has been widely used for this purpose.

It should also be noted that the A-weighting has been criticised on the basis that on average, most people lose high-frequency hearing as they get older and that low frequencies might then become subjectively even more important than implied by the standard equal-loudness contours. While there is logic behind this suggestion, there is only very limited experimental evidence that any other frequency weighting provides a better match to older people's hearing.

The noise-rating (NR) frequency-weighting curves (BSI 1999) (see Figure 6.7) were devised to represent equal-loudness contours where measurements have been made using octave-band frequency-analysis filters. Many types of sound level meters can be fitted or supplied with octave and narrower bandwidth frequency filters to facilitate this kind of measurement. The generically similar noise-criteria (NC) and preferred-noise-criteria (PNC) curves were devised in the United States to meet the same requirement as the NR curves in Europe. NR curves have been widely used to define targets and limit values for sound levels inside buildings, taking into account both noise from building services and noise break-in from external noise sources. Typically, NR 10–20 is specified for audiometric test rooms, NR 20–30 for classrooms, NR 30–40 for quiet offices, and NR 60–70 for industry. NR 70 is suggested for industry as a basis for avoiding interference with speech communication, although in some sectors of manufacturing industry it might be considered rather optimistic in terms of what can be achieved in practice.

Of the frequency-weighting schemes which are widely known, the most complicated are the so-called loudness (ISO 1975) and noisiness (BSI 1979) calculation schemes. These schemes aggregate together the assumed separate contributions to overall subjective 'loudness' and 'noisiness' of each

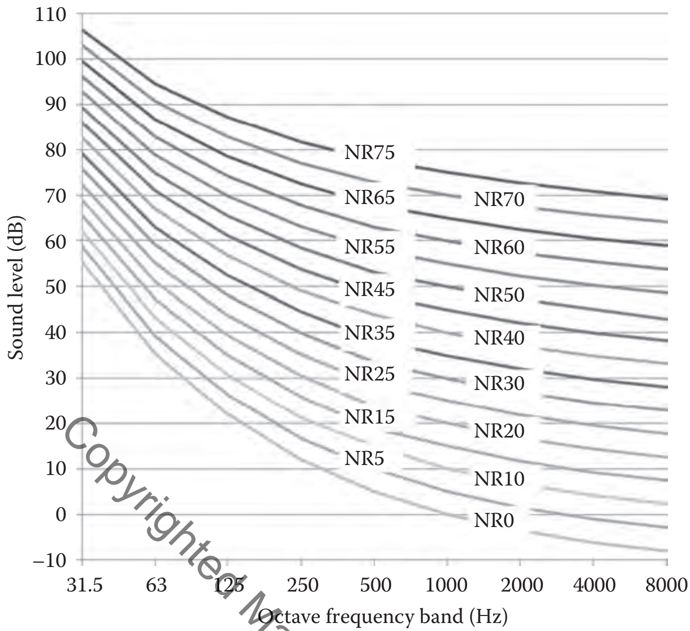

Figure 6.7 Noise-rating curves.

one-third octave band of the frequencies present. Zwicker's (European) and Stevens' (US) loudness-level calculation schemes have been the subject of considerable technical debate. All such schemes are based on more or less complex models of peripheral auditory processing. The results of experimental testing have not been universally conclusive. Unfortunately, no loudness calculation scheme yet devised seems capable of taking into account all possible acoustic features that might be present in any particular case.

Kryter's effective perceived noise level (EPNdB) calculation scheme has been the international standard for aircraft-noise certification purposes since the 1960s. It was originally modelled on Stevens' loudness-level calculation scheme, but was modified to deal with so-called perceived noisiness of aircraft flyover sounds rather than subjective loudness. Kryter's intention was that perceived noisiness should be understood as representing the objectionable or noisy features of aircraft flyover noise, in addition to the loudness, yet was not necessarily the same as disturbance or annoyance, which is felt by the individual person depending on the circumstances at the time, and could be argued to be outside the control of aircraft manufacturers. The EPNdB measurement scheme was developed to achieve a modest correlation with subjectively reported noisiness of aircraft types in service in the late 1950s and early 1960s, and takes particular account of the higher-frequency components common to the low-bypass-ratio turbojet engines used in that period. There is increasing evidence that this metric may not be optimum for representing the type of sound produced by more recent aircraft types powered by high-bypass-ratio engines or by advanced open-rotor-powered aircraft which might come into service in the future. Note that noise metrics such as EPNdB do not actually measure 'noise' anyway, because this is a matter of subjective perception.

Nevertheless, for assessment and regulation purposes, there is considerable merit in retaining standardised metrics such as EPNdB, even after they might have become less fit for purpose than at the time they were originally developed. Metrics become enshrined in legal contracts and specifications which could be difficult to change. In addition, it is not reasonable to expect any alternative metric to be able to perform any better over the longer term. Simply reducing the physical energy content in a sound will not necessarily lead to improvements in subjective sound quality, and if the wrong physical components or features of the sound are addressed, could even make matters worse. In some cases, annoyance can be reduced by addressing the psychological components of disturbance and annoyance without necessarily having to make any actual changes to the sound at all. There are also situations where actual physical reductions in sound level are too small to be directly noticeable and have no effect on disturbance and annoyance until people are told about them.

6.3.4 Time weightings and averaging

The selection of appropriate *time weightings* for objective measurements can be important. The time weighting determines the amount of time needed for a measurement. For completely steady, continuous sounds, the measurement averaging time has only to fit within the time for which the sound is present. However, most sounds that need to be assessed and possibly regulated are not steady. The ear requires a finite amount of time to register a sound. The various mechanical parts of the ear must first be set in motion. Then, the sensitive hair cells along the basilar membrane must be sufficiently stimulated to cause nerve impulses to be sent along the auditory nerve in sufficient quantity to alert the higher-level processing centres that something interesting is happening. Finally, the resulting neural activity must be sustained over a long enough fraction of a second that the central nervous system starts to process the new sensory input. If the sound then stops, any neural activity associated with the incoming sound then dies away quite rapidly.

The time-averaging response of the ear is best described by an approximately exponential averaging time of between 100 and 200 ms. Sound level meters can implement exponential or linear time averaging. Exponential averaging implies that the memory of recent events decays over time. Under linear averaging, however, everything happening within a fixed time interval makes an equal contribution to the average. A continuously updated exponential average or a sequence of very short-time linear averages is appropriate for showing a continuously varying sound-level history over time, whereas long-time linear averaging is appropriate for producing a single number to describe the average over longer time periods such as a 16 h day. Long-time-averaged sound levels such as $L_{Aeq,16\,h}$ are widely used for regulatory and contractual purposes for mainly administrative reasons, even though they do not reflect human hearing particularly well.

Transient sounds of shorter durations than around 100 ms are harder to detect and do not register the same loudness as continuous sounds of the same max or peak level. Figure 6.8 shows a generic curve of constant detectability for transient sounds with the sound level (when the sound is present) adjusted to maintain constant integrated acoustic energy (proportional to sound pressure squared × time) across the range of durations tested. Figure 6.8 shows that over the middle range of durations from about 15 ms up to about 150 ms, detectability appears to be a function of energy. As the duration increases beyond about 150 ms, detectability decreases. This is because the instantaneous sound level must decrease with increasing duration to maintain constant energy. Further increases in duration beyond 150 ms no longer offset the reduction in instantaneous sound level.

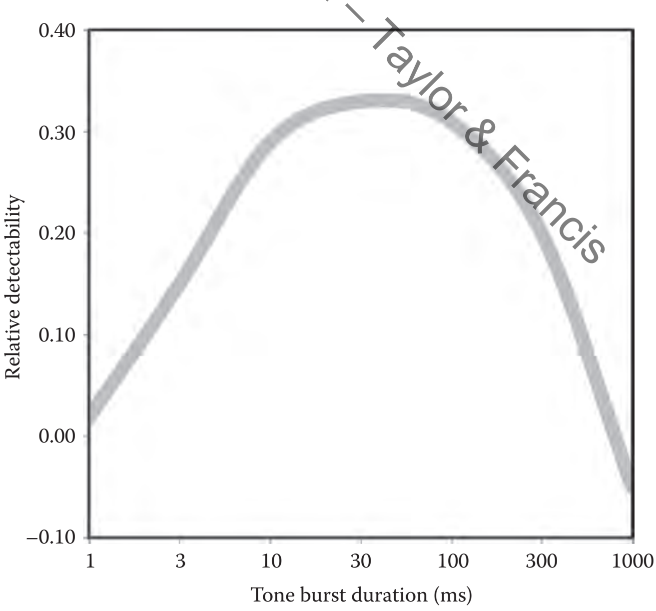

Figure 6.8 Relative detectability of constant energy short-duration sounds.

At very short durations below about 15 ms, detectability also decreases because time–frequency uncertainty causes the acoustic energy contained in the transient sound to be spread over an increasingly wide range of different frequencies. Consequently, there is less energy available to support detection in each separate part of the basilar membrane.

