1991

"An Investigation into the Perceived Urgency of Auditory Warnings

Hellier, Elizabeth Jane

http://hdl.handle.net/10026.1/1746

http://dx.doi.org/10.24382/4162
University of Plymouth

"An investigation into the Perceived Urgency of Auditory Warnings."

Elizabeth Jane Hellier.

Submitted to the Council for National Academic Awards in Partial Fulfilment for the Degree of Doctor of Philosophy.

Sponsoring Establishment :
Polytechnic South West
Department of Psychology

Collaborating Establishment:
RAE Farnborough

September 1991

Acknowledgments.

I would like to thank Dr Judy Edworthy and Dr Ian Dennis, my supervisors, for their help and encouragement throughout the last three years, and for their prompt and constructive criticism of earlier versions of this manuscript. My thanks are also due to all those with whom I have had discussion on the subject matter of this thesis. In particular, Sarah Loxley, Jonathon Elcock, Andrew Lee, Tina Meredith, Prof. M Treisman, Prof. T. Engen and M. Shailer.

Thanks are also due to the Science and Engineering Research Council for three years financial support , to Prof. S. Newstead and the RAE at Farnborough for providing the facilities and equipment that enabled me to carry out this work.

Declarations.

1. While registered for this degree, I have not been a registered candidate for another award of the CNAA or of a University.

2. This research was funded by a quota studentship award to the Department of Psychology, Polytechnic South West, from the Science and Engineering Research Council. The period of funding was from 1st October 1989 to 30th September 1991.

3. A course of advanced study has been completed, in partial fulfilment of the requirements for the degree. This consisted of guided reading in the area of alarms, supervised by Dr. Edworthy, and attendance at relevant professional conferences.

4. None of the material presented in this thesis has been published prior to submission. The material from Chapter 3 has been presented as a poster ('Scaling the Perceived Urgency of Auditory Warnings') at the conference of the International Society for Psychophysics, August 1990.

ABSTRACT

An Investigation into the Perceived Urgency of Auditory Warnings

by

Elizabeth Jane Hellier

This thesis considers the perceived urgency of sound, with specific reference to auditory warning design.

Psychophysical techniques were investigated as a means of measuring perceived urgency. The biases inherent in different techniques were reviewed. Free modulus magnitude estimation, fixed modulus magnitude estimation, category estimation and cross modality matching were used to scale perceived urgency. On the basis of the cross modality matching validation procedure it was recommended that free modulus magnitude estimation or cross modality matching were used to measure perceived urgency. The successful application of psychophysical techniques meant that the relationship between perceived urgency and objective changes in sound parameters could be quantified.

The effects of changes in four different acoustic parameters, speed, pitch, repetition and inharmonicity were investigated and quantified. It was shown that increases in all the parameters increased perceived urgency. The amount of change in each parameter that was required to communicate a unit change in perceived urgency was revealed.

An attempt was made to see what it was about different acoustic changes that resulted in changes in perceived urgency. In particular, perceived duration was considered as a determinant of perceived urgency. Acoustic parameters were varied in ways known to alter perceived urgency and the effect of these variations on perceived duration was noted. It was shown that one parameter change known increase perceived urgency, increases in speed, increased perceived duration whereas another, increasingly unresolved stimuli, decreased perceived duration. The Reiss Jones(1989) model of temporal contrast was used to explain these findings. It was suggested that changes in perceived duration were part of what makes changes in acoustic parameters communicate changes in perceived urgency. The nature of the relationship between perceived duration and perceived urgency may depend on the type of acoustic parameter used to communicate urgency.

CONTENTS

CHAPTER ONE

INTRODUCTION

Today auditory warnings are employed in many working environments, including hospitals, cockpits, factories and control rooms. They generally take the form of horns, bells, buzzers or sirens that sound either intermittently or continuously to communicate danger or potential danger. Despite the frequency of auditory warning use, little research has been conducted upon warning design or on the behavioural or psychological responses to warnings. As a result operators complain that the warnings are too loud, too numerous, confusing, startling and hard to localise. In this Chapter auditory warnings will be discussed and research on them reviewed. This will show why it is important to investigate the perceived urgency of sound.

The fact that auditory warnings are generally too loud was reported by Patterson(1982) when he investigated the flight decks of fixed wing aircraft. Pilots reported to him that the warnings were loud enough to disrupt essential communication and interrupt thought. This finding was elaborated by Thorning and Ablett(1985). The pilots that they interviewed complained that excessively loud warnings resulted in the immediate cancellation of the alarm sound and not immediate attention to the problem that was being signalled. Similar problems resulting from loud auditory warnings have been reported in helicopters (Rood 1989), and in hospitals where warnings in excess of 75 dBA are frequently recorded (Kerr 1985, Kerr and Hayes 1983, Stanford et al 1988).

According to Thorning and Ablett(1985) loudness contributes to another problem with existing auditory warnings, their confusibility. Confusion among warnings is also said to be caused by their numerosity (Patterson 1982, Kerr et al 1983), indistinctive temporal patterns in their design (Paterson 1982), and by the inconsistent employment of warning sounds between different working environments and faults (Federal Aviation Administration 1977, Kerr 1985). In hospitals the consequences of operators confusing alarm sounds and responding to the wrong one has been acknowledged, at least unofficially, since 1981. According to Cooper and Courvillion(1981), a doctor discussing a fatality explained that,

> "the ventilator alarm was confused with the alarm from the E.K.G. leads."

The problem of auditory warnings being confusible is compounded by the fact that operators report that they are hard to localise, that is, their source is not readily detectable (Parker et al 1984, Tomlinson 1987); by that fact that they are startling (Kerr et al 1983, Edworthy et al 1989); and also by that fact that they can be masked by the background noise (Stanford et al 1985, Szeto et al 1991). Furthermore, there exists a serious mismatch between the perceived urgency of many warning sounds, and the situational urgency of the fault which they are signalling (Momtahan and Tansley 1989). This means that if two, or more, warnings sound simultaneously the operator has no indication from the warnings themselves which fault to attend to first. O'Carroll(1986) is one of many authors who recommend a graded system of alarms whereby the alarm reflects the priority of the fault that is detected. Momtahan(1990) also recommends this arrangement and terms it 'urgency mapping'.

It is apparent that auditory warnings are in dire need of improvement, and calls that they be standardised have been made by, for example, Bock et al (1983), Kerr(1985), O'Carroll(1986) and Hoge et al (1988). Because the problems of auditory warnings generalise to so many working environments, and as the consequences of poor auditory warnings in an emergency situation are potentially so serious, research has been conducted to try to improve them.

Patterson(1982) proposed design principles for the construction of auditory warnings that aimed to counter many of the aforementioned problems. He recommended that warnings were played at 15-20dB above the threshold imposed by background noise, that the temporal characteristics of warnings were manipulated to make them more distinctive, that warnings were attenuated to avoid startle and that the spectral characteristics were manipulated to avoid masking. Pattersons' suggestion that these advanced auditory warnings were constructed from pulses and bursts of sound made the implementation of these recommendations possible. Digital technology such as that described in Appendix 1A has enabled pulses and bursts to be constructed.

The building block of the advanced auditory warning is a pulse of sound lasting from 100-300 ms. As shown in Appendix 1A the amplitude, length, frequency and harmonics of the pulse can to be specified. The pulse is then repeated at different amplitudes and pitches and with different time intervals between repetitions of the pulse. This collection of pulses is referred to as a burst of sound. A burst typically lasts approximately 2 seconds and is like a simple atonal melody. The advantage of

warnings constructed in this way is that the pulse and burst parameters can be varied according to the environment and situation for which the warning is designed and can be manipulated to convey differing levels of perceived urgency so that 'urgency mapping ' can be done. The pulse and burst components of a 'Patterson style' auditory warning are shown in Fig. 1.1.

In a complete warning, Patterson(1985) recommended that a burst was initially played once or twice with the parameters set at a moderate urgency level to attract the operators attention. He recommended that the burst was then played at a low level of urgency to allow operators to communicate. If the fault was not rectified after a specified period of time he said that the burst should be repeated at a high level of urgency which would interrupt the operator and demand immediate attention. Schreiber and Schreiber(1989) proposed a similar idea when they suggested that operating room alarms had advisory, cautionary and warning levels, each more urgent than the other. Kantowitz et al (1988) also recommended that the dimension of perceived urgency was incorporated into auditory warnings. Warnings constructed in this way would adhere to the recommendations made by McIntyre(1986) for the improvement of auditory warnings in operating theatres.

In order that urgency mapping can take place between a warning and the urgency of the fault that it is signalling, and in order that each individual warning can have low, medium and high urgency formats as recommended, knowledge of the effects of different sound parameters upon perceived urgency is required. That it is important for this research to be done experimentally rather than by relying on intuition was demonstrated by Halpern et al (1986) - they discovered that the acoustic determinants of a 'chilling' sound were counter intuitive. Some research in this area has been done, notably by Lower et al (1986), Patterson et al (1986), Momtahan(1990) and Edworthy et al (1988,1991).

On the basis of such research Patterson-style advanced auditory warnings have begun to be accepted into many environments. They form the basis of the British Standards Institute draft standard for intensive care and operating theatre alarms, and have been incorporated into military helicopters (Lower et al 1986, Edworthy et al 1989, Rood 1989). Where they have been introduced, the warnings have been accepted favourably. James and James(1989) found that

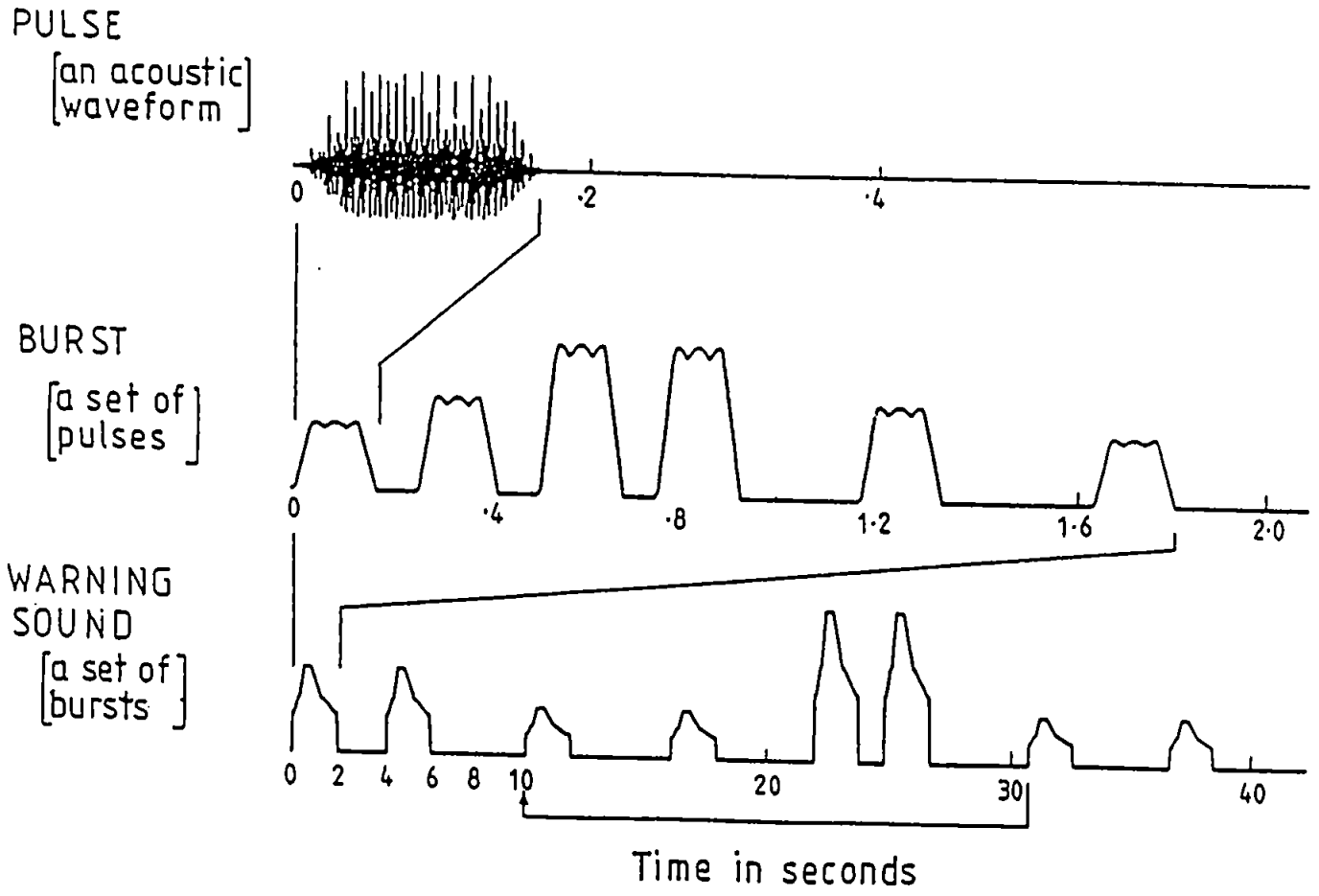


Fig. 1.1 : Components of a Patterson Style Auditory Warning.

helicopter pilots were able to respond faster to prioritised alerting tones than to more traditional warning systems.

Although it is recognised that there are many possible messages that warning sounds might be required to convey (Lazarus and Hoge 1986, Hoge et al 1988, Loxley 1991), this thesis will focus on urgency since it is crucial to the prioritising and urgency mapping of alarms - factors which have been recommended as methods of improving existing warnings. The present research will continue work on the perceived urgency of auditory warnings. Much of the previous research on the effects of individual sound parameters upon perceived urgency has been limited by the measurement techniques employed. Therefore, in Chapters Two and Three, psychophysical techniques will be investigated as a means of measuring subjective continua such as perceived urgency. In Chapters Four and Five and the effects of manipulating various sound parameters upon perceived urgency will be assessed. An attempt will be made to quantify objectively the changes in perceived urgency that result from different parameter manipulations. If this is achieved then warnings designers can use the information as a data base from which to specify the urgency of new warnings. In Chapter Six the time perception literature will be reviewed and links sought between perceived time and perceived urgency. Chapter Seven investigates these links experimentally. The possibility that non-acoustic, determinants of urgency exist will be investigated and the theoretical and practical implications discussed. The thesis thus aims to highlight a suitable methodology for researching perceived urgency, to show how different sound parameters relate to perceived urgency and to uncover possible non acoustic determinants of urgency.

It is well documented that operators respond to aversive auditory warnings by cancelling them and then attending to the problem in hand. According to Stanford et al(1988) the manufacturers of medical equipment are responding to this undesirable situation by introducing alarms that cannot be cancelled until the fault has been rectified. When one considers the possible consequences of forcing operators to conduct emergency procedures in the presence of, perhaps many, excessively loud and confusing alarms that they cannot silence it is obvious how timely it is to conduct research into the improvement of the warning signals themselves.

CHAPTER TWO

AN EVALUATION OF PSYCHOPHYSICAL SCALING METHODS.

2.1. Introduction.

This Chapter examines the claim that sensation can be measured and the claim that there is a universal psychophysical law. The administration and validity of various measurement techniques are discussed.

The greatest advocate of the Nomothetic Imperative (which states that an ordered input output function exists between stimuli and sensations) was S. Stevens. He divided stimuli into two categories, Prothetic and Metathetic. Stimuli that were Prothetic were said to be quantitative and to describe 'how much'. An example of this is loudness - as something increases in loudness it becomes more loud, it changes in amount. Metathetic stimuli on the other hand were said to be qualitative. An example of this is pitch - increases in pitch change the quality or nature of the stimulus, but not the amount of it.