The time range of constant detectability varies with frequency, as might be expected, being shorter for high frequencies and longer for low frequencies. Similar curves can be plotted at sound levels above the detectability thresholds by using subjective loudness instead of detectability as the outcome variable. In this case, the precise shape of the curve varies, depending on the actual experimental method used. The comparative loudness of transient tone bursts of varying duration is essentially subjective and the judgement criteria used by different listeners is likely to vary. The subjective sound quality of short transient sounds varies with duration in addition to subjective loudness, and differences in sound quality can be easily confused with differences in loudness in this situation. This is not the case for detectability at the threshold, which is much more dependent on the fundamental capabilities of the auditory system, although differences in listener effort and motivation can still have significant effects.

One practical consequence of the finite integration time of the ear is that the subjectively perceived loudness of very short transient sounds might not be commensurate with the damage risk associated with high instantaneous peak sound levels. This is a good reason for setting the action levels for peak sound pressures in the *Control of Noise at Work Regulations* (2005) conservatively. These regulations specify that the peak SPLs should be measured using the C-weighting, presumably to discriminate against any very low- or very high-frequency content which might also be present, and expressed in decibels relative to 20 μPa.

The time weightings used in sound level meters (see Chapter 8) reflect what is known about the averaging time of the ear while avoiding unnecessary or impractical complications. Many sound level meters include both F(ast) and S(low) time weightings. The F-time weighting has a 125 ms exponential time constant and is intended to broadly represent the equivalent integration time of the ear, although without taking into account any differences at different frequencies. For rapidly fluctuating sounds such as normal speech, the instantaneous readings of a sound level meter set to the F-time weighting can change up and down too quickly to be followed by a human operator reading the sound level meter display. The S-time weighting, with an exponential time constant set at 1 s, averages out the instantaneous readings from a rapidly fluctuating sound, so that the slowed-down fluctuations can be followed by a human operator reading the sound level meter display visually. On the other hand, the S-time weighting is too slow for the sound level meter display to follow all of the rapid fluctuations in normal speech, and where these fluctuations are important, the F-time weighting should be used.

In practice, the maximum and minimum instantaneous sound levels of short duration or rapidly varying sounds will be underrepresented by the S-time weighting to a greater extent than by the F-time weighting. The peak (instantaneous) sound levels of short transient sounds with durations <125 ms will be underrepresented by both the F- and S-time weightings, but the maximum instantaneous sound levels of longer transient sounds with durations between 125 ms and 1 s will only be underrepresented by the S-time weighting. The different time weightings can be used for different purposes. For example, when measuring the maximum sound levels of aircraft flyover events, it is standard practice to use the S-time weighting. This is because the maximum S-weighted sound level is generally considered to provide the most reasonable indication of the sound level produced by the aircraft flyover event, where the effect of rapid random fluctuations caused by atmospheric turbulence along the acoustic propagation path from the aircraft source to the receiver has been largely averaged out. Rapid fluctuations in sound level associated with atmospheric turbulence cause the F-time weighted sound-level time history to fluctuate both above and below the equivalent S-time weighted sound-level time history and are considered (by the industry) to be outside the aircraft manufacturer's and operator's control. The amount of atmospheric turbulence varies depending on meteorological conditions at the time.

While long-time averages can be calculated from a series of readings taken from the time-varying display of a sound level meter, it is better to use digital averaging to obtain the most accurate results. Early analogue sound level meters have been superseded by digital sound level meters, which can now accomplish via software what used to require very expensive analogue circuitry to perform.

The widely used long-time-averaged sound level or L_{Aeq} metric is formally defined as the time-average or equivalent continuous sound level, and is widely specified in various standards and regulations. In the acronym for L_{Aeq}, the 'A' denotes the use of the A-frequency weighting. The formal definition recognises that a continuous sound with the same L_{Aeq} as a time-varying sound would have the same energy content as the time-varying sound, provided that the averaging time or measurement duration is the same in both cases. The L_{Aeq} is calculated by dividing the integrated A-weighted sound pressure squared by the measurement duration. If the same amount of sound energy or number of events is averaged over a longer measurement duration, the resulting average will be numerically lower, although the sound being measured will have been unchanged. Extending the measurement duration to include periods of relative silence before and after a single noise event reduces the average. Under these circumstances, doubling the measurement duration or halving the number of separate noise events included within it decreases the L_{Aeq} by 3 dB. The duration of the measurement must always be specified or implied by the context in which the metric is used. It is good practice to add the measurement duration as a suffix, for example $L_{Aeq,16\ h}$ or $L_{Aeq,5\ min}$. Where there is diurnal variation over the 24 h, it is good practice to state the start and end times of the measurement as well.

For environmental and community noise, it has become standard practice to report L_{Aeq} over 12 h or 16 h daytime periods and 8 h night-time periods, or averaged over the full 24 h. The European Environmental Noise Directive (European Parliament/Council of the European Union 2002) specifies L_{den} and L_{night} as harmonised acoustic metrics. L_{den} is the 24 h average L_{Aeq} with the evening period (19:00–23:00 h) weighted by +5 dB and the night period (23:00–07:00 h) weighted by +10 dB. The daytime period (07:00–19:00 h) is not weighted. Regulatory authorities in the United States use a generically similar metric, the L_{DN}, which is the 24 h average L_{Aeq} with the night period (2200–0700 h) weighted by +10 dB. In the United Kingdom, and for largely historical reasons, the Department of Transport has used the term L_{eq} instead of $L_{Aeq,16\ h}$ for many years to describe 16 h daytime aircraft-noise exposure.

It should be understood that, for the L_{den} and L_{DN} metrics, the precise weightings (+5 dB, +10 dB) and the precise time periods over which they are applied are essentially arbitrary committee decisions. They are not based on definitive research. It is difficult to conceive of any possible research design which would be capable of determining exactly which time periods and weightings should be used in any particular acoustic metrics in any case. The reason for this is that field measurements using similar metrics are normally so highly correlated together that it is impossible to tell them apart statistically.

The sound-exposure level, or L_{AE} (the traditional abbreviation, SEL, is more easily remembered as the single-event level), is a specific metric applied to separately identifiable noise events. Clearly, the maximum instantaneous sound level on its own cannot reflect the duration or time variation during the event. L_{AE} is a measure of the integrated acoustic energy in the event normalised to a standard averaging time of 1 s and is calculated in the same way as L_{Aeq}, except that the integrated A-weighted sound pressure squared is not divided by the measurement duration.

The best way to appreciate the difference between L_{AE} and L_{Aeq} is to consider two extreme examples: a completely steady sound, and a single transient event preceded and followed by extended periods of silence. For a completely steady sound, the L_{Aeq} does not change with different measurement durations, because the average sound level during the measurement duration is constant. The L_{AE}, however, increases logarithmically as the total acoustic energy received at the measurement point increases with measurement duration. For a single transient sound surrounded by extended periods of silence, the L_{Aeq} reduces logarithmically as the long-time average acoustic energy received at the measurement point reduces with increasing measurement duration. For the same situation, the L_{AE} is independent of measurement duration as the total acoustic energy received at the measurement point is independent of the total duration. There is no difficulty in translating between L_{AE} and L_{Aeq} measurements, provided that the relevant measurement durations are known. For the measurement of short-duration sounds, L_{AE} has an advantage, because the onset and offset times of the measurement are not critical, provided that the whole event is included. However, for people who are not acoustical engineers, neither metric is as easy to understand as the simple L_{AFmax} or A-frequency and F-time weighted maximum sound level during an event.

The L_{Aeq} and L_{AE} metrics have become increasingly popular with administrators and regulators in recent years and have largely superseded a number of older metrics. The two main reasons for this, lower cost and greater convenience, have nothing to do with subjective response. The inclusion of software capable of calculating L_{Aeq} and L_{AE} within digital sound level meters results in instruments that are much more effective than earlier instruments based on analogue electronics and that are readily available at much lower cost than when these metrics were first introduced. L_{Aeq} and L_{AE} measurements are simple and robust, and facilitate straightforward calculations of the likely effects of noise-management action, which will often be entirely consistent with actual measurement. This kind of rational simplicity does not apply to many of the more esoteric noise metrics and indicators which were invented in the past and have now fallen out of use. On the other hand, rational simplicity and engineering convenience do not automatically imply consistency with subjective impressions. For example, according to the physics, doubling the number of similar noise events within a defined time period doubles the acoustic energy received which will increase L_{Aeq} by 3 dB. The same physical effect of a doubling of energy could be achieved by increasing the sound levels of the separate events by 3 dB while keeping the number of events constant. Typical human listeners are less likely to perceive 3 dB changes in the sound levels of events than doublings or halvings of the numbers of events, particularly if the change happens gradually over time.