In Stevens (1957) said that for Prothetic continua there was a general psychophysical law relating subjective stimulus magnitude to objective stimulus magnitude, whereby equal stimulus ratios produce equal subjective ratios. In mathematical terms,

$$S=kO^m$$

(Equation 2.1)

where **S** is the subjective variable, the sensation, **k** a free parameter that depends upon the units by which the stimulus and response are measured, **O** the physical variable and **m** is the exponent with a characteristic value for different sensory continua. Stevens recommended that this function be converted to logarithms so that it could be represented by a straight line on a log-log plot, the slope of which would represent the value of the exponent. Thus the form of the psychophysical function was said to be a power function for Prothetic continua. Stevens said that Metathetic continua should be measured in terms of the just noticeable difference.

As Warren(1981) pointed out, many bases for the Power Law have been proposed - for example, the input-output operating characteristics of sensory transducers, Stevens(1957); or biases in scaling techniques, Poulton(1968). Warren said that it was the result of subjects learning the physical correlates of sensory stimuli. Whatever its basis, the data from many sensory scaling studies appear to fit the power function. In 1960 Stevens claimed to have revealed over twenty four different continua for which the Power Law held, and more recently, Kowal(1987) reported that the relationship between the perceived and physical duration of note numbers and musical sequences was also a power function. Despite such support, Stevens(1960) was right to predict that,

> "the announcement of a presumed law in science will trigger prompt and vigorous attempts at its refutation."

Jones and Marcus'(1961) observation that the averaging of experimental data over subjects, which was necessary in the computation of the Power Law, could conceal important information was supported by the many authors who felt that the Power Law only held for a groups, not for an individual's data (on account of the fact that on some continua, large individual differences in the value of the exponent had been revealed, (for example by Green and Luce 1974). Stevens(1971) explained these individual differences by stating that they could reflect either differences in subjects understanding of relative magnitudes, or the different operating characteristics of their sensory transducers. In relation to the latter explanation, it should be noted that whilst individual exponent differences are often very large, it is barely conceivable that such large differences could exist between subjects sensory systems.

Another problem with the Power Law was noted by Engen(1971) who described how weaker stimulus magnitudes deviate from the power function near the threshold (which is dependent on experimental conditions). In 1960, Stevens recommended that stimuli should be measured in terms of their distance from the threshold, having observed that temperature could only be made to fit the power function if it was measured in this way. He also recommended the introduction of an additive constant to the stimulus side of the power function to bring the zero on the psychological and physical scales into coincidence and represent the effective threshold in the current experimental conditions. Other variations on the Power Law have been proposed by, for example, Ekman(1961), Galanter and Messick(1961), Atkinson (1982) and Peleg and Campanella (1988). It is apparent that although there is a certain amount of disagreement as to the precise

nature of the psychophysical law, many authors share Stevens' conviction that it is a power function. Furthermore although authors such as Refinetti(1989) have rejected the idea that there is a universal psychophysical law, many have accepted the devise as a descriptive if not a theoretical tool.

Weiss(1981) rejected the idea that any single psychophysical function could describe sensory intensity by stating that the exponent was dependent upon the way in which the stimuli are measured (when loudness was measured in pressure, exponent = 0.67; when it was measured in energy, exponent = 0.33). He felt that the exponent was meaningful as a descriptor of a single set of stimuli, but that exponents should not be compared across continua. McBride(1983a) re-plotted Steven's(1969) data in linear co-ordinates to check the Power Law prediction that exponents of more than 1 resulted in positively accelerating functions and exponents of less than 1 resulted in negatively accelerating function. Having found that this was not always the case, he also concluded that the power function should only be used descriptively and not theoretically.

Despite such criticisms, Steven's Power Law has replaced Fechners Law (1860) as the dominant psychophysical law, and the frequency with which it is supported is seen as support for sensory scaling. It has heralded the dawn of the new psychophysics, which virtually abandons the indirect measurement techniques of Fechner in favour of direct methods in which the quantitative property desired is stated to the subject in instructions. In 1962 Goude indicated that confidence in direct scale values was warranted for he demonstrated their additivity. As Marks(1974) pointed out whereas the old psychophysics employed the subject as a null instrument, the new discipline attempted to quantify sensory responses and say how they related to physical stimuli by assuming that subjects were capable of judging the magnitude of their sensory experiences.

The development and improvement of psychophysical methods has still failed to convince everybody that attempts to measure sensation are valid. A significant number of authors, for example Poulton(1982) and Fucci et al(1987) said that the results of psychophysical experiments were not simple measures of sensation, but were highly dependant upon the measurement technique employed and the biases specific to it. This chapter investigates this claim and will attempt to reveal the least biased psychophysical methods so that, by studying the relationship between perceived and physical magnitude, we might understand behaviour in relation to the physical energies that control it.

Although Fechner's indirect methods are still sometimes employed, this chapter will investigate only direct scaling procedures. We are thus adhering to Stevens(1961) assertion that subjects are capable of accurately reporting the sensations that they experience. In contrast to indirect scales which, as Engen(1971) pointed out, require supplementation with theoretical assumptions, his is the only assumption that it is necessary to make when employing direct scaling procedures. Acceptance of the Power Law heralded the emergence of ratio over interval scaling techniques, for it is only the former that provide evidence supporting the power function. Stevens(1971) asserts that, since the power function is the only true description of sensory events, then only measurement procedures producing data that fit the function are producing unbiased data. Despite this fact, both direct ratio and interval scaling techniques and the biases that affect each will be investigated in this chapter, so that Stevens assertions do not prejudice our findings. We will attempt to highlight the techniques that report the most valid relationship between physical and subjective stimulus magnitudes. Two authors have previously attempted this task, Warren(1970) and McRobert, Bryan and Tempest(1965). Unfortunately they achieved highly discrepant results, their proposed ratio estimation procedures yielded loudness exponents of 1.0 and 0.42 respectively.

2.2. Interval Scaling Techniques.

This section will review interval scaling techniques and some of the biases that effect them specifically.

Marks(1974) described interval scales as scales of dissimilarity which do not reveal absolute magnitudes, and which have two unspecified parameters, the unit and the zero. Although interval scales can be constructed from ordinal data by assuming that psychological preferences are normally distributed, as already stated, in the present discussion only direct interval scaling techniques, which require subjects to judge the magnitude of sensory intervals, will be considered. These procedures, were developed by Plateau in the 1850's for scaling the reflectance of greys. As Warren and Poulton(1962) noted, all direct interval scaling techniques since have been constructed from a variation of Plateau's procedure.

Despite Stevens(1971) claims that interval scales of Prothetic continua are biased because subjects are unable to make linear, only ratio, partitions on such continua, and that interval scales should only be used on Metathetic continua, the techniques are still used in both cases. For example, Ward(1972) used an interval scaling technique to collect loudness judgment of pure tones. The two most widely used interval scaling techniques are category production and category estimation. They are reviewed below.

2.2.1. Category Production.

Category production is the oldest of the interval scaling techniques. In 1971, Engen described how Plateau employed the category production method of bisection by requiring artists to mix a grey paint so that it was subjectively half way between black and white, in a test of Fechners Law. As is apparent, bisection requires subjects to produce stimuli that subjectively divide a continuum into two equal intervals. Another form of category production that is less widely used is equisection. This requires the division of the continuum into a specified number of more than two equally appearing intervals.

Stevens(1971) concedes that bisection studies have, to a limited extent, confirmed Power Law exponents derived from ratio scaling techniques, but points out that because the lower half of the bisection often appears larger, subjects lower their bisection point and thus tend to produce lower exponents than those arrived at by other means. His assertion that bisection performance represents a compromise between setting the mid-value between the end values in objective terms and setting the mid-value so that the distance between the mid and lower values appears to equal that between the mid and higher values was refuted by Masin(1983). Having required subjects to double and then bisect a sensory continuum, Masin was unable to find evidence for Stevens' claim.

Dissatisfaction with bisection procedures does however remain. Gage(1934) declared them incapable of producing scales of sensory magnitude; whilst Stevens(1955) suggested that they be employed only to test the generality of ratio scales.

2.2.2. Category Estimation.

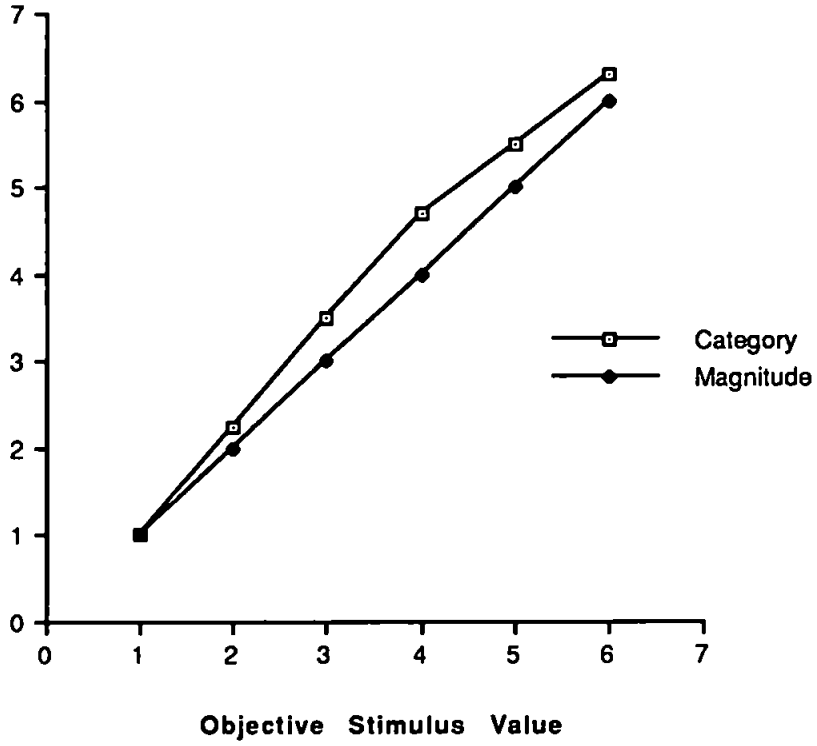
Engen(1971) described the category estimation method of interval scaling by stating that subjects were usually presented with one high and one low anchor stimulus, with its corresponding category number, to define the limits of the psychological continuum. Stimulus values between these anchors are then presented and the subject required to assign each to a numerical category reflecting its subjective value. For example, subjects might hear the slowest sound of a set called category 1, and a the fastest sound of a set called category 10, and be asked to place subsequent sounds in categories 1-10.

As Foley, Cross, Foley and Reeder(1983) observed, category estimation has revealed both a power function (Marks 1968) and a logarithmic (Montgomery 1975) relationship between physical stimuli and subjective magnitude. They said that methodological differences between category estimation experiments made it hard to conclude in favour of either relation. As Stevens(1971) pointed out, when power function relations have been revealed the exponents are often smaller than those achieved by ratio scaling methods - he calls them virtual, as opposed to actual, exponents. Stevens(1960) attributed the occasional discovery of a logarithmic relationship to the fact that category scales are biased by subjects differential sensitivity to stimulus differences over the continuum.

Galanter and Messick(1961) described how this differential discriminability could be utilised in conjunction with category scale values to produce the 'Processed Category Scale'. Typically differential discriminability results in a non linear relationship between category and magnitude scales of the same Prothetic stimuli, (see Fig. 2.1). They said that if scale values were determined by a Thurstonian model allowing unequal stimulus dispersions and category widths, then the nonlinear relationship between category and magnitude scales might vanish. In fact the Processed Category Scale accentuated the non-linear function so that the category scale was a logarithmic transformation of the magnitude scale.

Despite their assertion that category estimation scales were the least satisfactory interval scale and should be avoided, Stevens and Galanter(1957) proposed an iterative procedure to minimise the bias therein. They produced a category scale which met the subjects' expectation that each category should be employed equally often. They stated that this could be achieved if a group of subjects were exposed to

Fig. 2.1 : The Non Linear Relationship Between Magnitude and Category Scales



stimuli that were equally spaced on a scale created by a previous group, and the process repeated until no further change was required. In 1974, Montgomery and Eisler investigated iteration procedures further, and found that they produced category scales with virtually equal intervals between successive stimuli. Although small biases existed they claimed that it was a pure category scale in Stevens and Galanter's(1957) terms. The fact that a Fechnerian Integration Model fitted their data allowed them also to conclude that category estimation techniques produced discrimination scales.

Torgerson(1961) to advocated the use of category estimation techniques for assessing colour judgement. He felt that the location of colour in multidimensional space necessitated a scaling method that allowed distance relations to remain invariant through a change in direction. Despite such advocates a general dissatisfaction remains with all category scaling methods, and with the interval scales that they produce. This can partly be accounted for by the fact that several biasing factors are specific to interval scaling techniques. They are discussed below.

2.2.3. Biases Specific To Interval Scaling Techniques.

2.2.3.1. Number of Categories.

Although Stevens and Galanter(1957) said that category scales were unaffected by the number of categories that were used, in 1974 Eisler and Montgomery compared seven and fifteen point category scales and concluded that more categories resulted in a more linear scale.

2.2.3.2. Category Label.

In 1968, Marks revealed that category scales were affected by the numbers that the categories were labelled. He noted that, as the category number increased so did the exponent of the best fitting power function.

2.2.3.3. Stimulus Discriminability.

Stevens and Galanter(1957) found that stimulus discriminability, and to a lesser extent, stimulus spacing and frequency, altered the width of subjects' categories.

2.2.3.4. Anchor Effects.

Johnson and Mullally(1969) observed that the employment of anchors outside the stimulus set caused the displacement of judgment of stimuli near to the anchors. End stimuli were judged extreme in relation to the stimulus set and less extreme in relation to the anchors. Thus, anchors with a lower value than the stimulus pushed judgment down, whilst anchors with higher values raised judgment.

2.2.3.5. Hysteresis.

Hysteresis is the particular name given to assimilation effects when they occur in bisection studies, it causes the bisection point to be higher when stimuli are presented in an ascending, rather than a descending order.

Whilst conceding that category scales were useful for threshold determination, Stevens(1971) asserted that their high susceptibility to bias made them best avoided. He felt that since interval scale data rarely fitted the Power Law, which was the true expression of the stimulus/sensation relationship, it was biased, and noted that if interval scales were required they could be derived from ratio scales.

2.3. Ratio Scaling Techniques.

In this section a selection of the most commonly employed ratio scaling techniques are reviewed. Such techniques produce scales that contain interval, ordinal and nominal information and have a true zero and meaningful relationships between the scale values. Ratio scale values preserve the ratios between experimental stimuli. According to Stevens(1960) the only admissible transformation of a ratio scale is multiplication by a constant, for example a transformation from inches to feet. He claimed that more general transformations resulted in the loss of information.

Mashour(1965) doubted that scales constructed by numerical estimates could be considered ratio scales. He suggested that the variation of the Stevens Power Law exponent under different ratio scaling conditions meant that the response scale was only a quasi-ratio scale that produced power functions as if it were a ratio scale but empirically only met the requirements of an ordinal scale. A similar criticism was made by Hellman and Zwislocki(1968) who suggested that in

magnitude estimation and production subjects did not employ ratio, but absolute scales (which imply that numbers and sensations have absolute psychological values). They asserted that people could only follow a ratio scale within a restricted range of modulus. Schneider and Bisset(1988) on the other hand claimed that subjects found it easier to judge ratios of continua that were easily decomposable, than it was to judge differences (as is necessary in interval scaling techniques). They described decomposable stimuli as those which could be easily divided into smaller units mentally, such as line length. That ratio scales should be valid is not only important in terms of sensory scaling - as Mashour and Hosman(1968) observed, while their validity is in dispute the mathematical form of Stevens Power Law cannot be settled since data supporting the Power Law comes almost entirely from ratio scaling experiments.

Despite debate as to the validity of ratio scales, ratio scaling techniques have been widely used. In 1957, for example, Stevens constructed ratio scales of fourteen perceptual continua; furthermore, in 1975 J. Stevens described how ratio scale quantification had been introduced into sociology, criminology and politics.