The interpretation of metrics such as L_{Aeq} and L_{AE} requires caution, particularly when the effect of adding a new noise source into an existing acoustic environment is considered. If the existing acoustic environment and a new noise source both generate the same L_{Aeq} at the measurement position when measured separately, adding the new noise source will only increase the overall L_{Aeq} at the measurement position by 3 dB compared with what it was for the existing situation. Depending on the character of the new noise source compared with the existing acoustic environment, the addition could be highly noticeable in terms of subjective impressions, whereas an equivalent 3 dB increase in the L_{Aeq} sound level of the existing noise source measured separately might be completely unnoticeable. Assessors and regulators often mistakenly assume that a 3 dB change in

L_{Aeq} always has the same effect in terms of subjective impressions, regardless of what has actually happened to cause the change.

Complex time-varying acoustic environments can also be described by using what are known as *statistical levels*, L_n. Any particular L_n is the sound level (note that this would normally be A-weighted – as in L_{An}) over the defined measurement duration, which is exceeded for *n*% of the measurement duration. The L_{A10} and L_{A90} are the A-weighted sound levels which are exceeded for 10% and 90% of the measurement durations respectively. These statistical levels were widely used in the days before reliable and stable integrating-averaging sound level meters or systems became available, because they could be calculated by plotting out the statistical distributions of instantaneous sound levels obtained by using a mechanical statistical distribution analyser attached to a pen-recorder system. Modern digital sound level meter systems can perform the same calculations in software, but the whole concept of statistical levels has become largely obsolete, except for defined specialist applications.

The main problem with statistical levels is that they do not behave rationally when arithmetically manipulated. For example, to calculate the combined L_{50} from the addition of two noise sources of known L_{50}, it is necessary to know the cumulative distributions of the two time histories separately. The L_{A90} is still widely used in the United Kingdom and in one or two other European countries to indicate the steady background sound level attributable to mostly distant noise sources, but cannot be arithmetically manipulated in any rational way. The L_{A90} provides no information about how far above or below that sound level (the L_{A90}) that instantaneous sound levels rise or fall during the 90% of the time that sound levels exceed the L_{A90} or during the 10% of the time that they fall below the L_{A90}. On the other hand, the L_{A90} does have some advantages over other methods for indicating background sound levels, because it does not require any subjective decisions to be made about which noise sources are part of the background and which are part of the foreground. Almost any other metric used for this purpose requires the user to make what are essentially arbitrary subjective decisions about which noise sources are included in the background and which are not. The subjective importance of the steady background sound level is that it determines an objective threshold, above which more prominent noise sources can be administratively deemed to be intrusive and below which they can be deemed to be not intrusive. Note that this is an objective definition which might not necessarily be very representative of actual subjective impressions, but it can at least be applied consistently.

In the United Kingdom, the L_{A10} metric is still used (despite being technically obsolete) to determine entitlement to a noise-insulation grant in mitigation of increased road-traffic noise. The L_{A10} metric can be considered as behaving semi-rationally when applied to busy road-traffic noise, which has similar statistical properties anywhere in the country, but it does not behave rationally when applied to any intermittent noise source such as railway noise or aircraft noise, which might only be present for less than 10% of the time. Comparisons between different intermittent noise sources using statistical-level metrics are meaningless.

Other metrics have been proposed for intermittent noise sources (such as aircraft flyover events) such as the time-above x dB L_{pA} and the number-above y dB L_{pA}. Within a defined measurement duration, the time-above metric indicates the proportion of time during which a defined sound level x dB L_{pA} is exceeded, and the number-above metric indicates the number of events which exceed a defined sound level y dB L_{pA}. Both alternative metrics reflect different acoustic features from the *de facto* standard L_{Aeq}-based metrics and could therefore be useful in some situations. However, neither behaves rationally when arithmetically manipulated. The time-above and number-above metrics provide no information about what happens during the time periods that the x and y dB L_{pA} thresholds are exceeded. For example, the N_{65} metric (the number of events within a defined measurement period with maximum A-frequency and F-time weighted sound level, L_{AFmax}, exceeding 65 – assuming the A-frequency and F-time weightings are used) provides no indication of the number of decibels by which the qualifying events exceed the 65 dB L_{pA} threshold. Suppose there are 100 events per day at 66 L_{AFmax} (i.e. just exceeding 65 dB), the N_{65} would be 100. If the maximum sound level for each event was increased by 15 dB, the N_{65} would still be 100. The change in event sound levels from 66 L_{AFmax} to 81 L_{AFmax} would be highly noticeable, but the N_{65} metric would not have changed at all. If the maximum sound level for each event was decreased by only 2 dB, the N_{65} would fall to zero. The change in event sound levels from 66 to 64 L_{AFmax} would hardly be noticeable, but the N_{65} metric would have changed from 100 to 0.

6.3.5 Time-varying frequency spectra

The most comprehensive representations of complex acoustic environments can be provided by showing successive time-varying frequency spectra, but these are of limited use for assessment and regulation because of their complexity. No meaningful method has yet been devised for setting sound-level targets and limit values in terms of time-varying frequency spectra. The main problem is that there are no simple acoustic metrics which can always reflect all acoustic features present in any particular case which are, or could be, important for subjective response. In addition, anything too complex will be unworkable for assessment and regulation purposes. All objective acoustic metrics which have been adopted to date are compromise solutions which can only ever take a proportion of potentially relevant variables into account.

While time-varying frequency spectra are too complicated for use as sound-level targets and limit values, they can nevertheless be useful for identifying specific acoustic features (normally, either discrete tonal or harmonic components in the frequency domain or various fluctuations or modulations in the time domain) which appear to be making a disproportionate contribution to overall disturbance and annoyance, and should perhaps be targeted for engineering noise control effort. There are two main types of frequency-analyser system in general use for such purposes, depending on whether the frequency-analysis bandwidths are distributed linearly or logarithmically across the audio-frequency range. The fast-Fourier-transform (FFT) frequency analyser is the most common type and generates linearly spaced frequency-analysis bands (see Section 4.4.5). The resulting spectra can be displayed using a linear or a logarithmically spaced frequency scale irrespective of the underlying operating principle. A typical 1024 point FFT analyser running at a 25 kHz audio sample rate will display 400 linearly spaced frequency-resolution bandwidths from 0 to 10 kHz at 25 Hz spacing. Note that the actual frequency resolution is usually wider than the line spacing on the display, because this also depends on the averaging time and the type of time-domain window filter applied. At low frequencies (say 100 Hz) the frequency resolution of around 25 Hz is much coarser (25%) than the frequency resolution of the human ear. At high frequencies (say 10 kHz), the frequency resolution is the same at around 25 Hz and is much finer (0.25%) than the frequency resolution of the human ear. FFT analysers are good at detecting linearly spaced harmonics but are not well adapted to reflect human hearing.

The constant percentage bandwidth frequency analyser comprises a series of parallel one-third octave or narrower fractional octave band-pass filters (see Section 5.2.2.2). Traditional analogue equipment required separate filter circuits for each filter band, but modern digital systems can operate multiple parallel filters in real time without difficulty, and are consequently much better adapted to human hearing than FFT analysers. One-third-octave band-pass filters have been an industry standard for many years, and while they mimic the approximately logarithmically scaled frequency resolution of the human ear, the actual frequency resolution of one-third-octave band filters is still very much coarser than that of the human ear. This is one of the problems of the EPNdB metric used for aircraft-noise certification discussed in Section 6.3.3. The EPNdB metric uses one-third-octave band-pass filters (as do Zwicker's and Stevens' methods for calculating loudness level), possibly because that was the state of the art, using analogue electronics, at the time these metrics were originally developed. However, the normal human ear has about 20–25 times narrower frequency resolution than one-third octave band-pass filters, each of which span four semitones of a 12-semitone musical octave scale. Traditional one-third-octave band-pass filters would be useless for transcribing music.

Much narrower bandwidth parallel filters can be modelled in software-based systems that can more closely reflect the frequency-resolution capabilities of the human ear. For frequency analysis of rapidly time-varying sounds, the extended averaging times required for narrower-bandwidth frequency resolution have to be balanced against the time resolution required. One possible advantage of constant percentage bandwidth filters is that each filter can be set up with different averaging times to best suit requirements across the entire frequency range of interest. Low-frequency filters require much longer averaging times than high-frequency filters to obtain the same degree of statistical accuracy across the frequency range. FFT filters operate on discrete blocks of sample data, providing less flexibility for selection of the optimum averaging time in each frequency range, which is normally constrained to multiples of sample data blocks. Other more sophisticated methods of frequency analysis have been developed to solve some of the problems outlined in this section. They are outside the scope of this chapter, but they are discussed to some extent in Chapter 4.

6.3.6 Spatial hearing

The mechanisms underlying binaural hearing are discussed in Section 6.2.2, but the implications for spatial hearing require further consideration in this section. It is obvious that no sound level meter measurement system using a single precision-grade omnidirectional condenser microphone can reflect the spatial resolution capabilities of human hearing. The technical issue is whether or not it is possible to devise an acoustic metric that would reflect human spatial hearing capabilities in addition to trying to reflect the time- and frequency-resolution capabilities of human hearing. Binaural systems using microphones installed at the ear positions of a dummy head have some spatial capability, but are still not entirely satisfactory for use in metrics for the following reasons:

1. Human spatial hearing relies on the spectral and interaural differences associated with different directions of the source and how these differences change with head movement or movement of the source relative to the head. No instrumentation systems have yet been devised which can reflect these differences in ways that might be useable in standards and regulations.
2. The real-ear, head-related transfer functions for individual listeners vary depending on individual physiognomy. Dummy head systems could be standardised so as to represent an average physiognomy, or there could be three or more standardised systems designed to represent ranges of physiognomies. Any system proposed would not be able to overcome the problem of individual differences in physiognomy between different listeners.

On the other hand, it seems likely that the continuing technical development of spatial audio systems will lead to increasingly realistic auralisations. *Auralisation* is the technical term used to describe the process of generating realistic simulations of actual or hypothetical sound sources using computer synthesis techniques, and can be extremely useful for simulating hypothetical situations and developments in advance of actual manufacture or construction. Depending on the situation, however, subjective realism is not necessarily the same as objective realism, because it depends on human perception, which can be coloured by selective attention and expectation. At present, it seems unlikely that auralisation techniques can be developed to provide simple metrics suitable for objective assessment and regulation.

6.4 RANGE OF NOISE EFFECTS ON PEOPLE

6.4.1 Direct health effects

Very high sound levels can potentially cause localised heating or physical disruption of the human body. However, and depending on acoustic conditions, even an SPL of 120 dB has an acoustic intensity of only around 1 Wm^{-2}, which, even if all this physical energy was absorbed by a human body, would not be enough to cause any measurable damage. At the highest levels of environmental and community noise normally encountered outside people's houses (not normally more than 75–80 dB $L_{Aeq,16h}$ with occasional events up to 110–120 dB L_{AFmax}), there are no direct effects on the human body and no permanent effects on the physical environment, because the amount of physical energy is far too small. Most, if not all, of the observed effects of environmental and community noise at these and lower sound levels are caused by the way that the human body and the human brain react to or respond to noise as a sensory stimulus, and not because of any direct effects of the sound energy on the human body.

6.4.2 Hearing damage risk

Transient sounds with peak overpressures reaching a fraction of one atmosphere within a few milliseconds can cause direct physical disruption, particularly to sensitive tissues such as the eardrum and the lungs, but these sounds require extremely powerful sources such as gunfire, explosives or rocket motors, which are not encountered in normal community situations. The lower and higher peak sound-pressure action values of 135 dB (C-weighted) and 137 dB (C-weighted) specified in the *Control of Noise at Work Regulations* (2005) (referred to in Section 6.3.4) are intended to protect workers against the risk of hearing damage caused by repeated exposure to high-level transient sounds, and do not necessarily imply significant risk of physical damage associated with single or infrequent exposures at such levels.

For continuous sounds, and depending on duration, even short-time exposures at lower sound levels can lead to temporary loss of hearing sensitivity, known as TTS, and which may also be accompanied by tinnitus. Tinnitus, or 'ringing in the ears', is an audible sensation occurring when there is no actual sound present. One of the possible causes of TTS and associated tinnitus is fatigue and possible damage to the sensitive hair cells in the inner ear. Fatigue caused by overstimulation could cause some nerve fibres to lose sensitivity to actual input sounds and to start firing spontaneously even after the acoustic stimulation has stopped, or it could cause a diminution of function in other nerve fibres which would normally have an inhibitory role. After TTS, hearing sensitivity begins to return to normal immediately after the acoustic stimulation stops, although it may take several hours for full recovery to take place, depending on the degree of threshold shift in the first place.

If the exposure is sufficiently prolonged, this can lead to permanent loss of hearing sensitivity, which may also be accompanied by permanent tinnitus. Permanent hearing loss caused by excessive noise exposure is associated with progressive atrophy of the sensitive hair cells in the inner ear. The three most important factors in hearing damage risk are the sound level, the duration of exposure, and individual differences in susceptibility, although the duration and frequency of recovery periods between periods of high noise exposure may also be important. It has not yet been possible to devise reliable methods for identifying individuals with higher susceptibility to noise-induced hearing loss (NIHL) in advance of exposure. The biological mechanisms underlying progressive loss of hair cells are not precisely understood, but the measurable consequences of losing both sensitivity and frequency selectivity are easy to understand. Since it is unlikely that prolonged noise exposure at the sound levels known to be associated with NIHL would have been a significant feature of the environment before the development of industrialised societies, it also seems unlikely that individual differences in susceptibility would have been directly associated with any evolutionary mechanism, although they could be associated with unrelated genetic factors connected with other functions and systems.

The *Control of Noise at Work Regulations* (2005) require both employers and employees to carry out a range of preventive actions to reduce the risk of NIHL in the workplace, depending on whether the specified lower and higher action values (80 and 85 dB $L_{Ep,d}$: equivalent to $L_{Aeq,8\ h}$) are exceeded. There is a general duty to reduce noise exposure as far as is practical; employers are required to offer advice and preventive options at the lower action value and above; and both employers and employees have specific duties to prevent the higher action value from being exceeded. There is no similar mandatory protection against NIHL in domestic and leisure activities. This means that in, for example, a live-music venue or discotheque, management would be required to protect employees such as waiters and bar staff against excessive noise, perhaps by being required to wear ear protection, while there is no obligation to protect customers against the same risk. Generally speaking, customers are assumed to be at much lower risk of NIHL than employees because of infrequent attendance, but there is no obligation for venues to assess this. Attendance and resulting noise exposure for customers is voluntary, and hence, any associated risk of NIHL is, presumably, a matter for the customers themselves, although the venue does have at least a moral duty to inform them of risk. Similarly, while manufacturers of personal stereo systems have some obligation to warn users against the risk of NIHL associated with excessive use, any actual risk is outside the manufacturer's control because the duration of exposure is entirely a matter for the user. Actual risk is assumed to be independent of whether the exposure is incurred through a voluntary leisure activity or as an unavoidable condition of employment, but is nevertheless treated differently under existing laws.

It is generally accepted that the risk of NIHL reduces at lower cumulative sound levels. The current World Health Organization (WHO) guidelines for community noise (WHO 2000) accepted the advice given in the then current version of ISO (1999) that long term exposure up to 70 $L_{Aeq,24\ h}$ will not result in significant hearing impairment and quote this value in the summary tables of guideline values. Note that 70 dB $L_{Aeq,24\ h}$ is approximately equivalent to 75 dB $L_{Aeq,8\ h}$ for an 8 h working day if there is no other significant exposure during the rest of the 24 h day.

6.4.3 Causal mechanisms for other adverse health effects

Outdoor measured sound levels at community and residential locations are normally in the range from 45 to 75 dB $L_{Aeq,16\ h}$. In this range, acoustic stimulation causes a direct response in the auditory nerve and higher neural processing centres and this can, in turn, lead to endocrine and autonomic system responses such as temporary increases in blood pressure and concentrations of stress hormones such as adrenaline and cortisol. Normal autonomic and endocrine responses are essentially temporary adaptations by the body to meet changes in the external environment and are not necessarily harmful in themselves. A number of researchers, mainly from Europe, have suggested that temporary or acute adaptations could lead to more prolonged or chronic negative effects such as raised blood pressure, cardiovascular disease or mental health problems. There are important differences between temporary or acute effects, which might possibly be harmful on a short-term basis if sufficiently severe but do not persist after the noise has ceased, and long-term permanent effects, which continue after the noise exposure has ceased, such as NIHL. It is not presently known to what extent noise-induced temporary or transient changes in physiological variables such as blood pressure or stress hormone concentrations contribute to permanent changes which are harmful. It seems likely that individual susceptibility to long-term permanent effects varies. If only a minority of the population are in fact susceptible, it might be very difficult to demonstrate any such hypothesised causal links by statistical means, or, on the other hand, to demonstrate that they do not exist.

Difficulties arise because of the large number of situational, dietary, lifestyle and other variables which are known to be associated with adverse health effects. Some evidence exists of increased prevalence of circulatory disorders in noisier areas (Niemann and Maschke 2004; Babisch 2006), but causal relationships remain unproven because of the large numbers of uncontrolled possible confounding factors. For example, it is usually difficult to eliminate the possibility that people with higher-risk factors for certain diseases might also tend to prefer different types of residential areas for reasons other than noise, but which happen to be noisy to a sufficient extent that statistical associations can be observed. Alternatively, it is equally possible that people with higher-risk factors for noise-related diseases might tend to avoid noisier areas, thereby reducing any statistical correlations that might otherwise have been observed. In addition, in most cities, houses in the noisier areas are usually also exposed to higher levels of air pollution, which might provide an alternative explanation for any adverse health effects found. In general, most city noise is a consequence of higher population density in areas where large numbers of people choose, or prefer, to live.