2.3.1. Magnitude Production.

Stevens(1955) described the method of magnitude production by stating that numbers were presented irregularly to the subject who was required to adjust the stimulus to a level that matched the numerical value. It was recommended that anchor stimuli should be avoided, (to prevent the task becoming one of category estimation), and that the numbers presented should approximate a geometric progression. Stevens and Poulton(1956) went on to recommend that when producing loudness subjects should only use the upper three quarters of the decibel attenuator because adjustments in the lower quarter were not considered fine enough to reflect small changes in magnitude. They elaborated this point to show that inexperienced subjects were highly influenced by the apparatus that they used in magnitude production tasks. It was recommended that the amount of dial movement should be proportional to the magnitude of the stimulus that subjects were producing. Stevens(1955) also pointed out that subjects' aversion to extreme stimuli was likely to effect the way that adjustments were made, and that exponents arrived at by magnitude production were usually higher than those arrived at by magnitude estimation.

Stevens and Guiaro(1962) on the other hand compared magnitude estimation, magnitude production and category production and concluded that the magnitude production data that best fitted the straight line logarithmic plot predicted by Stevens' Power Law.

2.3.2. Magnitude Estimation.

Magnitude Estimation, which was probably first employed by Richardson and Ross in 1930, is the most direct approach to sensory scaling. It attempts to avoid all restrictions and biases imposed by the experimenters control of the subjects response system - working on Stevens(1965) principle of 'minimum constraint' subjects are only instructed as to what scale unit to use, all other response decisions are their own. There are several variations of the magnitude estimation procedure, but in all instances subjects are required to assign numbers to stimuli in proportion to the magnitude thereof.

An early magnitude estimation procedure was described by Stevens(1956) who said that a comfortable stimulus level should be employed as a standard and stimuli be presented above and below that level. Randomised order of stimulus presentations, short experimental sessions (about ten minutes) and an easily multiplied and divided modulus were also advocated. Marks(1974) provided a more detailed, but similar description. He recommended that stimulus size and intensity should be varied on each presentation and that stimuli should be equally spaced in logarithmic steps. Marks also said that faster stimulus presentations were more advantageous, that each stimulus should be presented twice (unless there were a great deal of stimuli) and that the number of subjects should be increased with the number of stimulus parameters that were being varied. He said that usually, ten or twelve practised subjects were sufficient. Marks recommended that experimenters should avoid employing a standard or a modulus.

Although Stevens(1971) claimed that data was not affected by subjects' use of different moduli if a modulus equalisation procedure was employed to adjust judgment to a common modulus, Marks rejection of the standard and modulus was reflected in the 'free modulus' magnitude estimation procedure proposed by Engen(1971), and supported by Green and Luce(1974), and Foley, Cross and O'Reilly(1990). This procedure, designed to avoid any variance caused by subjects choice of different modulus was introduced by three, often unidentified, practice trials, and data was transformed to eliminate any inter/intra individual

variance without affecting individual slopes and intercepts. The free modulus method of magnitude estimation was used successfully by Wilson and Stelmack(1982) to measure loudness.

In 1974 Marks discussed the advantages of magnitude estimation procedures. He claimed that they were fast and easy to employ, could consistently be applied to Stevens Power Law, could be used when it was impossible to give subjects the control of the stimulus level (as is necessary in production techniques), and that they avoided any test x comparison stimulus interaction because only one stimulus had to be presented at a time. He said that the ratio scales produced by magnitude estimation were internally consistent. Stevens(1959) stated that it was magnitude scales that most closely reflected the input-output functions of the sensory transducers.

Magnitude estimation has recently been employed by, for example, Haverland(1979) to investigate human factors variables; Fucci, Harris, Petrosino and McMath(1987) to investigate suprathreshold sensation magnitudes; by Kowal(1987) to investigate the perceived length of musical sequences; and by Fagot and Pokorny(1989) to study judgment of loudness and heaviness.

2.3.3. Ratio Production.

Stevens(1957) described how the ratio production procedures of multiplication and fractionation required subjects to adjust a stimulus until it was a required ratio, for example a third, of the standard. In fractionation the adjusted stimulus was smaller than the standard and in multiplication it was larger. Stevens, J. And Tulving(1957) varied the ratio production procedure by setting two lights at a given ratio to each other and requiring subjects to adjust two tones to represent the same ratio in loudness.

Although both ratio production techniques are simple - the subject only has to consider one ratio at a time, Garner and Hake(1951) noted that they may be particularly susceptible to context effects. Stevens(1971) said that the biases in fractionation and multiplication were reflected by the fact that neither procedure produced data fitting the power function. He said that it was no longer necessary to employ ratio production techniques for better scaling methods had been developed.

2.3.4. Ratio Estimation.

Although ratio estimation was first suggested by Metfessel(1947), the most popular specific procedure was that developed by Ekman(1958). He instructed subjects to divide 100 points between a pair of stimuli according to their subjective intensity. This has also been called the constant sum method (Engen 1971). Ekman noted a variation of the procedure whereby subjects were required to state directly the ratio between a pair of stimuli, as did Stevens(1971), who said that for the sake of simplicity, subjects could be asked to state what percentage one stimulus was of another. It should be noted that despite his advocacy of ratio scaling techniques, Stevens is implying in this suggestion that subjects may find ratio judgment hard to make.

Specific drawbacks to the ratio estimation procedure have been noted. Ekman(1958) felt that subjects tended to use a constant range of numbers regardless of the range of stimulus values, and that the task appeared artificial to the subject; whilst Luce and Green(1974) suspected that subjects employed categorisation strategies during ratio estimation tasks.

Despite the fact that ratio estimation results have been shown to agree with those of ratio production (Guilford and Dingman 1954), ratio estimation is rarely employed, but is usually rejected in favour of magnitude estimation.

2.3.5. Random Production.

Although Stevens may classify Banks(1974) method of random production as a partitioning technique since it produces smaller exponents than the method of magnitude estimation, Banks insisted that random production could produce a ratio scale for any continua with qualities that could be separated by the subject. His procedure required subjects to generate a specific number of stimulus values, evenly spaced between the top and bottom of the continuum. In experimental demonstrations of random production, Banks(1974) scaled force of hand grip, intensity of electric shock and size of area - in each instance the Power Law exponent was found to be lower than that achieved by magnitude estimation.

Having seen that the common direct scaling procedures can be divided according to whether they scale ratios or intervals, and that within each of these dichotomies, techniques either require production of stimuli (to fit a category, or reflect a

numerical value, or a ratio, or to divide a continuum equally) or estimation (of a category, a number or a ratio) it is possible to look at the biasing factors that effect all of the techniques to varying degrees. Although only a few specific biases have been discussed in relation to ratio scaling procedures, many of the biases that will be mentioned affect ratio as well as interval techniques - premature conclusions as to the superiority of ratio scaling techniques are therefore unwarranted.

2.4. Biasing Factors in Psychophysical Techniques.

In this section potential sources of bias for interval and ratio scaling techniques will be discussed.

In order to obtain accurate judgment of sensory intensity it must be ensured that subject's judgments only result from the stimuli that are presented, and not from extraneous factors. Stevens(1971) answered criticisms of his Power Law, based on the observed variability of exponents, by stating that this variability was caused by biasing factors which could be eliminated. Poulton(1982) and Mellers and Birnbaum(1982) also said that judgment could be made free of bias.

To eliminate bias from judgment of sensory intensity, it is important to know which scaling techniques are affected by which biases. In 1979 Poulton specified the way in which different types of response were differentially susceptible to bias. He said that when subjects responded in familiar physical units their judgment were virtually free of bias; that responses in named or numbered categories were only slightly affected; that magnitude judgment were more biased and that cross modality matching (to be discussed) was the most highly biased scaling technique. On the other hand, supporters of Stevens Power Law have maintained that category judgment are more biased than ratio judgment and that cross modality matching is an unbiased validation procedure - the argument can only be resolved by a closer look at biasing factors.

There are two main types of bias, context effects and response bias.

2.4.1. Context Effects.

In 1974, Birnbaum said that ratio and interval judgments were dependant upon context from outside and inside the laboratory. Although other authors, for

example Helson(1964), supported Birnbaum by stating that contextual bias could affect all scaling behaviour, Poulton(1968) pointed out that many authors underestimated its importance and did not adequately report experimental conditions. In 1956 Stevens refuted claims that context effects were powerful enough to account for his Power Law by stating the Power Law was supported by subjects' very first judgment (which are unaffected by context). Warren and Poulton (1962) also said that context effects could be avoided entirely if subjects were only ever required to make one judgement.

Avoidance techniques of different types have been proposed by many authors. In 1964 for example Aiba and Stevens showed how the use of qualitatively identical stimuli could lower contextual bias, while Poulton(1979) recommended the use of a complete between subjects design, logarithmic stimulus spacing and the avoidance of examples to prevent context effects in magnitude estimation. In support of Stevens (1966) and Birnbaum(1974), who felt that context effects could only be eliminated in ratio scaling techniques, Ross and Dilollo(1971) did not talk about avoiding, only of minimising, context effects. They felt that this could be achieved by providing subjects with a general context that would allow the experimental context to fall into place. They said that this could be achieved by training subjects to measure the relevant property. Similarly, Mellers(1983) examined the study of Zwislocki and Goodman(1980) who claimed that the unconstrained procedures they employed allowed subjects to use an absolute scale unaffected by context; but found that evidence of context effects remained in their data.

It appears that context effects influence subjects when they are making interval and ratio responses, and that these effects can be minimised, perhaps more readily in ratio than in interval scaling procedures. Context effects come in many different forms which differentially affect different scaling techniques, they are reviewed below.

2.4.1.1. Effect of the Standard.

In 1955, Stevens conceded that the level of the standard stimulus altered the value of the exponents arrived at by magnitude estimation, bisection and ratio estimation. This conclusion was supported by Poulton(1968) who noted that lower stimuli produced steeper slopes if a low standard was employed, by Ross and Dilollo(1970) who attributed their failure to achieve a power function fit for the

heaviness of lifted weights to the effect of the standard, and by Mellers and Birnbaum(1982) who found that in magnitude estimation tasks a higher value for the standard resulted in higher responses.

Jones and Woskow(1966) said that the magnitude of the standard exerted its influence by altering the range over which subjects distributed their responses; whereas Macmillan et al(1974) claimed that the standard had an effect merely as a result of its presence, not as a function of its size. More recently, Fagot and Pokorny(1989) used a model of relative judgement to account for the effects of the standard.

In 1956 Stevens and Poulton described other manifestations of the effect of the standard. They found that in ratio production tasks, the distance between the standard and the first stimulus influenced the results. If the first experimental stimulus was close to the standard fractional estimates revealed steeper slopes, for multiple estimates however a close standard and first stimulus lessened steepness. These results were confirmed by Poulton(1969).

The obvious solution to the aforementioned problems, as advocated by Engen(1971), Marks(1974), Zwislocki and Goodman(1980) and Wilson and Stelmack(1982), is the abandonment of the standard. However Stevens(1956) noted that in such cases, wide ranges of numbers were employed by subjects, making averaging data difficult. Furthermore he stated, as did Poulton(1968), that in the absence of a standard subjects may treat the first stimulus that they hear as a standard. As an alternative strategy for avoiding the effects of the standard Poulton and Simmonds(1963) recommended combining the results from multiple and fractional estimates (which they felt were affected equally, but in opposite directions by the standard) to cancel out any effects. They discovered however that the two forms of judgement did not always result in effects that were equal and opposite, and so later recommended that only the first judgement from each subject should be analysed.

Although no particularly specific avoidance strategies appear to have been recommended to deal with the effect of the standard (bar Engen's 1971 free modulus magnitude estimation), Poulton(1968) said that the standard exerted less of an effect than other factors, and in 1989, Fagot and Pokorny were unable to find any effect of the standard on ratio estimates. Stevens(1959) went on to say

that the effect of the standard upon exponent values was only second order. He said that normalising data would enable the effects of the standard to be assessed.

2.4.1.2. Effect of the Modulus.

In 1957, J. Stevens and Tulving observed that the size of the exponent obtained was often related to the size of the modulus (the number assigned by an experimenter or subject to the standard). Hellman and Zwislocki(1961) supported this observation and showed how, in ratio production, changing the modulus by a factor of one hundred changed the median fractional estimate by a factor of ten, and the median multiple estimate by a factor of two.

In 1968, Poulton described this effect. He said that since increasing the modulus increased the set of numbers available for fractional estimates, the resulting exponent was increased. He described the opposite effect for multiple estimates.

Lane, Catania and Stevens(1961) attempted to determine a scale of autophonic output (subjects' estimation of their own vocal response) by magnitude estimation and production. They observed a larger effect of the modulus in the latter task and said that it could be reduced by a data normalisation procedure to make subjects productions more normally distributed. The free modulus magnitude estimation technique again offers a means of avoiding this source of potential bias. If that technique were unsuitable, given that the effect of the modulus is said to be very small (Poulton 1968), then the proposed data treatment should be a sufficient precaution against it.

2.4.1.3. Effect of Stimulus Spacing.

Poulton, Edwards and Fowler(1980) described stimulus spacing bias as a non-linear bias that occurred when subjects responded as if all stimuli were equally spaced geometrically and equally probable. They said that both category and magnitude judgment could be affected. Although Eisler and Montgomery(1974) said that of the two it was magnitude judgment that was more susceptible, the majority of authors, for example Marks(1974), claim that category judgment is more highly affected by stimulus spacing bias. Marks said that the effect occurred because closely spaced stimuli caused a stretching of the numerical responses in that area, which resulted in a local increase in the power function. When scaling loudness and brightness, Stevens and Galanter(1957) found that more uniform

stimulus spacing reduced the steepness of the resulting power functions at the low end.

Stevens(1957) felt that experimental iteration could be used to produce a stimulus spacing that did not bias experimental results, allowing the production of a 'pure' category scale. Poulton(1982) also suggested that theoretically unbiased stimulus spacing should be employed if it were known. A specific iterative technique for neutralising stimulus spacing bias was proposed by Pollach(1964). He said that more stimuli should be put in areas where the slope of the rating scale was steep, to ensure that each category could be used equally often (as subjects expected).

Although Pollach also suggested using category production to avoid stimulus spacing bias, many authors would find this unnecessary for they find little evidence to suggest that such a bias even exists. McBride(1983b) for example, found that category scales of taste were unaffected by stimulus spacing bias; while Pradham and Hoffman(1963), who used nine different stimulus spacings for the magnitude estimation of weight, agreed with Stevens(1956) statement that stimulus spacing only played a small part in determining estimates.

2.4.1.4. Stimulus Bias.

Stimulus bias refers to any aspect of the stimulus that can affect judgement. One such aspect is stimulus duration. Raab and Osman(1962) commented that very short stimuli made magnitude scaling harder, whilst Van Orden, Sturr and Taub(1987) found that when short and long flashes of light were intermixed in a study, brightness judgment of the short flashes resulted in steeper power function slopes.

In 1956 Stevens pointed out that stimulus level could also effect the exponent. He found that subjects underestimated faint sounds and overestimated loud ones. It seems that Marks(1988) was right to state that care should be exercised in the selection of all aspects of experimental stimuli.

2.4.1.5. Transfer Effects.

Transfer effects describe what happens when the effects of judgement under one condition carry over to effect judgement in another condition. According to Poulton

and Freeman(1966) symmetrical transfer occurs when subject's performance is always better in a second condition. This could be considered an effect of practice, which Eisler(1974) says increases the size of the exponents in magnitude estimation. Poulton and Freeman(1966) suggested that a simple within subjects balanced design could avoid this effect, and the effect of negative asymmetrical transfer (whereby subjects performance always deteriorates in a second condition).

Apparently biases such as these are not avoided by employing one scaling method rather than another, but by consideration of experimental design in relation to the task that subjects are required to perform.