The technical uncertainties in this area create problems for policymakers and administrators. Noise control is not usually without cost, which could exceed the benefits if they have been overestimated by applying overprecautionary assumptions about the health effects of noise. On the other hand, it is also possible that significant numbers of citizens are seriously affected and not enough has been done to reduce the risks. Policymakers have to decide these issues on the basis of incomplete evidence and are rarely qualified to be able to make these assessments. Acoustical engineers and physicists may be asked to make recommendations without necessarily being aware of the wider consequences of those recommendations if followed up. Engineering noise control carried beyond what is strictly necessary to minimise disturbance and possible health risks could have economic consequences which could be more damaging than the noise effects to which it was addressed.

6.4.4 Speech interference

Speech communication is one of the most important characteristics of the human species. The evolutionary development of speech processing within the auditory system appears to be intimately linked to the development of the vocal tract and associated neural processing involved in producing speech. Two main methods exist for estimating speech intelligibility based on objective measurements of interfering background noise; the *speech intelligibility index* (SII) (ANSI 1997) and the related *speech transmission index* (STI) (BSI 2003a).

The classic articulation-index (AI) (ANSI 1969) calculation procedure, on which the current US SII calculation procedure is based, compared the long-time r.m.s. average one-third octave-band frequency spectrum of the wanted speech signal against the long-time r.m.s. average one-third octave-band frequency spectrum of the masking noise. The measured signal-to-masking noise ratio in each of the 14 one-third octave bands from 200 Hz to 5 kHz was assumed to contribute to overall intelligibility, according to weightings defined on the basis of laboratory listening tests. The standard weightings assumed that the 200 Hz frequency band contributed the least to overall intelligibility and that the 1.6 and 2 kHz frequency bands contributed the most. Note that the standard weightings were derived using native English-speaking talkers and listeners and might not be optimum for different languages or for talkers with unusually high- or low-pitched voices, or for electronically processed speech. The frequency bands and weightings were modified slightly for the current SII procedure.

The measurement of the speech signal-to-masking noise ratio has to take into account that the instantaneous sound level of any typical speech signal varies significantly over time. Obviously, the relative peaks in the speech signal contribute much more to overall intelligibility than the periods of relative silence in between the speech sounds. Similarly, and depending on their duration, any relative peaks in the masking noise contribute more to the masking effect than any quieter periods in between. The louder speech phonemes will be more easily distinguished during quieter periods in the background noise and vice versa. The maximum sound levels of consecutive peaks in the speech signal vary over a wide range, according to the phoneme being spoken and according to the talker. Standardised *peak programme meters* (PPM) have been used for monitoring speech sounds for specialist radio broadcasting and recording purposes, but the PPM standard has not been widely applied elsewhere. Most recording and broadcast monitoring is carried out using simple volume-unit short-time r.m.s. averaging meters, which operate similarly to conventional sound level meters used in the environmental and community noise-management fields. Long-time r.m.s. averages are used in the AI/SII measurement procedures to quantify both the speech-signal and masking-noise sound levels, assuming that instantaneous speech peaks rise 12 dB above the long-time r.m.s. average. In practice, this is a reasonable assumption for normal running speech, but might not apply to shouted or whispered speech, or for talkers who speak much faster or much slower than normal, and definitely does not apply where the dynamics of running speech have been compressed either by intentional audio processing or by unintended nonlinearities and distortions in an electro-acoustic speech communications system.

The unweighted contribution made by each one-third octave bandwidth of the speech signal to overall intelligibility (before multiplication by the standard weightings described) were assumed to vary from zero at speech signal-to-masking noise long-time r.m.s. ratio of –12 dB up to a maximum of unity at speech signal-to-masking noise long-time r.m.s. ratio of +18 dB. According to the assumption that the peaks rise 12 dB above the long-time r.m.s. average, this is equivalent to assuming a zero unweighted contribution at 0 dB speech peak-to-masking noise long-term r.m.s. ratio, increasing to a unity unweighted contribution at +30 dB speech peak-to-masking noise long-term r.m.s. ratio.

Guidance included with the original AI procedure assumes that an AI score of 0 represents zero intelligibility, that an AI score of 0.5 is acceptable for general purposes and that an AI score of 0.7 or above is desirable for electro-acoustic speech communications systems 'to be used under a variety of stress conditions and by a large number of different talkers and listeners with varying degrees of skill'. Clearly, the AI/SII calculation procedure can only deal with the masking effects of interfering steady background noise on undistorted normal running speech. It requires that the speech signal and the masking noise be measured separately, which would, of course, be rather difficult in a sports stadium or similar indoor situations where most of the masking effect is contributed by crowd noise or reverberation excited by the speech signal. The AI calculation procedure is of limited value in many practical situations because it does not address distortion of the wanted speech signals that might occur in electro-acoustic speech communications systems (such as telephone systems) other than reductions of speech signal-to-masking noise ratio caused by uncorrelated masking noise. It effectively assumes that the speech signal would be 100% intelligible if there were no masking noise. This is not always realistic. In addition, speech communications systems with no audible distortion and no background noise can still be ineffective if the talker does not speak clearly, breathes heavily into the microphone, or speaks in a language which is unknown to the listeners, when the actual intelligibility could then drop to zero.

The STI was invented by Houtgast and Steeneken at the Netherlands Organisation for Applied Scientific Research (TNO) (Houtgast et al. 1980) and is similar to the AI and SII but uses a special test signal to allow measurements to take place over a much wider range of circumstances. The STI test signal (and the simplified rapid speech transmission index (RASTI) test signal) is made up from cosine amplitude-modulated narrowband noise signals with similar dynamics and frequency content to normal running speech, but which allows speech signal-to-noise ratios to be inferred from measured modulation transfer functions in each frequency band. The STI measurement system is an improvement to the AI measurement and calculation procedure, because it does not require the speech signal and the masking noise to be measured separately, which is impossible where the masking noise is mainly reverberation. Modulation transfer functions can also be calculated from measured or calculated system impulse-response functions, which can be very useful in system design where the overall acoustic performance can be predicted in advance using image-source and ray-tracing models.

Note that the median AI or STI score of 0.5 can be obtained where the long-time r.m.s. average speech signal-to-masking noise ratio is only +3 dB across the one-third octave frequency bands from 200 Hz to 5 kHz. Experimental data reported in the original ANSI standard (ANSI 1969) shows that 95% sentence intelligibility (measured using real listeners) can be achieved where the AI or STI score is only 0.5. Sentence intelligibility is always higher than the intelligibility of separate words, which in turn is always higher than the intelligibility of random nonsense syllables. This is because of the redundancy and additional context information in meaningful words and sentences as compared with nonsense syllables. If a speech system is to be used where the intelligibility of unfamiliar words is a priority, a higher AI or STI score than 0.5 will be required.

Actual intelligibility varies above and below intelligibility estimated from objective measurements, reflecting individual differences in talking and listening skills, and also differences in the level of motivation among listeners. Selective attention and other higher-order perceptual processes affect individual test results. All objective speech intelligibility calculation and measurement procedures apply to ideal or standardised test conditions, which may not be representative of actual conditions.

Further uncertainties arise because real speech is far more complex than the one-third octave frequency bandwidths represented by the AI procedure or by the cosine amplitude-modulated one-third octave bands of random noise included in the STI procedure. Real speech comprises a sequence of separate phonemes, each of which is made up from a series of harmonically related voiced frequencies generated by the vocal cords, together with mostly higher-frequency broadband components generated by turbulent airflows through constrictions in the vocal tract. The overall speech frequency spectrum is then filtered through variable formant frequency bands generated by changing the articulation of the vocal tract. Transitions and silences between consecutive phonemes can be important for meaning. One-third octave-band spectrum methods such as AI and STI ignore the periodic and harmonic nature of typical vowel sounds, which contribute a large part of the intelligibility of normal running speech. The fundamental and harmonic frequency components in vowel sounds are controlled by the airflow through, and muscular tension in, the vocal cords, and can be heard as voice pitch, as in singing. Changes in voice pitch contribute information about the emotional state of the talker, in addition to the semantic content contributed by word and sentence intelligibility.

To summarise, objective procedures such as AI/SII and STI can be more useful for comparing the relative intelligibility capabilities between generically similar speech communications systems than for absolute assessments of overall performance where any significant degree of precision is required. The only way to determine actual speech intelligibility is to carry out behavioural measurements using real talkers and real listeners, and even then the results will vary depending on the actual talkers, listeners, word lists or sentence materials and test procedures used. Real speech using running sentences has considerable redundancy, which allows listeners to achieve relatively high intelligibility even where there are significant masking or distortion effects.