2.4.1.6. Sequential Effects.

A. Assimilation.

Holland and Lockhead(1968) described the relationship that they had found between the response and the immediately preceding stimulus as 'assimilation'. In 1983, Lockhead and King proposed a model to describe assimilation. They said that each stimulus was assimilated to the memory of the previous one, and that the previous stimulus and the memory thereof differed from each other in predictible ways.

This process has been witnessed by many authors, for example Cross(1973) found that in magnitude estimation, judgments were assimilated towards previous stimuli. Cross claimed that assimilation resulted in underestimation of the exponent. Allen(1983) also found that assimilation occurred when subjects judged the length of tones, and it appears that only Marks(1988) claims that the effect is negligible - he found no evidence for assimilation in his matching studies (to be discussed).

B. Contrast.

According to Holland and Lockhead(1968) contrast is a consistent inverse relationship that has been seen to exist between responses and the average value of all preceding stimuli, and responses and the immediately preceding stimulus. Lockhead and King(1983) felt that it occurred because subjects tried to keep track of the labels that they had used for previous stimuli, so that they could employ a reliable response scale. When assimilation necessitated a shift in the response scale, contrast occurred a few trials later. They also said that if no feedback was

provided, contrast would occur because subjects would feel that they had not used all of the available response space. Contrast has been demonstrated in many experimental settings, for example by Dilollo(1964).

Holland and Lockhead(1968) were among the first to investigate sequential effects in psychophysical judgement. They examined subjects responses to auditory stimuli, and concluded that assimilation and contrast were the response consequences of memory. They attributed the biases to subjects faulty memory of the standard stimulus, a memory that was contaminated by other stimuli and by the decaying memory traces thereof. They felt that responses could be assimilated to previous stimuli and either assimilated to or contrasted from those further back, depending on the presence or absence of feedback.

Since then many different hypotheses have been proposed to account for sequential effects such as Ward and Lockhead(1971), Cross'(1973) Response Ratio Hypothesis, Jesteadt , Luce and Green(1977), King and Lockhead(1980) and Ward(1985). The methods proposed to avoid sequential effects have depended on which of these theories was supported. Jesteadt et al(1977) for example, considered sequential effects unimportant, especially in magnitude estimation. Despite their finding that sequential effects did not operate in ratio estimation tasks, the authors continued to use magnitude estimation, for the sake of experimental convenience. Lockhead and King(1983), on the other hand, believed that reducing the computations that subjects had to perform would reduce sequential effects, and so advocated the use of successive ratio tasks, in which subjects were required to judge the ratio between current and previous stimuli. Atteneave(1962) said that sequential effects in bisection studies could be balanced out by an iterative procedure that produced a stimulus spacing encouraging the use of all categories with equal frequency.

2.4.1.7. Effects of Feedback and Instructions.

The effects of feedback on judgement were documented by Siegel(1972) when he noted that performance decreased at specific retention intervals if feedback was not provided. His data supported the idea that feedback influenced subjects' decision rule, not their sensitivity to stimulus differences. Also by Kreuger(1984) who discovered that feedback greatly reduced the variability of individual exponents in magnitude estimation and production. The decision to employ feedback should be

taken bearing these effects in mind, for no strategies have been proposed to avoid them.

As Stevens(1958) recommended, instructions to subjects should clearly specify what it is that they are required to judge, unfortunately however it has been shown that judgement can be affected by the wording and examples used in instructions. Teghtsoonian(1965) for example found that asking subjects to judge actual or apparent size resulted in different exponents . Furthermore, despite Macmillan et als'(1974) claim that instructions had no effect upon judgement, Goldner, Reuder, Riba and Jarmon(1971) found that ego-orienting instructions increased response variability, and, as Stevens(1971) predicted, lowered the exponents. Interestingly this effect only occurred when the experimenter was of equal or higher status than the subject.

The type of example employed in instructions has also been shown to affect judgment. Robinson(1976) discovered that if a larger range of numbers were used as examples then larger exponents would result from subsequent judgment. This finding was replicated by Mellers and Birnbaum(1982). Similarly, McBride(1983c) stated that the method of presentation of taste stimuli effect the exponent that was obtained. He reported a study by Meiselman(1980) which showed that the 'sip' method of presentation resulted in higher exponents than other methods.

Great care should be taken in the preparation of experimental instructions to ensure that the task is performed in the required manner. Empirically validated standardisation could probably eliminate experimental variability caused by different instructions and examples.

2.4.2. Response Bias.

Jones and Woskow(1966) described subjects' use of a characteristic set of responses regardless of the stimulus set as response bias. Garner and Hake(1951) said that this occurred when the subject appeared to have an idea of the set of responses that it was reasonable to use at the start of the study. Stevens(1971) held response bias in interval scaling techniques responsible for the fact that ratio and interval scaling techniques did not produce consistent results when used to scale the same Prothetic continua. One example of response bias was highlighted by Louge(1961) who hypothesised that the temporal stability

of exponents that he discovered over an eleven week delay following the magnitude estimation of loudness was a response bias caused by learning. Many other forms of response bias exist, affecting different aspects of the response process.

2.4.2.1. Logarithmic Bias.

Logarithmic bias, which is introduced by the subject and affects even their very first judgment, occurs when responses involve a step change in the number of digits used, (usually from 1 to 2 digit numbers, e.g. from 9 to 10). If responding linearly subjects would use numbers 1-9, followed by numbers 10,11,12 etc., so that 2-digit numbers would be used ten times more often than single digit ones, and 3-digit numbers would be used ten times more often still. Poulton(1982) said that in fact subjects responded logarithmically, using number 10 followed by 20, 30, etc., so that single digit numbers are used as often as 2-digit and 3-digit ones. Poulton, Edwards and Fowler(1980) said that this shrank the upper part of the numerical scale. They said that bias free data was not available for numerical judgment unless a step change in the number of digits available to the subject was not allowed. They recommended that the bias could be avoided in category rating experiments by using less than ten categories. Since a comparable strategy for ratio scaling techniques would mean that a magnitude estimation task could only use the numbers 1-9, it can be said that logarithmic bias is only satisfactorily avoided in interval scaling techniques. In 1986, Teghtsoonian and Teghtsoonian contradicted Poulton's claims and said that subjects used number in a linear manner, at least when judging loudness. They claimed that Poultons' data was derived from experiments using too small a range of experimental stimuli for him to be able to tell accurately which response type was being employed.

2.4.2.2. Effect of Number Use.

Ekman et al(1968) said that an important source of response bias was subjects interpretation and handling of numbers. Similarly, Jones and Marcus(1961) claimed that subjects used number in an individual and consistent manner that resulted in each producing a characteristic range of responses in magnitude estimation tasks. In 1962, Atteneave explained the curvilinear relationship that is often found to exist between ratio and interval scales (Fig. 1), in terms of number use. He said that magnitude estimation required subjects to represent subjective magnitudes in terms of number, despite the fact that number may itself have a subjective value, non-linearly related to arithmetic number. He felt that

the power function should match subjective magnitude to the subjective numerical value. Like Duda(1975), he proposed a two stage model of magnitude estimation in which the sensory input and output of numbers were described by separate power transformations. This model was supported by Curtis, Atteneave and Harrington(1968) for individual data. Their group data however supported Stevens(1971) claim that subjects used number in a linear fashion.

Number use does not only affect judgement because it may be non-linear, but also because subjects have preferences for specific numbers and for using round numbers, a fact noted by J. Stevens(1975) and by Schneider(1981). J. Stevens, (1975) conceded that the number continuum was warped but felt that some biases could be eliminated, for example, that caused by subjects' reluctance to employ fractions and large numbers. To do this he suggested presenting subjects with such faint stimuli in practice trials that they were forced to use fractions to describe them. Zwislocki and Goodman(1980) explained this reluctance by stating that subjective number scales arise from an early awareness of numerosity that does not involve fractions. They advocated the use of magnitude production techniques to overcome problems of number use. Poulton(1979) said that the use of practised subjects and familiar measurement units would avoid such problems.

Although Stevens(1956) stated that subjects number preferences only had a small effect on judgement, as has been demonstrated, there is evidence to suggest that subjects' number use alters experimental results. This evidence questions the assumption that underlies magnitude estimation - that subjects can use number to make judgments. It is rarely claimed that category judgments are affected by number use. It should be noted however that even in ratio scaling techniques such effects are not reported by all authors - in 1989 Higashiyama and Tashiro scaled perceived distance and found that subjects were not reluctant to use extreme numbers in magnitude estimation.

Other strategies, besides number use, have been found to contribute to experimental variability. Stevens(1956) found that those subjects who did not visualise a linear scale whilst making judgment gave atypical responses; whilst Milewiski and Iaccino(1982) concluded that subjects formed situation specific response strategies. The judgement strategies that subjects might have employed should perhaps be examined with experimental data; alternatively, instructions might be designed so as to ensure that an appropriate and uniform strategy is used by all subjects.

2.4.2.3. Range Effect.

Garner's(1953) observation that psychophysical judgement often depended on the range of the stimuli that were employed has been supported by many authors. Poulton(1968) found that large stimulus ranges resulted in small exponents, as did Stevens(1971), Gravetter and Lockhead(1973) and Mellers and Birnbaum(1982). Poulton(1968) said that this effect was stronger towards the end of an experiment. The range effect appears to be pervasive for Teghtsoonian(1973) stated that range affected all continua, and Mellers and Birnbaum(1982) said that all scaling techniques were susceptible to the effects of stimulus range. Range effects have been demonstrated by Marks(1968) in category scaling techniques, by Poulton and Stevens(1956) in ratio production and estimation, and by Teghtsoonian and Teghtsoonian(1978) in magnitude estimation and production. McBride(1986c) said that the effect was inevitable. He said that when the power function was fitted to data, the exponent was a reflection of the ratio between the log. stimulus range and the log. response range, so that a large response range and a small stimulus range would always produce a large exponent.

An attempt to explain the range effect was offered by Parducci's(1974) Range Frequency model which saw judgement as a compromise between adjustment to stimulus range and to differential stimulus frequencies. Teghtsoonian(1971) observed that exponents were inversely proportional to stimulus range. He said this was because different sensory systems respond over different dynamic ranges, but produce approximately the same range of sensory response. This model was consistent with Stevens' well documented idea that exponents reflect the operating characteristics of sensory transducers, and with Pradham and Hoffman's(1963) observation of a range x subject interaction. Robinson(1976) had a similar idea to Teghtsoonian's - he said that the range effect was partly caused by the fact that subjects were always presented with stimuli that spanned a proportion of their dynamic range. It is possible that the dynamic range is part of the phenomena that is being measured. In 1978, Teghtsoonian and Teghtsoonian said that the exponent was a reflection of subjects range of sensitivities. In 1973, Gravetter and Lockhead supported Pollack's(1952) idea that discrimination decreased as stimulus range increased. They felt that this accounted for the range effect, and cited the fact that stimulus repetition decreased the effect as evidence for their model.

Because range effects are so prevalent, many avoidance procedures have been proposed. In 1956 Poulton and Stevens found that the range effects in ratio production and estimation were the mirror image of each other. They therefore proposed a balanced design, involving both procedures, and taking the true exponent as the point where the range lines of the two procedures crossed. Teghtsoonian(1973) observed that free modulus magnitude estimation techniques reduced the range effect in loudness experiments, whilst Stevens(1971) stated that stimulus repetition should be avoided if the stimulus range were small, for in such circumstances repetition was found to increase range effects (when the stimulus range is not especially small, repetition of stimuli has been shown to reduce the range effect, Gravetter and Lockhead, 1973). Stevens(1971) obviously felt that such avoidance procedures could be successfully implemented for he did not entertain the possibility that range effects could detract from the validity of sensory scales. He said that their existence could not detract from the very real differences that existed between exponents of different sensory continua.

Two specific types of range effect that have been discussed are the equalising biases.

A. Stimulus Equalising Bias.

Stimulus equalising bias was described by Poulton(1982) as occurring when subjects used the full range of responses, regardless of the stimulus range. They therefore magnified their response scale to fit a large stimulus range or shrank it to fit a small stimulus range. Poulton(1979) said that this resulted in category ratings that were dependant upon the range of the values employed, and in magnitude estimations distributed over the entire range of stimuli (so that if the range were small, a steeper slope would result). Stimulus equalising bias was thought to have a particularly pronounced effect on stimulus dimensions were unfamiliar.

Poulton(1979) proposed ways of lessening the effect. He said that stimulus and response scales that were linked by well-known rules were less affected, and that stimulus and response scales of the same subjective size should be employed, (this is very difficult to achieve in category ratings and magnitude estimation, for it is hard to know the subjective size of the response scale).

B. Response Equalising Bias.

Response equalising bias, the inverse of stimulus equalising bias, was described by Poulton(1982) as occurring when subjects distributed their responses over the entire stimulus range, regardless of the size of the response range. This meant that what ever the size response range a subject was given, he or she would use all of it to describe the stimuli, using a larger response range when one was available, regardless of the size of the stimulus set. He claimed that this bias was unavoidable in category judgement techniques (where it was especially prevalent if only a few categories were being judged), but that it could be avoided in magnitude estimation if the choice of response range were left to the subject.

2.4.2.4. End Effect.

Marks(1988) pointed out that the intensities that define the end-points of stimulus presentations affect judgment of sensory equivalence. In 1968 he noticed how the particular values of the end stimuli had flattened functions derived from category judgment of brightness. Siegel(1972) said that Erikson and Hake's(1957) Subjective Standard Hypothesis best explained the effect of end stimuli. The hypothesis said that subjects stored the end points of stimulus sets in memory and used them as reference standards. The subsequent prediction, that end effects should be minimal when retention intervals were shorter, was supported by Siegel(1972). Eisler and Montgomery(1974) said that end effects could be avoided if extreme stimuli were closer together, and the end points thus less discriminable.

2.4.2.5 Regression.

Central tendency, contraction bias, or regression, occurs when subjects centre their range of responses on the stimulus range so that responses regress towards the mid-point of the stimulus range. Marks(1988) said that judgments made to each qualitatively different subset of stimuli were shifted towards the average perceived magnitude of the other subset. Stevens and Greenbaum(1966) claimed that the effect could be made worse by harder tasks and by incommensurate response ranges (such as those that could be found in magnitude production if the range of possible adjustments on one variable were limited by the apparatus). Johnson and Mullaly(1969) said that comparison of the slope of the regression lines from estimation and production tasks could be used to assess the amount of regression that had occurred.

Several different causes for this effect, which according to Stevens(1970) results in magnitude estimation exponents being underestimated, have been proposed. Central tendency in category judgment is usually explained in terms of Johnson and Mullaly's(1969) Correlation and Regression Model which saw the stimulus-response relationship as an example of statistical regression. In 1973, Cross took up a suggestion made by Garner and Hake(1951), who suggested that sequential effects contributed to regression. His experiments revealed assimilation and an underestimation of the exponent, and he was able to conclude that the presence of an order bias was a sufficient condition to cause regression. Stevens, J. (1975) supported this idea when he stated that there were four main contributors to regression; subject preference for comfortable stimulus levels, experimental noise, stimulus order and subject caution (which made them unwilling to use wide ranges of the variable under their control). Teghtsoonian and Teghtsoonian(1978) modified Stevens and Greenbaum's(1966) explanation of regression to say that at small ratios, magnitude estimation led to an overestimation of the exponent (not an underestimation), and magnitude production led to an underestimation (not an overestimation) because subjects tended to avoid extreme judgmental ratios. In 1984, Kreuger refuted claims that individual differences in subjects judgement ranges caused regression. He said that , were this the case, a negative, not a positive, correlation would be found between magnitude estimation and production exponents.

Regression has been demonstrated frequently, by for example, Tulving(1954), Stevens and Marks(1965) and Dawson and Brinker(1971), and is considered by Stevens(1971) to be the most obstinate bias - it is no surprise therefore to note that many strategies aimed a reducing it have been proposed. Stevens, J.(1975) advocated interchanging the fixed and adjustable stimuli (by using magnitude estimation and production) or avoiding regression by for example matching continuum A to continuum B, followed by matching continuum A to continuum C, so that the ratio of these exponents would be that derived by matching B to C directly, and no regression could have occurred. His suggestion that the exponents of two regression lines could be combined by their geometric mean to produce an unbiased exponent is less satisfactory because the true exponent could lie closer to one regression line than the other. Marks(1988) echoed the idea that magnitude estimation and production should be employed together, and their results combined if regression were likely to occur, while Poulton(1979) claimed that the effect could be avoided in category judgment by providing anchors or by only using

subjects first judgment. In 1982 he went on to suggest that the use of a wide variety of stimulus ranges could also help to lessen the effect. Apparently methods exist for reducing the regression effect in all forms of judgement.

Thus far Wards'(1987) psychophysical paradigm of examining variability in judgement across methods and subjects has been followed. By examining the details of technique and bias in this way it is possible to find the scaling method best suited to the continuum being measured and to which the most practical bias avoidance strategies can be applied. Before such a decision is made, it is worth considering the recommendations that have been made and the instances in which scaling techniques have been compared empirically.

2.5. Comparison of Techniques.

Although Helson(1964) claims that there is an absence of criteria for determining the validity of one type of scale over another and denies that it is possible to decide upon a best scaling technique, many direct comparisons between scaling methods have been made. In 1965 for example, Eisler stated that ratio scales and thus ratio scaling techniques were preferable to interval scales and scaling techniques on account of the fact that the former involved simpler substantive and measurement theories. Stevens(1971) supported this preference for ratio scaling techniques by claiming that it was only they that measured the actual, as opposed to the virtual, exponent.

Although usually vague about the means by which to select a procedure, recommending that those scales which best measure the attribute of interest should be employed, Stevens and Galanter(1957) said that experimenters should choose a category or a magnitude scale depending on the particular continua being measured. They pointed out, for example, that the Munsel scale was a useful category scale of lightness in practical terms, and that ratio estimates provided useful measures of duration. Category scales, which often support Fechners Law (whereas ratio scales more usually support Stevens Power Law), are affected by hysteresis, stimulus spacing bias (except in category production), response bias and the end effect to such an extent that the biases are hard to eliminate. In ratio scales the most persistent biases are logarithmic bias, regression, the effect of the standard and the effect of number use (except in magnitude production). Although both scaling techniques are effect by bias, ratio scaling techniques may be considered

the most useful for the ratio scales that they produce contain more information than interval scales. Furthermore if it is accepted that Stevens Power Law represents the true psychophysical function then ratio scales are more valid for they more usually support that law.

Of magnitude scales, it is magnitude estimation and magnitude production that have been most frequently compared. Stevens(1956), Kreuger(1984), and Stevens, J. and Mack(1959) found that production techniques agreed with estimation techniques but generally yielded smaller exponents. Stevens, J. and Mack(1959) also found that when measuring force of hand grip ratio production and magnitude production gave similar results - however, Stevens(1956) warned that the reliability of results achieved by a scaling technique should not be interpreted as an indication of its validity (for it is possible that techniques could be reliably biased). In this instance agreement was found between two different techniques. This indicates that, unless both were biased in precisely the same way, each is reasonably valid.

Assuming that a ratio scale of sensation is required, (and ignoring Poultons'(1979) recommendation that most forms of bias could be avoided if each subject was only required to make one judgement, on the grounds that it is highly impractical in terms of subject numbers and experimental duration); this chapter recommends that complementary procedures be employed. Because the free modulus method of magnitude estimation is the technique considered to be least biased (no effects of standard, modulus, response equalising bias or unusual task, and a lessened effect of range), it is suggested that this technique be used together with magnitude production. If the exponents of the two procedures are compared then mid point exponent cancelling another effect, regression, could be deduced. If carefully selected stimuli, iterated stimulus spacing and standardised instructions were also employed then almost all sources of contextual bias and some response bias would be eliminated. If it is impractical to give control of the continua to the subject, as is necessary in production, then is suggested that free modulus magnitude estimation is used alone.

In the case of urgency scaling it is impractical to give control of that parameter to the subject. If subjects had access to the acoustic changes that were expected to result in urgency changes then they would be artificially aware of those changes and more likely to guess the experimental hypothesis. Furthermore the technical difficulties in allowing subjects to change for example, the speed or pitch of a

stimulus in a laboratory are huge. They would not be able to alter the level along a continuum by moving a dial, but would have to re-create stimuli specifying the new parameter levels. They would thus have to know about the levels and measurements of the parameter they were manipulating.
It is therefore recommended that magnitude production is avoided in urgency scaling and free modulus magnitude estimation used alone.

In order that any effects of number-use (including logarithmic bias) be identified it is recommended that a cross modality matching procedure is employed with the above techniques, as a validation procedure. This technique does not assume that subjects can use number to make judgements. In a cross modality matching study, the subject is required to adjust one continua, e.g loudnesss to match the magnitude of a stimulus presented from another continua, e.g. brightness. The slope of the resulting matching function should equal the ratio of the exponents obtained when each continua, brightness and loudness, is scaled independently by a direct method. If the matching function (unaffected by subjects' number use) and the ratio of the exponents obtained by combining magnitude estimation and magnitude production on each of the continua (which is free from the biases previously mentioned) match, then the scales are validated for virtually all forms of bias have been eliminated.

Cross modality matches have been used in this way by for example, Stevens(1971), Marks(1974) and Fucci, Petrosino, Harris and Randolf-Tyler(1988). Fucci et al (1988) concluded explicitly that cross modality matching provided an unbiased verification of magnitude estimation. Variations on the procedure, such as Stevens(1971) ratio matching (whereby subjects adjusted one of a pair of stimuli from the same modality so that the ratio between them matched the ratio of a pair of stimuli presented from another modality); or Stevens, J. and Marks(1980) magnitude matching (whereby alternative presentations of two modalities are estimated on a common scale to produce pairs of stimulus values that match), are equally advantageous.

In conclusion, it is felt that if magnitude estimation and production, or magnitude estimation alone, are used to scale a continuum, and if the mid point exponent is validated by cross modality matching, then that exponent can be considered stable, and not an artifact of bias. If such exponents can be achieved then critics of Stevens Power Law will no longer be able to argue that exponent variability casts

doubt upon the psychophysical function (for such variability will have been shown to be due to biases in the scaling techniques), and it
will be possible to create meaningful ratio scales of sensation. Furthermore ratio scales will be established as the preferred scaling technique for they will reflect the underlying psychophysical law. That is not to recommend that Stevens Power law should be accepted as *the* psychophysical law just because ratio scaling data fit it. It seems better to conceive of the power function as a descriptive device for a set of data, as has previously been suggested, by for example Weiss(1981) and Mcbride(1983c), rather than the definitive psychophysical law. It seems that the conditions under which the law are supported are too specific for it to be considered a general psychophysical law. We are thus following in the footsteps of, for example, Poulton(1989) and Schneider(1989) both of whom said that more than one psychophysical law might be capable of describing sensory experience. The search for an unbiased method of sensory scaling has not been in vain even if the objective is not to uncover the true nature of the psychophysical law. Even if the power function is only used to describe data it is still important to ensure that the description is as valid as possible. The proposed procedures should ensure that is the case.

CHAPTER THREE.

A COMPARISON OF DIFFERENT TECHNIQUES FOR SCALING PERCEIVED URGENCY.

3.1. Introduction.

In this chapter four experiments are reported. In three studies, different scaling techniques were used to measure perceived urgency. In the fourth, the exponents measured by these techniques were validated by cross modality matching.

In order to place alarms in order of priority in terms of their urgency it is important to investigate the effects of different sound parameters upon perceived urgency. We thus need to vary the sound parameter of interest and to scale the perceived urgency. A previous attempt to scale sounds in terms of subjective criteria was made by Fidell and Teffeteller(1981). Their subjects used a five point category scale to judge the annoyance of intrusive sounds. They did not however make any attempt to ensure that their chosen scaling technique could provide a valid sensory scale. It was selected without acknowledged reference to the scaling literature. Given that in the case of perceived urgency it could, in an applied setting, be necessary to place up to eight alarms in order of priority, it is important that the scaling technique employed to scale perceived urgency is suitable for the continua and is valid.

The study of the literature (Chapter Two) showed that magnitude estimation and production used together and validated by cross-modality matching probably represented the least biased means of sensory scaling. It was acknowledged that to some extent, the choice of scaling method should depend upon the continuum under investigation (Stevens and Galanter 1957); and for this reason it was decided not merely to accept the recommendation of the previous chapter but to compare different scaling methods experimentally so that the most advantageous and valid technique for measuring perceived urgency could be revealed, and then employed in future studies. Three experiments were conducted to compare different methods of scaling perceived urgency. They were compared in terms of their ability to scale

the same set of stimuli, the existence of a practice effect, the fit of their data to the power function and the validity of the power function exponent. The latter point was evaluated by a fourth experiment that used cross modality matching to validate the urgency exponents.

It is necessary to find a technique capable of measuring perceived urgency in particular because perceived urgency is different from the continua usually measured in psychophysical studies. In typical studies, what is manipulated by the experimenter, for example brightness, is what the subject is asked to scale, 'How bright is this?' When scaling perceived urgency however, the experimenter manipulates for example stimulus speed but does not ask 'How fast is this?', but, 'How urgent is this?' Thus it is not speed that is measured directly, but its affect upon perceived urgency. Similar studies were conducted by Hellman and Zwicker(1990) when they varied loudness and scaled annoyance. Unless experimentation is conducted we have no way of knowing that the recommended techniques for 'first order' scaling are applicable to 'second-order' scaling such as this.

It was decided to compare two magnitude estimation procedures because they are convenient to administer and produce an informative ratio scale. As previously recommended 'free modulus' magnitude estimation was used (Experiment 1), as, to facilitate comparison, was its opposite, the most restricted form of magnitude estimation (Experiment 2). Despite the risk of regression, magnitude production was not employed because it was too difficult to give control of perceived urgency to the subject. It was felt that since the present studies are not being used to create a perceived urgency scale, only for the purposes of comparison, the presence of a regression effect is not serious enough to justify the practical problems that using magnitude production to avoid it would entail. (For discussion of the problems see Chapter 2, Sections 2.3.1 & 2.5).

An interval scaling technique, category estimation (Experiment 3), was also employed. If urgency is a Prothetic Continuum (quantitative), and if the Prothetic/Metathetic division applies in instances of 'second-order' scaling then this technique is not expected to scale urgency successfully for Stevens(1957) claimed that interval scaling techniques could only measure Metathetic Continua (qualitative). It was employed so that the results could be compared to those of the magnitude estimation studies to tell us something about the nature of perceived urgency in terms of Stevens' continua divisions. Furthermore, it was felt that if

the category estimation technique was found to be a satisfactory measure of perceived urgency, then its simplicity would make it a valuable device to use when interval scales of urgency were sufficient.

A cross-modality matching study was employed in Experiment 4 so that the exponents measured by the three previous techniques could be validated as Marks(1974) suggested by the slope of the theoretically unbiased cross modality matching function (see Chapter Two, Section 2.5). The ratio between the exponent for urgency (as measured by each technique) and the exponent of the matching parameter should equal the slope of the matching function. The most accurate urgency exponent should produce the ratio closest to the slope of the matching function.

The stimuli employed in Experiments 1-4 communicated increases in perceived urgency through increases in stimulus speed as suggested by Patterson(1982) and validated by Edworthy, Loxley, Geelhoed and Dennis(1988); Hellier(1988) and Edworthy, Loxley and Hellier(1989). On the basis of this work it was predicted that faster stimuli would be perceived as being more urgent.

3.2. Experiment One : Scaling Perceived Urgency by Free Modulus Magnitude Estimation.

3.2.1. Introduction.

This experiment employed the 'free-modulus' method of magnitude estimation which Engen(1971) said was the best way obtaining subjective responses to stimuli. This direct ratio scaling technique allowed subjects to respond using any numbers that they chose, and did not require the presentation of a standard stimulus or a modulus. Engen(1971) said that the technique imposed the fewest possible number of restrictions on the subject, he claimed that this prevented data from being biased by the experimenters' choice of response system. The technique was also recommended by Gescheider(1990) for eliminating context effects. It was expected that this scaling technique might eliminate biases in the data caused by a restricted response range, a standard, a modulus and by requiring subjects to perform an unusual task.

3.2.2. Method.

3.2.2.1. Subjects.

Seven male and five female subjects were paid £1 each for volunteering to participate in this study. Subjects were undergraduate or postgraduate students in Psychology or Transport at Polytechnic South West, their ages ranged from 18-35 years. Six of the subjects had previously participated in similar psychophysical studies, none reported having present or previous hearing problems.

3.2.2.2. Materials.

Two adjoining laboratories were made available for the duration of the study.

In Laboratory One was a Tandon PCA20 microcomputer linked to a Cambridge Electronic Design 1401 interface and 1701 low-pass filters set to a cut off of 4 kHz. The Tandon had previously been used to generate, and now stored, the experimental pulse and bursts. The components of these stimuli are shown in Appendix 3A. The bursts were approximately the same length, they varied in terms of the number of pulses they contained, their pulse rate. Stimuli with a higher pulse rate contained more pulses per unit time and were faster. Pulse rate was measured by dividing the maximum stimulus length (2500 ms.) by the pulse-pulse time (ms. from the start of one pulse to the start of the next). In the present Experiments, because all the stimuli were approximately the same length, it would have been possible to measure pulse rate by just counting the pulses in each stimulus. However such a measurement would not have been applicable to all stimuli, only to those of the same length. The former, more generalisable method of measuring speed was therefore adopted. The Experimenter sat in this laboratory during the experiment, she sent stimuli from the Tandon to the subject in Laboratory Two.

Subjects sat at a desk in Laboratory Two, under a Marantz speaker through which they heard the experimental stimuli. They were given type written instructions, a response sheet and pencil.

3.2.2.3. Procedure.

Subjects were run one at a time. They were told the broad nature of the study and were asked to read the following instructions adapted from Engen (1971);

"I am going to present you, in irregular order, a series of sounds. Your task is to tell me how urgent they are by assigning numbers to them. When you have heard the first sound, give its urgency a number - any number that you think appropriate. I will then present another to which you will also give a number, and a third etc. Let high numbers represent high urgency and low numbers represent low urgency. Try to make the ratios between the numbers that you assign to the different sounds correspond to the ratios between the urgency of the sounds. In other words try to make the numbers proportional to the urgency of the sound as you hear it. Remember that you can assign any number. There is no limit to the number that you assign. There is no right or wrong answer. I want to know how you judge the urgency of the sounds.
Any questions?"

When subjects were ready to begin, the Experimenter sent the first stimulus from the Tandon to Laboratory Two. When the subject indicated that he or she was ready the next stimulus was sent, and so on. Bursts 1-7 were played eight times each, in a different random order to each subject. Multiple stimulus presentations were employed so that a practice effect, if it existed, could be identified. As recommended by Engen(1971) three different stimuli, bursts A,B and C, were presented (in the same order to all subjects) on the first three trials. These trials were considered practice trials and the data thereof discarded. Subjects made 59 judgments in all.

When subjects had completed the task they were asked to comment on the study and these comments were recorded. They were debriefed, paid, and allowed to leave.

3.2.3. Results.

Subjects ranked the stimuli in the order predicted from previous work, Edworthy, Loxley, Geelhoed and Dennis(1988); Hellier(1988) and Edworthy, Loxley and Hellier(1989), with the faster stimuli being judged more urgent after the first two, and after all eight of their judgments, (Tables 3.1 & 3.2).