6.4.5 Activity disturbance and personal preference

It is well known that excessive noise can interfere with or disturb wanted activities such as rest and relaxation, private study or concentration on important tasks such as desk work. However, the amount of interference or disturbance can also be highly dependent on the sensitivities and sensibilities of the individuals concerned. Some people can ignore distracting noise to a much greater extent than others. Motivation can be strongly influenced by penalties and rewards. Because there are so many other factors involved, there are no reliable methods for predicting activity disturbance based on objective sound levels alone, except for specific listening tasks which depend on good speech intelligibility, in which case AI-type indicators may suffice.

In recent years, there has been a considerable amount of interest in the adverse effects of noise in the classroom. Excessive noise break-in from external sources such as aircraft, road traffic and industrial or commercial sources can interfere with speech and music in the classroom, and various authorities have been concerned about the extent to which this might affect academic progress. The recent European road-traffic and aircraft-noise exposure and children's cognition and health (RANCH) study (Stansfeld et al. 2005) found statistically significant associations between aircraft noise measured outside schools and reading comprehension and recognition memory, based on pooled analysis from survey data collected in three European countries. Higher external-noise sound levels (mainly road traffic) have also been associated with poorer performance on standardised assessment scores in primary-school children in London (Shield and Dockrell 2008). However, academic progress can also be affected by many other factors, and it is not yet clear to what extent these findings actually demonstrate cause and effect. It is difficult or impossible to control for all possible confounding variables such as differences in socioeconomic class and language spoken at home between different schools in different areas, or for exposure to other noise sources when not at school.

A number of different mechanisms have been proposed which could potentially explain the observed effects of external noise break-in on child development and attainment. External noise break-in can directly interfere with speech communications, and while this kind of disruption can be overcome by increased vocal effort or repeating sentences, it could nevertheless interfere with concentration and attention or reduce motivation. Complex tasks requiring high cognitive demands, as opposed to simple repetitive tasks, tend to be the most affected by environmental stressors such as noise, and it is possible that some children could be particularly sensitive to noise interference at key development stages. On the other hand, it is not known to what extent delayed development at key development stages can be recovered subsequently. Different children develop at different rates. The current WHO guidelines (WHO 2000) suggests that children and elderly people should be treated as 'vulnerable' groups, according to the assumption that both groups are more susceptible to adverse effects of noise, although the evidence to support this suggestion seems to be lacking.

In more general terms, people differ in their preferences for either complete silence or some form of masking noise or music which might prevent otherwise distracting noise from being heard. Depending on sound quality, a certain amount of noise can contribute to arousal and help to maintain alertness, whereas too much noise can be distracting. Feeling annoyed by the presence of what is assumed to be unnecessary or avoidable noise is unlikely to assist concentration, particularly on difficult tasks. The provision of artificial background noise in open-plan offices provides an interesting example. Depending on the type of work being done and the motivation of the people doing the work, introducing controlled background noise in open-plan offices at an appropriate level can increase performance. The secret is to provide just enough controlled background noise to mask distracting intrusions and maintain speech privacy without the controlled background noise becoming objectionable in itself. Most people have heard background music in public places such as railway stations and shopping malls. It is only ever provided because management perceive it to be beneficial in terms of sales or efficiency. The provision of military band-style music during the morning rush hour at railway stations can assist in encouraging people to disperse rapidly, while more restful music at less busy periods or in airport departure lounges can help to calm people who might otherwise find the experience stressful. Different types of music at the supermarket checkout can have significant effects on customer behaviour, but the effects are variable depending on individual personality and preference. If people were all the same, then they would all like to listen to the same things, which is clearly not the case.

The fashion industry is particularly sensitive to the type of music reproduced over public-address systems inside shops and shopping malls, and controlled experiments in these areas are rarely reported because of commercial sensitivity. Service companies provide online recorded music programmes tailored to specific markets. In addition to music-on-demand services for domestic consumers, they also provide specific programmes for different types of retail outlets based on market research. Retail management buys into these services because of the effect they have on sales. The particular types of music programme reproduced in different stores are designed to appeal to the type of customers who are most likely to purchase the type of products sold in that store. Management can immediately tell if the selected music is commercially beneficial or harmful because of the effect on sales. It is, of course, impossible to predict the likely effects of different types of music on sales turnover just by measuring the L_{Aeq} sound level. Personal preference is not an engineering quantity, and there is hardly any relationship between observed effects and measured sound levels in this area.

6.4.6 Sleep disturbance

Sleep disturbance is interesting, because everyone knows that noise can disturb sleep, but nobody really understands what sleep is actually for. The most important effect of loss of sleep from any cause seems to be increased sleepiness the next day. Anything that increases the risk of falling asleep while engaged in safety-critical tasks such as long-distance lorry driving is obviously dangerous. However, if the person has not actually fallen asleep, task performance also depends on application and motivation and can be relatively unaffected. Anything that interferes with normal sleep patterns can be very annoying, and may cause people to believe that their performance or mood the next day has been adversely affected, even if there is no objective evidence of any adverse effect. Disturbed sleep during one night can be compensated for by more time spent in deep sleep on following nights, and most people habituate to familiar sounds such as regular railway-train pass-bys to some considerable extent. Most people are more sensitive to noise while sleeping in an experimental laboratory situation than in their own homes, especially on the first night.

Most people sleep for around 7–8 h every night, although there is considerable variation above and below the average. Normal sleep passes through successive stages from light sleep to deep sleep and back to light sleep again in approximately 90 min cycles. People are least sensitive to external stimuli when in deep sleep, which can be identified from different brainwave patterns recorded via electroencephalogram (EEG) electrodes placed on the scalp. When a person is awake, EEG records tend to be random and irregular. EEG records become more regular with increasingly identifiable low-frequency wave patterns with increasing depth of either relaxation or sleep. There is normally a short awakening period between each sleep cycle which is not usually remembered the next day unless something interesting was happening at the time. Short periods of REM (rapid eye movement) sleep associated with dreaming are also most likely to occur in between the separate sleep cycles. The proportion of time spent in deep (slow-wave) sleep decreases with successive sleep cycles, while the proportion of time spent in REM (dreaming) sleep increases. Loss of sleep on previous nights is generally compensated for by an increased proportion of time spent in deep sleep on subsequent nights.

The peripheral auditory system remains active during sleep, but higher-level cognitive processing systems shut down unless an incoming signal is sufficiently intense or unexpected to require cognitive attention. Sensitivity varies depending on sleep stage and the number of sleep cycles which have already been completed at that time. Neural response to transient sound can lead to partial or complete arousal depending on the sound level and character of the sound. Transient arousals identified from changes in EEG records, other physiological measurements (heart rate, blood pressure, endocrine responses, etc.), changes in posture or other movements do not necessarily lead to behavioural awakening, although they may increase the probability that another sound following soon after causes behavioural awakening.

Habituation to familiar or expected sounds reduces the probability of behavioural awakening without necessarily reducing the prevalence of transient arousals. Scientific opinion is divided about the possible effects of transient arousals which do not progress to behavioural awakening. Transient arousals could be considered simply as indicators of normal biological function.

Babisch (2006) observed an increased risk of cardiovascular disease in people living in areas with high road-traffic noise sound levels, and speculated that the increased risk could have been associated with an increased prevalence of transient arousal while asleep. Repeated transient arousal leads to increased blood concentrations of cortisol and adrenaline hormones, the so-called stress hormones. The main function of these hormones is to facilitate action in response to an opportunity or threat. If there is no action because the person is asleep and has not actually been awakened, the stress hormones might still (in theory) have some effect on metabolic homeostasis (European Environment Agency 2010) and thereby contribute to an increased risk of cardiovascular disease over the longer term.

The current WHO guideline values (WHO 2000) for sleep disturbance at 30 dB L_{Aeq} and 45 dB L_{ASmax} are mostly based on laboratory data, which consistently shows transient arousals associated with noise events (and not necessarily leading to behavioural awakening) at quite modest sound levels down to around 40–50 dB L_{ASmax} measured indoors. The most recent

WHO Europe night noise guidelines document (WHO 2009) proposes even lower guideline values, based on recent field data showing small increases in the probability of transient awakening associated with aircraft-noise events (measured using EEG records) down to around 35 dB L_{ASmax} measured indoors. However, the most recent large-scale field study of aircraft noise and sleep (Jones and Rhodes 2013) carried out in residential areas around Heathrow, Gatwick, Manchester and Stansted airports in the United Kingdom found that disturbance measured by using wrist-worn actimeters was minimal where outdoor aircraft-event sound levels were below 80 dB L_{ASmax}. The probability of the average person being awakened at higher aircraft-event sound levels up to 95 dB L_{ASmax} measured outdoors was only about 1 in 75. A subsequent small-scale field and laboratory combined-methodology trial study (Robertson et al. 1999) using EEG recordings found similar results. Taking into account the typical sound-level difference between outdoors and indoors for average British houses of around 15 dB with the bedroom windows open for ventilation and around 25 dB with the bedroom windows closed, the current WHO guideline values appear to be somewhat conservative.