STIMULUS.	3.69	4.98	5.71	7.87	9.63	10	11.93
SUBJECT.		(pulse rate).					
1	32	27.5	23	48.5	64.5	56.5	80.5
2	4.5	6	7	9	10	10.5	12.5
3	2	3	3	4.5	5	5.5	7
4	1.5	2	3.5	4.5	6.5	7.5	11
5	16.5	17.5	23.5	31.5	39	46	60
6	17.5	17.5	30	37.5	45	50	75
7	1.5	6	5.5	7	10.5	10	12.5
8	12.5	22.5	22.5	55	70	77.5	82.5
9	1.5	2.5	2.5	4	7.5	7.5	12.5
10	6.5	8	17.5	12.5	30	27.5	42.5
11	3	25	12.5	30	60	50	85
12	16.5	16.5	16.5	25	37.5	45	45
MEAN	9.6	12.8	13.9	22.4	32.1	32.5	43.8
ST.DEV	9.2	9.5	18.1	24.1	24.7	31.8	32.7

TABLE 3.1 : MEAN OF FIRST TWO JUDGMENTS OF EACH STIMULUS.

STIMULUS / SUBJECT.	3.69	4.98	5.71	7.87	9.63	10	11.93
		(pulse rate).					
1	27.6	36.6	39.1	50.8	65.8	65.6	83.5
2	4.5	6.1	7.25	8.5	9.8	10.1	12.2
3	2.5	3.37	3.37	5	5.8	6	6.87
4	2.3	3.06	3.75	5.25	6.5	6.87	11.56
5	12.3	12.5	19.8	27.1	38.7	37.7	54.6
6	28.1	31.2	37.5	46.5	49.5	50.6	68.1
7	2.12	4.25	4.5	6	9.62	9.25	11.12
8	13.7	26.2	33.7	61.8	71.8	75.6	85.6
9	1.5	2.25	3.75	4.62	10.7	9.25	15.8
10	6.37	9.75	14.7	15.1	23.5	22.5	34.8
11	2.62	20.6	14.4	38.7	53.7	57.5	91.2
12	15.2	18.2	19.7	30.6	39.3	46.2	48.7
MEAN	9.9	14.4	16.8	25	32	33.1	43.6
ST.DEV	9.68	12.1	13.4	20.5	24.2	25.5	32.5

TABLE 3.2 : THE MEAN OF EIGHT JUDGMENTS OF EACH STIMULUS.

The mean of each subjects first two judgments to each stimulus was regressed against the mean of their last six judgments to each stimulus, to produce a regression equation for each subject, (Appendix 3B). If the two judgments were identical then we would expect the regression line to be a straight line passing through zero. Thus the intercept (a) would be 0, and the slope of the line (b) would be 1. As shown in Appendix 3B, two t-Tests were performed. The first t-Test was performed upon the intercepts of the regression equations, against a mean of zero. It showed that these components did not differ significantly from 0, ($t= -1.33$, $p=0.21$). A second t-Test, against a mean of 1, was performed on the slopes of the regression equations. These did not differ significantly from 1, ($t= 1.22$, $p=0.25$.). These findings indicated that the mean of subjects' judgments after two presentations of each stimulus did not differ significantly from the mean of their judgments after the last six presentations of each stimulus.

Examination of each subjects' standard deviation of judgement, Table 3.3, showed that all but subjects Four, Six and Ten judged the stimuli in the middle of the urgency range least consistently (the highest standard deviations were found here). In order to see if subjects became any more or less consistent with repeated stimulus presentations, a two way randomised block factorial ANOVA (stimulus by number of judgements) was performed upon each subjects standard deviation of judgement after the first two and after the last six presentations of each stimulus (see Table 3.4). There was a significant effect of number of judgments upon standard deviation, ($F(1,11)=5.79, p=0.035$); but no significant interaction, ($F(6,66)=0.94, p=0.472$). Examination of the raw data indicated that subjects' judgments became less consistent with repeated presentations of the stimuli, their standard deviation of judgement were higher after eight than after two stimulus presentations.

Having investigated the effects of repeated stimulus presentations, the remaining analysis was performed upon subjects first two judgments to each stimulus, as is usual psychophysical procedure (Engen 1971). Before the data was fitted to Stevens(1957) Power Function, Engen's(1971) logarithmic transformation to eliminate inter- and intra- subject variability was performed on each subjects' first two judgments of each stimulus, (Table 3.5).

STIMULUS / SUBJECT	3.69	4.98	5.71	7.87	9.63	10	11.93
		(pulse rate).					
1	10.2	17	20.9	15.4	10.2	10.9	4.3
2	0.53	0.99	1.03	0.75	0.64	0.64	0.70
3	0.53	0.51	0.70	0.75	0.64	0.53	0.35
4	1.21	1.01	1.03	1.19	1.85	1.24	2.44
5	2.82	3.11	6.75	5.59	8.24	9.39	5.95
6	15.3	13.5	10.3	11.9	9.83	10.8	4.58
7	1.24	1.66	1.41	1.60	1.50	0.88	3.36
8	4.43	10.9	12.4	10.6	8.43	4.17	3.20
9	0.53	0.88	1.58	1.99	3.02	3.20	2.47
10	3.1	3.20	4.56	4.09	5.76	5.1	5.94
11	3.29	10.5	7.76	12.4	22	13.8	17.2
12	2.95	2.18	4.68	9.04	7.76	4.4	3.54

TABLE 3.3 : INDIVIDUALS STANDARD DEVIATION OF EIGHT JUDGMENTS.

SOURCE	SUM OF SQU	df	MEAN SQ	F	P
Subjects	2743.102	11	249.372		
Stimulus	178.587	6	29.764	1.34	0.248
Stimulus* Subj.	1455.872	66	22.058		
Judgement	49.519	1	49.519	5.79	0.035
Judgement * Subj.	93.928	11	8.538		
Stim.*Judg -ement	62.024	6	10.337	0.94	0.472
Stim.*Judg -ement* Subject	725.45	66	10.991		

TABLE 3.4 : TWO WAY ANOVA, STIMULUS BY NUMBER OF JUDGEMENTS

STIMULUS	3.69	4.98 (pulse rate)	5.71	7.87	9.63	10	11.93
MEAN LOG	0.75	0.94	0.99	1.16	1.34	1.35	1.48
ANTI LOG	5.62	8.82	9.65	15.58	21.88	22.38	30.84

TABLE 3.5 : MEAN JUDGMENTS TO EACH STIMULUS AFTER TRANSFORMATION.

As suggested by Engen(1971) the method of least squares was used to fit the transformed data to Stevens Power Function (see Appendix 3C).

In fitting the data to the power function, speed (pulse rate) was used as the objective stimulus measure and subjects' judgements were used as the subjective stimulus measure.

The resulting Stevens Power Function for perceived urgency took the form,

Perceived urgency =

1.170*pulse rate^1.43

(Equation 3.1)

A linear regression of the log. subjective stimulus values against the log. objective stimulus values was performed to test the goodness of fit of the data to Stevens' Power Law. 99.2% of the variance was accounted for by a straight line, which represents a good fit. The plot of Stevens Power Function is shown in Fig. 3.1. When the same data was plotted and regressed in linear co-ordinates 98.3% of the variance was accounted for by a straight line.

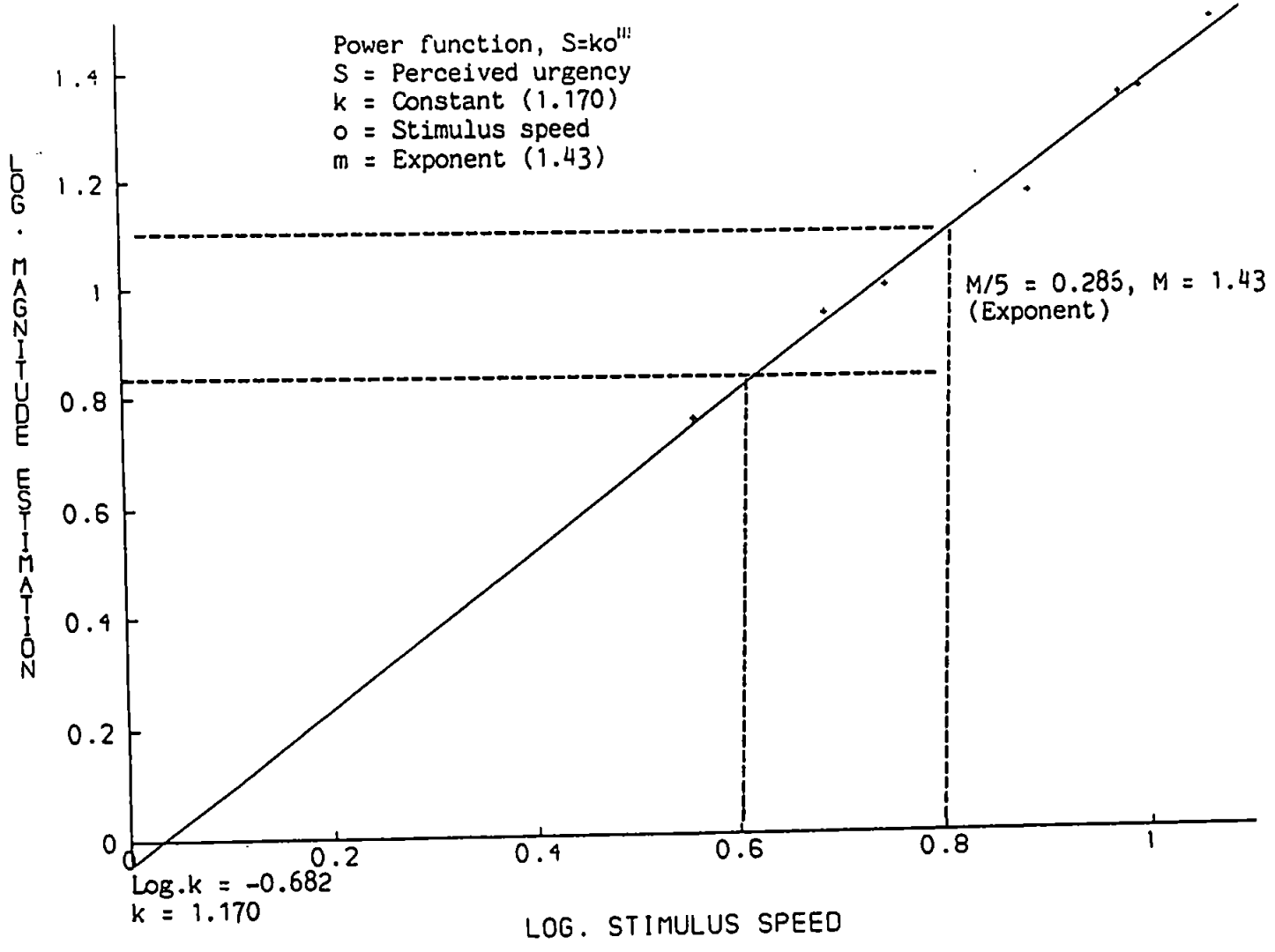
3.2.4. Discussion.

Although subjects' comments indicated that they found the task hard, the stimuli were ranked in the order predicted from previous research, with faster stimuli being perceived as more urgent. Furthermore, although they said that they wanted the limits of their responses more restricted and more guidance on the range of numbers that it was reasonable for them to use, all subjects used whole numbers between 1 and 100. This finding may indicate that in the 'Free Modulus' method of magnitude estimation subjects' response range, although not restricted by the experimenter, may remain restricted by the subjects own expectations about acceptable numbers to use.

the last six

Although the regression of mean judgments after two and eight trials indicated that mean judgments to the stimuli were not significantly different after the first two trials compared to after all eight, the Anova on individuals standard deviations after the first two and after the last six judgments indicated that subjects judgments became less consistent with repeated stimulus presentations. This effect can be considered to be the effects of fatigue or

Fig. 3.1 : Stevens Power Function (Experiment 1).



confusion, and provides a rationale for adhering to the normal psychophysical procedure of using subjects first two judgments only (Engen 1971).

The mean standard deviation of judgement to each stimulus was high between subjects because they were able to choose different moduli and response ranges (Table 3.1, 3.2). Some subjects appear to have used a response range 1-15, whereas others used 1-100. The group's standard deviations were lowest for low urgency stimuli because then, subjects were using similar numbers which ever response range they adopted. For high urgency stimuli, the responses of subjects using the different response ranges became more divergent, and so the standard deviation of judgement between subjects was higher. Despite the lack of experimenter induced restriction, logarithmic bias seems to have been exhibited by those subjects that used the response range 1-100. The remaining five subjects did not exhibit logarithmic bias in their responding.

The data fitted Stevens' Power Law well. The log.-log. plot indicated that perceived urgency was related to stimulus speed by an exponent of 1.430. The log.-log. plot was well fitted by a straight line (99.2% of the variance accounted for). Although a regression showed that the linear plot of the data could also be fitted by a straight line with 98.3% of the variance accounted for, Stevens Power Function remains the best fitting line, as a higher percentage of the variance was accounted for.

3.3. Experiment Two : Scaling Perceived Urgency By Fixed Modulus Magnitude Estimation.

3.3.1. Introduction.

A variation of the 'fixed modulus' method of magnitude estimation used by Pradham and Hoffman(1963) was used in this study. Besides setting the standard and modulus as Pradham and Hoffman(1963) had done, the Experimenter also restricted the range of numbers that subjects could use in their responses. Although more usually imposed in category rating studies, number limits were used here so that the present magnitude estimation technique, with a prescribed standard modulus and response range, would represent the most restricted form of the task. The aim was to compare the results of this study with those of

Experiment One, which employed the least restricted magnitude estimation procedure.

Although previous authors have suggested that subjects judgments can be affected by the standard and the modulus (Poulton 1979), and by Response Equalizing bias which can be caused by restrictions on subjects response range (Poulton et al 1980), this potentially biased magnitude estimation technique was selected, not only for the purpose of comparison, but also because subjects in Experiment One stated that a more restricted task would make judgement easier. Furthermore, it is not known whether such biases affect second order scaling in the same way that they do first, or whether the effects are of the same magnitude.

3.3.2. Method.

3.3.2.1. Subjects.

Nine male and three female subjects were paid £1 for volunteering to participate in this study. All subjects were undergraduate students in Psychology or Engineering at Polytechnic South West, their ages ranged from 20-35 years. Eight subjects had previously participated in similar psychophysical studies, none reported having present or a history of hearing problems.

3.3.2.2. Materials.

The materials employed were identical to those used in Experiment One.

3.3.2.3. Procedure.

The procedure differed from Experiment One only in the specified ways.

Subjects read the following instructions adapted from Pradham and Hoffman(1963);

"I am going to present you with a series of sounds and your task is to estimate their urgency. You will do this by assigning a number to each of them proportionate to its urgency. You can use any number between 1-100, decimal, fraction or whole number - the only restriction being that the number you use should be proportional to the urgency of the sound; that is, if the sound appears twice as

urgent as the standard say 100, if half say 25, if one hundredth say 0.5, and so on. The standard tone will be played before each stimulus on every trial, and shall be called 50. Estimate the urgency of the second sound that you hear on each trial in relation to the first (the standard). Do not try to be consistent. Make your judgments independent of what you have done in the past. Every time, compare the given sound with the standard and write down the number proportionate to its urgency.
Any questions?"

Burst 6 (pulse rate = 7.87) was selected as the standard stimulus since it was in the middle of the stimulus range, as suggested by Pradham and Hoffman(1963).

When subjects were ready, the Experimenter sent the first stimulus pair, (the standard and the experimental stimulus separated by a one second gap), from the Tandon to Laboratory Two. Bursts 1-7 were played, preceded by burst 6, eight times each, in a different random order to each subject. Although not part of Pradham and Hoffmans' (1963) procedure, practice trials were employed in this experiment because it was felt that they helped to familiarise subjects with the task. They were identical to those used in Experiment One except in this instance each of the practice stimuli was preceded by the standard.