Which data should be relied on for assessment and regulation purposes? There is no clear-cut answer to this question. In the United Kingdom, Department of Transport policy on night-time aircraft noise continues to be informed by the Ollerhead et al. (1992) study. Airport noise objectors would probably prefer policy to be based on the WHO guidelines. One of the main difficulties here is that most kinds of transient arousals and other sleep disturbances associated with noise events also occur naturally throughout the night anyway, even if there is no noise at all. Even transient arousals or other sleep disturbances occurring at the same time as a noise event may have been about to happen anyway. A person who claims to have been woken up by an early-morning aircraft-noise event cannot be certain that it was actually the aircraft-noise event complained of that caused them to be woken. On the other hand, if there is a statistically significant increase in the overall amount of sleep disturbance in noisy areas compared with quiet areas, then this would be good evidence in support of an effect.

Policymakers tend to be influenced by public perceptions, individual complaints and community action unless there is scientific evidence to persuade them otherwise. The available evidence suggests that, while most people habituate to noise at night to a much greater extent than claimed by community action groups, there may also be a small minority who are persistently disturbed and may even be suffering from long-term health consequences as a result. Not surprisingly, not many people are very good at reporting what happened while they were asleep. It is very easy to misidentify the actual cause of any wakening event, particularly if the offending road vehicle or aircraft has moved on by the time the sleeper is sufficiently awake to try to work out what happened.

How can we best summarise what is known about noise and sleep disturbance? Sounds which cannot actually be heard because of masking by other sounds cannot disturb sleep, although it is possible that some other sensory input, such as car headlights shining into the bedroom, could be the actual cause of disturbance which might later be wrongly attributed to the noise made by the car. However, any sound which is audible to a person who is awake can potentially disturb sleep to an extent determined by the character of the sound, its relative prominence and onset rate compared with the background-noise sound level, and the unexpectedness or familiarity of the sound. Any audible sound can contribute to an observable transient arousal. Transient arousals may or may not lead to behavioural awakening, depending on the changing sensitivity of the person during different stages of the sleep cycle and the extent to which the sound can be recognised by subconscious processing as being of no interest without requiring cognitive appraisal, which can only be performed if the person is awake. Because of the number of variables involved it is almost impossible to predict whether any particular sound will actually wake a person up or not, except if the sound is very loud or particularly strident. This is the principle adopted for hotel fire alarms, which are required, according to a completely different concept from the WHO night noise guidelines, to have a minimum sound level of 75 dB L_{AFmax}, measured at the bedhead (BSI 2013), to have a reasonable probability of waking hotel guests.

6.4.7 Annoyance

By far the greatest amount of effort and resource expended on noise control and noise management is directed towards the reduction of noise annoyance. Somewhat curiously, noise annoyance is also the least well defined out of all possible adverse effects of noise. This is probably because there are no direct physical or physiological correlates of noise annoyance, so there are no possibilities of absolute calibration. As a subjective concept, noise annoyance is fundamentally intangible. Most people have a generic understanding of the concept, but there are large individual differences between different people in different situations. Actual annoyance (whatever that is) probably varies, and so also does different people's understanding of what the word really means. This is one of the main reasons why noise assessment and regulation is usually based on sound levels measured or calculated according to objective procedures rather than on subjective opinions, which can always be challenged on the basis of alleged or actual individual bias. Noise makers can usually estimate the engineering and financial costs and inconvenience of noise control and noise management action, whereas noise sufferers cannot quantify the reduction in annoyance likely to be achieved by noise control and noise management action in any comparable way.

The prevalence of noise annoyance can be measured by the numbers of noise complaints in any particular area. For example, most major airports record and tabulate noise complaints which are often about specific noise events, but can also be more general in nature or about other topics such as track keeping and air pollution. Not all complaints can be related back to the specific events complained of; for example, air traffic control radar records might show there were no aircraft flying in a particular area at the time of the complaint. Any sudden increase or decrease in the overall numbers of complaints could be associated with some actual change in operations, or alternatively with a change in awareness resulting from proposed developments or media attention. Complaint statistics can provide an indication of possible changes in the general attitudes of the population towards the organisation or facility receiving the complaints. However, the attitudes and opinions of the majority of the population who do not register complaints are often more interesting and could be as, or more, relevant for policy. People who do not complain may have nothing to complain about; or they could be annoyed but do not know how to complain or be concerned about possible victimisation, or they may simply feel it would be a waste of time. Sometimes people may be annoyed but are sufficiently reassured by the knowledge or belief that the noise maker has treated them fairly and is prepared to take complaints seriously that they are happy to accept the situation without actually complaining. This is the concept of perceived fairness, whereby people can be more inclined to accept a limited amount of disturbance and annoyance if they consider that preventive action has been commensurate and reasonable and that any residual disturbance is unavoidable.

Reported annoyance measured using standardised questionnaires administered to statistically representative samples of the population is assumed to be representative of underlying actual annoyance, whatever that might mean to the individuals concerned. Reported annoyance is a record of a person's objectively observable behaviour when responding to or completing a questionnaire. In cross-sectional studies, the effects of subjective bias can be controlled, at least to some extent, by averaging reported annoyance across statistically representative samples of survey respondents exposed to noise in different and objectively measurable ways, and by standardising the procedures according to the specification set out by an ISO committee established for that purpose.

The relevant ISO recommendation (ISO 2003) defines noise annoyance as 'a person's individual adverse reaction to noise'. The definition in the standard is amplified by two notes as follows: 'NOTE 1 The reaction may be referred to in various ways, including for example dissatisfaction, bother, annoyance, and disturbance due to noise; NOTE 2 Community annoyance is the prevalence rate of this individual reaction in a community, as measured by the responses to questions specified in clause 5 and expressed in appropriate statistical terms.'

The standard specifies two alternative direct questions which should be presented towards the beginning of any questionnaire to avoid bias from questions on other or more detailed topics which could then be presented later.

Verbal scale

Thinking about the last (… 12 months or so …) when you are here at home, how much does noise from (… noise source …) bother, disturb, or annoy you:

Not at all Slightly Moderately Very Extremely (annoyed)

Numerical scale

Next is a zero to ten opinion scale for how much (… source …) noise bothers, disturbs or annoys you when you are here at home. If you are not annoyed at all choose zero, if you are extremely annoyed choose ten, if you are somewhere in between, choose a number between zero and ten.

Thinking about the last (… 12 months or so …) what number from zero to ten best shows how much you are bothered, disturbed, or annoyed by (… source …) noise:

Not at all annoyed 0 1 2 3 4 5 6 7 8 9 10 Extremely annoyed

The ISO committee was unable to agree on a single recommendation for either the verbal or the numeric scale and therefore offered both. The adjectives on the 5-point verbal scale were selected as being the most evenly spaced from 'not at all' to 'extremely', based on psycholinguistic testing. The main criticism of the 5-point verbal scale was that it does not allow finer discriminations between the scale points, so the 11-point numeric scale was offered as an alternative. However, statistical comparisons of the 11-point numeric scale against the 5-point numeric scale showed that the extra scale points in the numeric scale merely increase the variance without having much effect on statistical precision. The 11-point numeric scale is most useful for laboratory-type tests where listeners are asked to rate a series of different sounds, one after the other, in repeated measures experimental designs. For single one-off judgements of the sound outside a respondent's house, the 5-point verbal scale seems to be perfectly adequate and is generally easier to explain.

The ISO standard annoyance scale avoids the problem of not being able to define annoyance in any precise way by instead specifying the way in which it should be measured. Changing the wording of the questionnaire or changing the way in which the questionnaire is administered could change the results, but this uncertainty is avoided if the standard format for the questionnaire is always used in the same way. The ISO standard encourages further questionnaire items to investigate any specific issues in more depth, but only after the standard annoyance scale has been dealt with first to avoid bias.

The WHO defines health as ‘a state of complete physical, mental and social well-being and not merely an absence of disease or infirmity’. Most people in the United Kingdom would not normally consider noise annoyance to be a health effect as such, but reported annoyance reflects (or is assumed to reflect) a degree of adverse reaction in terms of perceived ‘physical, mental, and social well-being’ as defined in the WHO constitution. This means that, according to the WHO, reported annoyance measured using the ISO standard annoyance scale can be considered as a legitimate health effect. The seemingly unlikely but nevertheless theoretical possibility that nonhabituated physiological response to noise while asleep or otherwise unaware could lead to potentially harmful long-term consequences was discussed in Section 6.4.6. A more plausible consideration is that high concentrations of stress hormones could arise simply from a person’s continuing frustration and perceived lack of control to be able to do anything about their situation – and that these high concentrations could lead to more serious adverse health consequences over the longer term. For any individuals for whom this is the case, any measure which reduces annoyance would also reduce associated stress and any subsequent adverse health consequences. For the reduction of any direct effects, any action which reduces physical exposure could be effective. However, for the reduction of annoyance/stress-related effects, psychologically-based methods which do not necessarily involve any actual reduction in physical sound levels could nevertheless be equally or more effective.