3.3.3. Results.

Subjects ranked the stimuli in the order predicted from previous work, Edworthy et al(1988), Hellier(1988) and Edworthy et al(1989) with faster stimuli being perceived as more urgent after the first two judgments of each stimulus and after all eight, see Tables 3.6 and 3.7.

As in Experiment One the mean of each subjects first two judgments of each stimulus was regressed against the mean of each subjects last six judgments, to produce a regression equation for each subject, and t Tests were performed upon the intercepts and slopes of the equations, (Appendix 3D). The intercepts did not differ significantly from 0, ($t=-0.73$, $p=0.48$); and the slopes did not differ significantly from 1, ($t= 0.63$, $p=0.54$). The findings indicated that the

STIMULUS. SUBJECTS	3.69	4.98 (pulse rate).	5.71	7.87	9.63	10	11.93
1	15	12.5	5.25	50	100	100	100
2	26	36.5	40	52.5	69	67.5	82.5
3	4	5	6	10	11	11	17.5
4	20	25	37.5	55	57.5	57.5	75
5	20	20	35	50	65	70	72.5
6	22.5	35	37.5	47.5	70	80	95
7	17.5	25	27.5	49.5	55	56.5	65
8	27.5	25	32.5	51	60	60	76
9	17.5	30	32.5	42.5	70	77.5	95
10	15	22.5	27.5	50	77.5	80	100
11	25	25	35	50	60	62.5	100
12	25	22.5	30	65	55	65	100
MEAN	19.5	23.6	28.8	47.7	62.5	65.6	81.84
ST.DEV	6.50	8.68	11.5	13	20.43	21.1	23.8

TABLE 3.6 : MEAN OF FIRST TWO JUDGEMENTS OF EACH STIMULUS.

STIMULUS / SUBJECTS	3.69	4.98	5.71 (pulse rate).	7.87	9.63	10	1.93
1	9.56	13.18	19.4	50	82.5	91.2	98.75
2	29.5	36.25	39.37	53.62	66.65	71.2	78.87
3	4.7	6.25	7.8	10	11.6	13.2	18.75
4	15.6	25	28.5	51.2	61.2	65	78.12
5	15.6	21.25	28.7	50.6	56.8	62.8	74.37
6	31.2	37.8	41	49.5	65.1	69.3	89.37
7	16.2	27.5	35.6	49.7	51.5	57	65.62
8	21.2	28.7	38	51.1	60.6	65.7	85.25
9	17.5	30	35.6	46.2	66.8	73.7	93.12
10	16.8	23.7	27.5	47.5	71.7	75	100
11	25	25	32.2	50	58.1	60.6	86.25
12	21.8	26.2	34.3	64.3	55	64.3	90.62
MEAN	18.7	25	30.6	47.8	58.9	64	79.92
ST.DEV	7.6	18.75	9.3	12.7	17	18.3	21.71

TABLE 3.7 : MEAN OF SUBJECTS EIGHT JUDGEMENTS OF EACH STIMULUS.

mean of subjects judgments after two presentations of each stimulus did not differ significantly from the mean of their judgments after the last six presentations.

Examination of each subjects standard deviation of judgement, Table 3.8, showed that presentations of Burst Six (pulse rate 7.87), the standard, resulted in the lowest standard deviation of judgement, for all bar subjects Two, Five and Ten. This result was expected because subjects were told what number to assign the standard stimulus, and they had heard it on every trial.

A two-way Anova,(stimulus by number of judgements), was performed on each subjects standard deviation of judgement after the first two and after the last six presentations of each stimulus, as shown in Table 3.9. No significant differences between the figures were found ($F(1,11)=3.5, p=0.087$), and no significant interaction ($F(6,66)=0.86, p=0.527$). This shows that subjects' judgments did not become any more or less consistent with repeated presentations of the stimuli. There was an effect of stimulus on standard deviation ($F(6,66)=2.61$, $p=0.025$).

The remaining analysis was conducted upon subjects first two judgments of each stimulus. Engen's(1971) data transformation to eliminate inter- and intra-subject variability was conducted (Table 3.10). As suggested by Engen(1971) the method of least squares was used to fit the transformed data to Stevens Power Function, (see Appendix 3E). The resulting Stevens Power Function for perceived urgency took the form,

$$\text{Perceived urgency} = 2.985 * \text{pulse rate}^{1.34} \qquad \text{(Equation 3.2)}$$

A linear regression of the log. subjective stimulus values against the log. objective stimulus values was performed to test the goodness of fit of the data to Stevens Power Function. 96.4% of the variance was accounted for by a straight line. The plot of Stevens Power Function is shown in Fig.3.2. When the same data was plotted in linear co-ordinates regression showed that 98.1% of the variance was accounted for by a straight line.

STIMULUS / SUBJECTS	3.69	4.98	5.71	7.87	9.63	10	11.93
		(pulse rate).					
1	9.21	8.31	10.7	0	13.3	7.9	13.54
2	3.59	3.01	1.40	6.70	5.28	5.39	5.89
3	0.70	1.38	1.95	0	1.68	1.98	2.31
4	4.96	6.55	9.23	3.54	3.54	5.98	4.58
5	6.23	6.94	7.44	4.17	17.31	8.43	3.20
6	9.16	6.24	6.82	2.13	13.36	12.08	7.76
7	3.54	7.56	7.76	0.88	4.34	2.50	4.96
8	8.35	9.16	11.1	1.80	5.06	6.78	7.46
9	3.78	7.56	4.96	3.20	3.72	5.18	4.58
10	7.53	9.16	4.63	7.07	7.53	6.55	0
11	0	4.63	6.54	0	5.30	6.78	7.07
12	3.72	4.96	4.17	6.05	7.29	9.43	
	6.94						

TABLE 3.8 : INDIVIDUALS STANDARD DEVIATION OF EIGHT JUDGEMENTS.

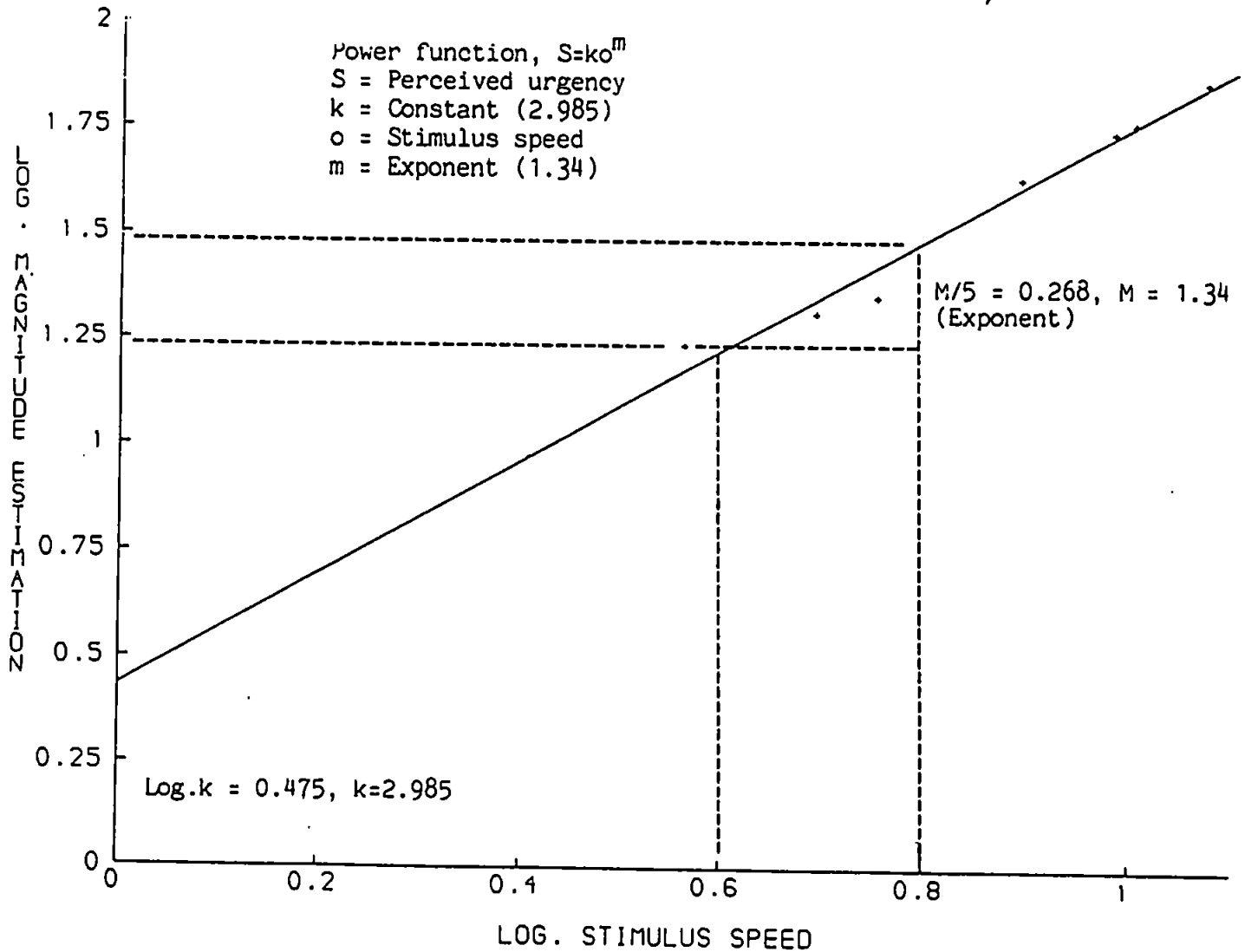
SOURCE	SUM OF SQU	df	MEAN SQ	F	P
Subjects	510.579	11	46.416		
Stimulus	238.295	6	39.715	2.61	0.025
Stimulus* Subj.	1004.193	66	15.215		
Judgement	82.194	1	82.194	3.53	0.087
Judgement* Subj.	255.92	11	23.265		
Stim.*Judgement	36.837	6	6.139	0.862	0.527
Stim.*Judgement* Subject	469.91	66	7.119		

TABLE 3.9 : TWO WAY ANOVA, STIMULUS BY NUMBER OF JUDGEMENTS

STIMULUS	3.69	4.98 (pulse rate).	5.71	7.87	9.63	10	11.93
LOG	1.24	1.32	1.36	1.64	1.75	1.77	1.87
ANTILOG	17.70	21.21	23.17	44.43	56.84	59.55	75

TABLE 3.10 : THE MEAN JUDGEMENT AFTER TRANSFORMATION.

Fig. 3.2 : Stevens Power Function (Experiment 2).



3.3.4. Discussion.

The stimuli were ranked in the order predicted from previous research, with faster stimuli being perceived as more urgent. Subjects' comments indicated that they found the task easy to perform, because their response range was more restricted. Subject Three however stated that the task was easy, yet appeared not to understand it - he consistently ranked the standard 10 instead of 50, and used the numbers 1-20 instead if 1-100. His data was not excluded from analysis because it may reflect the fact that for some subjects the task might prove hard to understand.

There is further evidence that the task was not performed properly. The standard deviation for judgments of stimulus pairs when Burst 6 appeared twice (as the standard and as the experimental stimulus) was only 0 for four of the twelve subjects. If, as should have happened, the standard was always recognised as such and ranked 50 the standard deviation of judgement on these trials would have been 0 for all subjects. If subjects perform differently on the same task (through for example variations in their understanding of instructions) then a valid scale of sensation cannot be constructed using that task.

The Anova on standard deviations after the first two and after the last six judgments and the regression of the mean judgments after the first two and after all six presentations of each stimulus indicated that no effects of practice, learning or fatigue existed. The significant effect of stimulus upon standard deviation can be accounted for by the lowered standard deviation associated with the standard stimulus.

The mean standard deviation of judgement between subjects was again highest for the high urgency stimuli. However it was not lowest for the standard stimulus, as would be expected if subjects had recognised it, but for the least urgent stimulus. Standard deviation between subjects were not quite as high as in Experiment One, perhaps because in this instance subjects response range was explicitly specified and restricted.

The data fitted Stevens Power Function well. The log.-log. plot indicated that perceived urgency was related to stimulus speed by an exponent of 1.340. The log.-log. plot was well fitted by a straight line (96.4% of the variance accounted

for) as Stevens said it should be if the power function fitted. A linear regression accounted for a higher percentage of the variance than did the Power Function (98.1%). This implied that the data was slightly better fitted by a straight line in linear co-ordinates than by Stevens Power Function.

It is interesting to note that although subjects preferred performing the 'fixed modulus' magnitude estimation task, and found it easier they did not follow the instructions as well as in the 'free modulus' task (which also produced data that better fitted the Power Law).

3.4. Experiment Three : Scaling Perceived Urgency by Category Estimation.

3.4.1. Introduction.

A variation of the category estimation methods used by Ward(1972) and by Curtis(1970) was used in this study. Although category estimation procedures have been highly criticised, especially with respect to the measurement of Prothetic continua (Stevens 1957); the present study employed the procedure so that such criticisms could be validated in relation to scaling perceived urgency and so that interval and ratio scales of the same stimuli could be compared. Subjects were shown the most extreme stimuli at the beginning of the study so that the regression affect could be avoided. This procedure was suggested by Stevens and Galanter (1957).

3.4.2. Method.

3.4.2.1. Subjects.

Five male and seven female subjects were paid £1 each for volunteering to participate in the study. All were undergraduate or postgraduate students in Psychology or Engineering at Polytechnic South West, their ages ranged from 18-26 years. Six of the subjects had previously participated in similar psychophysical studies, none reported present or a history of hearing problems.

3.4.2.2. Materials.

The materials employed in this study were identical to those used in Experiment One.

3.4.2.3. Procedure.

The procedure only differed from that used in Experiment One in the ways stated. Subjects were asked to read the following instructions, adapted from Curtis(1970) and Ward(1972);

"In this experiment (category judgments of urgency), I would like you to judge the urgency of some sounds that will be presented to you. Try to divide the range between the most and least urgent stimulus into ten equal intervals or categories numbered 1-10. If it seems very non-urgent give it a number like 1 or 2 etc. In a moment I will show you the least urgent stimulus which I would like you to call 1, then I will show you the most urgent stimulus which I would like you to call 10. On subsequent judgments, each sound that is presented should be rated on a scale of 1-10. There are no right or wrong answers, I am only interested in your subjective impressions of urgency. There is no need to try and be consistent. Any questions? "

Subjects questions were answered and the least urgent stimulus, Burst 1, was played, followed by the most urgent stimulus, Burst 4. It was explained that these were the most extreme stimuli that they would hear. Practice trials which were the same as those in Experiment One were used.

3.4.3. Results.

Subjects responses, which ranged between 1 and 10, were transformed so that they ranged from 1-100, to allow comparison with previous experiments. The transformation implied that if the original ten categories were stretched to a scale 1-100, then each category would be ten digits wide. A judgement from the original category scale was transformed so that it was a number in the middle of the stretched category. Thus an original judgement of category 1 became 5 because that is the number in the middle of the first stretched category,(1-10); 2 became 15, 3 became 25, 4 became 35, etc.

Subjects ranked the stimuli in the order predicted from previous work, Edworthy et al(1988), (Hellier 1988) and Edworthy et al(1989) with faster stimuli being perceived as more urgent, after the first two and after all eight of their judgments (Tables 3.11 and 3.12).

The mean of each subjects first two judgments to each stimulus was regressed against the mean of their last six judgments to each stimulus to produce a regression equation for each subject. As shown in Appendix 3F two t-Tests were performed upon the intercepts and slopes of the regression equations as in Experiment One. The intercepts did not differ significantly from 0, ($t=-1.36, p=0.20$).The slopes did not differ significantly from 1, ($t=1.79, p=0.10$). These findings indicated that the mean of subjects' judgments after two presentations of each stimulus did not differ significantly from the means of their judgments after the last six presentations of each stimulus.