What does noise annoyance really mean? In open-ended qualitative interviews, many respondents in areas where a majority report relatively high levels of aircraft-noise annoyance in standardised questionnaire surveys admit to having ‘mostly got used to’ the noise. The most straightforward explanation for this apparent inconsistency is that when reporting high annoyance in a standardised questionnaire, the respondent is expressing an opinion about the local noise environment, which, even if that respondent has personally adapted to it to at least some extent, does not necessarily mean that the reported annoyance should be ignored by policymakers. Most residents around airports (or near main roads or other noise sources) appreciate a range of both advantages and disadvantages of continuing to live where they do, and noise is usually only one of several disadvantages. It is perfectly reasonable for individual residents to describe the noise as very or extremely annoying (perhaps on a Sunday afternoon when they have friends round for an outdoor barbecue) but nevertheless acceptable within the overall balance of advantages and disadvantages of the place where they live. In the past, some research studies have attempted to conceal these inconsistencies by not disclosing the true purpose of the questionnaire until the last possible moment. In this way, they attempted to discover if respondents would mention aircraft noise as a disadvantage of living where they do without having been prompted by being told the true purpose of the research. Quite apart from the ethical issue of requiring informed consent before taking part in any research of this type, it is not obvious that this approach would have acquired data of any more use for informing policy. If people report that they are annoyed or even very or extremely annoyed by aircraft noise but not (separately) to the extent that it is unacceptable in terms of their everyday lives, then that should be of interest for policy.

Most surveys have shown a general tendency for average reported annoyance to be higher in higher-noise areas, and researchers have attempted to combine the various results to obtain aggregated and possibly more generally applicable noise dose–annoyance response relationships (Fidell 2003; European Commission 2000). In the past, researchers generally preferred to devise their own questionnaire wording and experimental designs, either to test particular hypotheses or simply because they did not particularly agree with other researchers' designs. The resulting diversity created particular difficulties for subsequent researchers attempting to combine different databases in the expectation of finding underlying trends not observable in individual studies because of sample size limitations. Even if exactly the same questionnaire wording has been used in different studies, this does not necessarily mean that the questionnaire responses are directly comparable, because there could have been other material differences in design and procedure. If different questionnaire wordings were used, the data becomes even less comparable. Meta-analysis of combined data sets is only feasible by making various normalising assumptions to enforce comparability of the data. Unfortunately, the only way to devise appropriate normalising assumptions is to assume underlying consistency between the different data sets. The normalising assumptions then become self-justifying. It is equally plausible that there is no *a priori* reason to expect consistency between different surveys, and that observed differences in noise dose–annoyance response relationships between different surveys represent normal and understandable diversity (Fidell et al. 2011).

In summary, it seems that the main causal factors responsible for reported noise annoyance can be divided into three main categories, in order of importance as follows:

1. The sounds themselves (this is *not* necessarily the same thing as a sound level attributed to the sound)
2. The psychological orientation of the listener towards the sound within the context in which it is heard
3. A large and essentially unpredictable or random component

Administrators and regulators may be able to devise and assess policy which can reduce the noisiness or potential to cause annoyance of objectionable or complained-of sounds, either by reducing sound levels measured according to established acoustic metrics, or by improving the sound character or sound quality. Whichever method is the most cost-effective should be used, but it should be noted that cost-effectiveness is difficult to assess in this area because of a lack of directly comparable metrics.

Alternatively, administrators and regulators may be able to devise and assess policy which addresses the psychological orientation of listeners towards objectionable or complained-of sounds. Methods which address psychological orientation can be as or more effective in reducing reported annoyance as methods addressed to physical or objectively measurable noise reduction. However, it seems unlikely that methods based solely on information exchange and good public relations would have any long-term effect unless they are also accompanied by physical measures that respondents can actually see and hear.

It is unlikely that administrators and regulators will ever be able to do anything about the unpredictable or random components of reported annoyance. There are many types of sound which some people enjoy and others hate, and this is often completely unpredictable. Exactly the same sound can be enjoyable one day and annoying the next, depending on whatever else is going on at the time. There would seem to be little or no point in trying to predict the unpredictable.

REFERENCES

ANSI (American National Standards Institute) (1969) S3.5 R 1986 *American National Standard Methods for the Calculation of the Articulation Index*, American National Standards Institute, New York.

ANSI (American National Standards Institute) (1997) S3.5 *American National Standard Methods for the Calculation of the Speech Intelligibility Index*, American National Standards Institute, New York.

Babisch, W. (2006) *Transportation Noise and Cardiovascular Risk*, Umweltbundesamt, Berlin.

BSI (British Standards Institution) (1979) *BS 5727: Method for Describing Aircraft Noise Heard on the Ground*, British Standards Institution, London.

BSI (British Standards Institution) (1999) *BS 8233: Sound Insulation and Noise Reduction for Buildings. Code of Practice: Annex B Noise Rating*, British Standards Institution, London.

BSI (British Standards Institution) (2003a) *BS EN 60268-16: Sound System Equipment Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index*, British Standards Institution, London.

BSI (British Standards Institution) (2003b) *BS ISO 226: Acoustics: Normal Equal Loudness Level Contours,* British Standards Institution, London.

BSI (British Standards Institution) (2013) *BS 5839-1: Fire Detection and Fire Alarm Systems for Buildings*, British Standards Institution, London.

Control of Noise at Work Regulations (2005) SI 2005/1643, Statutory Instruments, The Stationery Office, Norwich.

European Commission (2000) Position paper on dose-response relationships between transportation noise and annoyance. COM(2000) 468 final, Office for Official Publications of the European Communities, Luxembourg.

European Environment Agency (2010) Good practice guide on noise exposure and potential health effects, EEA Technical Report No. 11, European Environment Agency, Copenhagen.

European Parliament/Council of the European Union (2002) Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002 relating to the assessment and management of environmental noise, Office for Official Publications of the European Communities, Luxembourg.

Fidell, S. (2003) The Schultz curve 25 years later: A research perspective, *Journal of the Acoustical Society of America*, **114**(6) Part 1, 3007–3015.

Fidell, S. et al. (2011) A first-principles model for estimating the prevalence of annoyance with aircraft noise exposure, *Journal of the Acoustical Society of America*, **130**(2), 791–806.

Houtgast, T., Steeneken, H.J.M. and Plomp, R. (1980) Predicting speech intelligibility in rooms from the Modulation Transfer Function. I. General room acoustics, *Acta Acustica United with Acustica*, **46**(1), 60–72.

ISO (International Organization for Standardization) (1975) *ISO 532: Acoustics – Method for Calculating Loudness Level*, International Organization for Standardization, Geneva.

ISO (International Organization for Standardization) (2003) *ISO/TS 15666: Acoustics – Assessment of Noise Annoyance by Means of Social and Socio-Acoustic Surveys*, International Organization for Standardization, Geneva.

Jones, K. and Rhodes, D.P. (2013) Aircraft noise, sleep disturbance and health effects: A review, ERCD REPORT 1208, Civil Aviation Authority, The Stationery Office, Norwich.

Morfey, C.L. (2001) *Dictionary of Acoustics*, Academic Press, London.

Niemann, H. and Maschke, C. (2004) WHO large analysis and review of European housing and health status (LARES): Final report noise effects and morbidity, WHO Europe office, Berlin.

Ollerhead, J.B., Jones, C.J., Cadoux, R.E., Woodley, A., Atkinson, B.J., Horne, J.A., Pankhurst, F.L., et al. (1992) Report of a field study of aircraft noise and sleep disturbance, Department of Transport, London.

Robertson, K.A., Wright, N., Turner, C., Birch, C., Flindell, I., Bullmore, A., Berry, B., Jiggins, M., Davison, M., Dix, M. (1999) Aircraft noise and sleep 1999 UK trial methodology study, Report 6131 R01, ISVR Consultancy Services.

Shield, B.M. and Dockrell, J.E. (2008) The effects of environmental and classroom noise on the academic attainments of primary school children, *Journal of the Acoustical Society of America*, **123**(1), 133.

Stansfeld, S.A., Berglund, B., Clark, C., Lopez-Barrio, I., Fischer, P., Öhrström, E., Haines, M.M. et al. (2005) On behalf of the RANCH study team, Aircraft and road traffic noise and children's cognition and health: A cross-national study, *The Lancet*, **365**, 1942–1949.
WHO (World Health Organization) (2000) Guidelines for community noise, WHO, Geneva.
WHO (World Health Organization) (2009) Night noise guideline for Europe, WHO Regional Office for Europe, Geneva.