Examination of each subjects' standard deviation of judgement (Table 3.13) showed that all bar subjects 1,2 and 12 judged the stimuli in the middle of the urgency range least consistently, (the highest standard deviations were found here). A two-way Anova (stimulus by number of judgements) was performed upon each subjects standard deviation of judgement after the first two and after the last six presentations of each stimulus(see Table 3.14). There was a significant effect of number of judgments, ($F(1,11)=27.1$, $p=0.000$); but no significant interaction ($F(6,66)=0.47$, $p=0.825$).

Having investigated the effects of repeated stimulus presentations, the remaining analysis was conducted upon each subjects' first two judgments to each stimulus. Engens(1971) data transformation to eliminate inter- and intra- subject variability was performed, (Table 3.15). As suggested by Engen(1971) the method of least squares was used to fit the transformed data to Stevens Power Function, (Appendix 3G). The resulting Stevens Power Function for perceived urgency took the form,

$$\text{Perceived urgency} = 0.469 \cdot \text{pulse rate}^{2.22}$$

(Equation 3.3)

STIMULUS / SUBJECT	3.69	4.98 (pulse rate).	5.71	7.87	9.63	10	11.93
1	5	20	15	50	35	40	60
2	5	10	5	55	75	75	85
3	5	10	5	45	55	55	85
4	5	25	40	65	70	75	95
5	20	30	50	60	75	75	95
6	5	15	45	45	60	60	95
7	5	15	15	50	35	45	70
8	5	15	20	65	80	85	95
9	5	20	50	80	80	80	95
10	5	20	25	30	60	60	95
11	5	15	20	40	65	55	95
12	25	25	50	90	95	95	95
MEAN	7.91	18.3	28.3	56.2	65.4	66.6	88.3
ST.DEV	6.89	6.15	17.6	16.9	17.8	16.6	11.7

TABLE 3.11: MEAN OF FIRST TWO JUDGEMENTS OF EACH STIMULUS.

STIMULUS / SUBJECTS	3.69	4.98 (pulse rate).	5.71	7.87	9.63	10	11.93
1	13.7	20	23.7	41.2	42.5	42.5	55
2	5	8.7	15	61.2	73.7	77.5	91.2
3	6.2	13.7	15	35	50	52.5	71.2
4	8.7	30	37.5	61.2	63.7	62.5	86.2
5	11.2	30	45	56.2	68.7	73.7	93.7
6	8.7	13.7	37.5	33.7	55	65	92.5
7	6.2	20	23.7	51.2	61.2	55	85
8	4.5	15	21.2	43.7	65	63.7	83.7
9	10	35	45	66.2	82.5	82.5	91.2
10	10	22.5	25	46.2	66.2	65	93.7
11	5	13.7	21.2	32.5	51.2	55	90
12	46.2	52.5	58.7	87.7	85	87.5	88.7
MEAN	11.3	22.9	30.7	51.3	63.7	65.2	85.1
ST.DEV	11.3	12.2	13.7	16.1	12.8	13.1	11.3

TABLE 3.12 : MEAN OF SUBJECTS EIGHT JUDGEMENTS.

STIMULUS / SUBJECT	3.69	4.98 (pulse rate)	5.71	7.87	9.63	10	11.93
1	9.91	5.35	8.35	7.44	7.07	4.63	5.35
2	0	7.44	14.1	5.18	9.91	11.6	5.18
3	3.54	9.91	10.6	9.26	13	12.8	10.6
4	5.18	13	13.8	11.8	12.4	20.5	11.2
5	7.44	7.56	14.1	13.5	7.44	8.35	3.54
6	5.18	6.41	10.3	12.4	10.6	11.9	4.63
7	3.54	7.56	13.5	15.9	16.8	23.3	11.9
8	1.41	0	7.44	13.5	14.1	15.5	8.35
9	5.35	15.1	5.35	14.5	10.3	7.07	5.18
10	7.56	10.3	11.9	17.2	15.5	13	3.54
11	0	6.4	7.44	8.88	15.9	9.26	5.35
12	22.3	21.21	8.4	10.3	14.1	7.07	9.16

TABLE 3.13 : INDIVIDUALS STANDARD DEVIATIONS OF EIGHT JUDGEMENTS.

SOURCE	SUM OF SQU	df	MEAN SQ	F	P
Subjects	915.404	11	83.218		
Stimulus	755.45	6	125.90	3.19	0.08
Stimulus*Subj.	2601.75	66	39.42		
Judgement	1039.47	1	1039.47	27.1	0.00
Judgement*Subj.	421.17	11	38.28		
Stim.*Judgement	67.32	6	11.22	0.47	0.825
Stim.*Judgement *Subject	1562.32	66	23.67		

TABLE 3.14 : TWO WAY ANOVA, STIMULUS BY NUMBER OF JUDGEMENTS

STIMULUS	3.69	4.98	5.71	7.87	9.63	10	11.93
		(pulse rate).					
LOG	0.78	1.21	1.33	1.72	1.79	1.80	1.94
ANTILOG	6.13	16.28	21.57	53.67	62.07	64.23	87.45

TABLE 3.15 : THE MEAN JUDGEMENT AFTER TRANSFORMATION.

A linear regression of the log. subjective stimulus values against the log. objective stimulus values was performed to test the goodness of fit of the data to Stevens Power Function. 96.2% of the variance was accounted for by a straight line. The plot of Stevens Power Function is shown in Fig.3.3. When the same data was plotted in linear co-ordinates regression showed that 98.1% of the variance was accounted for by a straight line.

3.4.4. Discussion.

Stimuli were ranked in the predicted order, Edworthy et al (1988), (Hellier(1988)) and Edworthy et al(1989) , with faster stimuli being perceived as more urgent. Subjects reported that the task was easy to perform.

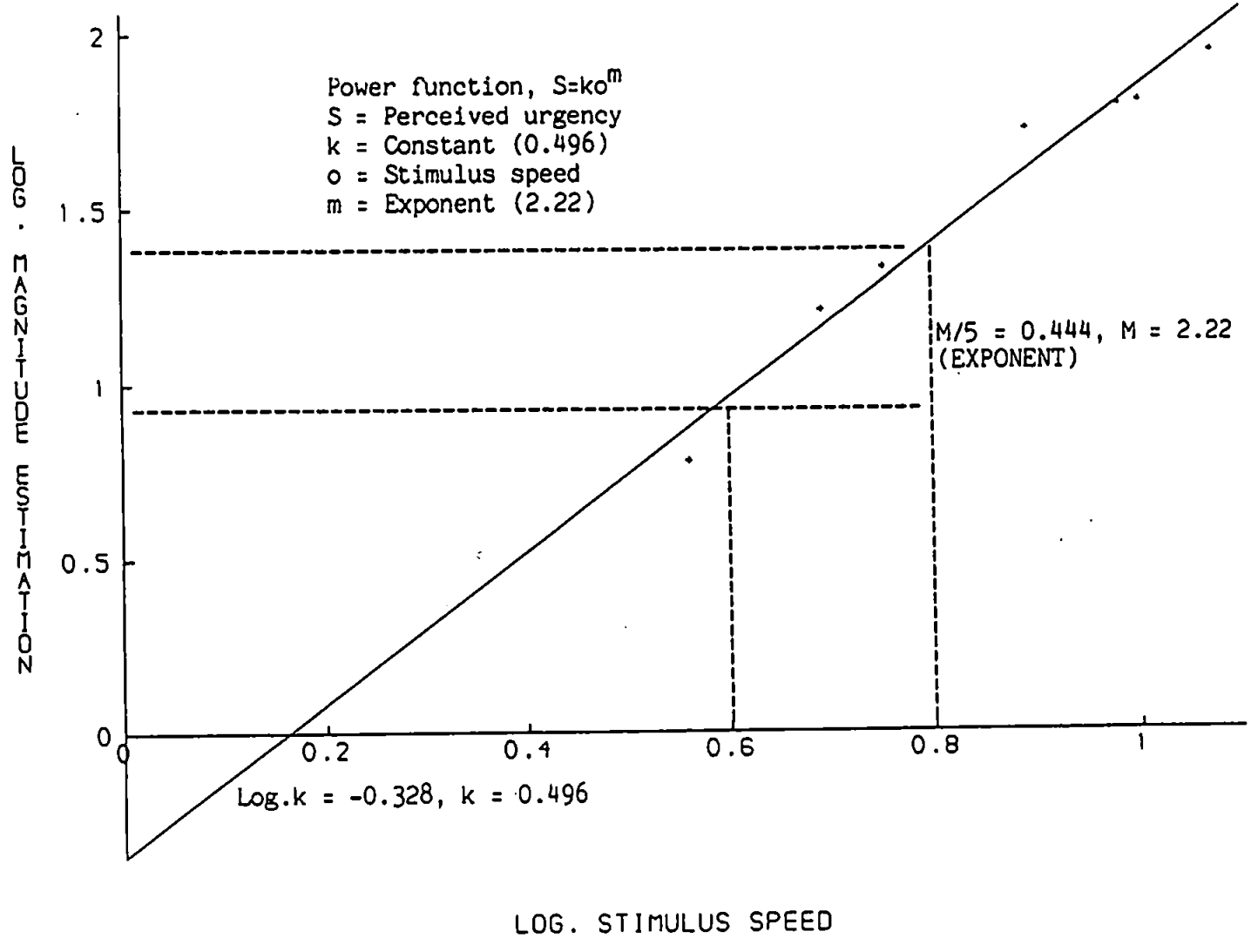
the last six

Although the regression of the mean judgement after two presentations of each stimulus against the mean judgement after eight presentations of each stimulus indicated that the means were the same, the Anova on individual standard deviations after two and after six presentations showed that subjects became less consistent with repeated presentations of the stimuli. This is possibly the result of subjects becoming fatigued, bored or confused with repeated stimulus presentations.

The mean standard deviation of judgement to each stimulus was highest for the mid-range stimuli. Individual standard deviations also tended to be higher for the mid-range stimuli. Standard deviations were higher than were expected considering that subjects originally responded using a much smaller response range than in the magnitude estimation tasks.

The fact that the data fitted the Power Function was not expected, given Stevens'(1971) claim that interval scaling techniques would not support the Power Law. In fact 96.2% of the variance was accounted for by a straight line on the log-log plot. What is more important however is that the Power Function does not represent the best fitting function of the data, 98.1% of the variance was accounted for by a linear plot. This supports Stevens predictions for interval scales.

Fig. 3.3 : Stevens Power Function (Experiment 3).



Another surprise is the size of the exponent, 2.22. Stevens(1971) stated that when category scales did fit the Power Function, the exponents that they produced were 'virtual' exponents, smaller than the 'actual' exponents derived from ratio scaling procedures. In fact the exponent derived from this category estimation technique is much larger than that derived by the magnitude estimation procedures. It is possible that the data treatment in this experiment has given rise to uncharacteristic results. The fault could lie either in my own transformation of the data, or in the application of Engen's(1971) data transformation to category data. In the literature, Engens' transformation has only previously been used on data from ratio scaling procedures. Engens' transformation was conducted so that the data from all the Experiments was treated the same to facilitate comparison.

3.5. Experiment Four : Scaling Perceived Urgency By Cross Modality Matching.

3.5.1. Introduction.

This study employed cross modality matching (between urgency and line length), which can be used to validate the exponents measured by the other scaling techniques. The study was conducted so that, as Stevens(1960) suggested, the slope of the equal sensation function produced by cross modality matching could be used to validate the exponents from the scaling techniques employed in Experiments 1-3. The ratio between the exponent for urgency, (1.43, 1.34 or 2.22 from Experiments 1,2 and 3), and the exponent for line length, (1.1 from Stevens and Galanter 1957) should predict the slope of the matching function.

The cross modality matching procedure employed here required subjects to match line length to perceived urgency. The procedure was similar to that used by Mashhour and Hosman(1968) when they matched line length to noise, grey, texture and weight. Also to that used by Kuwano and Namba(1990) who matched line length to helicopter noise. As Stevens and Galanter(1957) pointed out, line length is a convenient continuum to manipulate because no special apparatus is required, it is a familiar medium and most people are used to making judgments of length. They stated that subjective line length was nearly a linear function of apparent length, for its exponent was 1.1.

Stevens(1969) warned that regression could occur in cross modality matching studies when subjects shortened the range of the variable under their control. He said that this could be avoided if each of the two continua served once as the adjusted variable and once as the criterion stimulus. In this way subjects would, for example, adjust line length to match urgency, and then adjust urgency to match line length. The geometric mean of the exponents from the two matching functions was said to represent the true slope, free from regression. However he also pointed out that regression was minimal in studies involving line length as either the adjustable variable or the criterion stimulus. In view of that fact and the technical difficulty in allowing subjects to manipulate urgency to match length (Chapter Two, Section 2.5), it was not considered necessary to employ this 'balanced design' procedure.

3.5.2. Method.

3.5.2.1. Subjects.

Four male and eight female subjects received one participation point as part of their coursework requirement when they volunteered to participate in this study. All were undergraduate Psychology students at Polytechnic South West. Subjects ages ranged from 18-21 years. Four subjects had previously participated in similar psychophysical studies, none reported present or past hearing problems.

3.5.2.2. Materials.

The laboratory arrangements and stimuli were as reported in Experiment One.

On the desk in Laboratory Two subjects had type-written instructions, a pencil and a response sheet. The response sheet allowed subjects to draw horizontal lines up to a maximum length of 394mm.

3.5.2.3. Procedure.

The procedure differed from Experiment One only in the ways specified. Subjects were asked to read the following instructions, (adapted from Mashour and Hosman 1968);

"This experiment is concerned with your subjective experience of urgency. You will be presented with a series of sounds in random order. You are requested to match the length of a line to the urgency of each sound that I present by drawing a line so that its subjective length is equal to the subjective urgency of sound. Let short lines represent low urgency and longer lines represent high urgency. Do not try to be consistent, it is only your immediate impressions that are of interest. Any questions?"

Subjects were asked to draw the lines as straight as they could, without a ruler. Bursts 1-7 were heard twice each. An effect of repeated presentations was not tested for because cross modality matching was only being employed as a validation procedure, it was not being evaluated as a scaling technique in its own right. In the previous evaluations of scaling techniques, the existence of an effect would have had implications for the worth and administration of the method. The first three Bursts that each subject heard were the practice trials described in Experiment One.

3.5.3. Results.

All scores are line measurements in mm., accurate to 1mm.

As shown in Table 3.16, subjects ranked the stimuli in the order predicted from previous work, Edworthy et al(1988), Hellier (1988) and Edworthy et al(1989), with faster stimuli being perceived as more urgent - they were represented by longer lines.

Given the large response range available to the subjects, mean standard deviations to each stimulus were predictibily high. As in previous experiments, (One and Two), the highest mean standard deviations of judgement were for the most urgent stimuli.

Engen's(1971) data transformation to eliminate inter- and intra- subject variability was applied to subjects judgments (Table 3.17). The method of least squares, Appendix 3H, was used to fit the data to a straight line to reveal the slope of the matching function, 1.35. The matching function is represented graphically in Fig. 3.4. A linear regression was performed to test the goodness of fit of the data to a straight line, when it was plotted in log-log co-ordinates. 98.2% of the variance was accounted for.

STIMULUS	3.69	4.98	5.71	7.87	9.63	10	11.93
SUBJECT		(pulse rate).					
1	87	99.5	94.5	213	224.5	239.5	293.5
2	48	63.5	85	125	153.5	173	279.5
3	9	21.5	15	49.5	76.5	65.5	148
4	32	37	28.5	91.5	123	137.5	146.5
5	44	58.5	48.5	75	88	98.5	141.5
6	69	52	176	162.5	301	324.5	356.5
7	66.5	118.5	115.5	165.5	233	226.5	294
8	101	127	172.5	196.5	200	197.5	290
9	20	22	26	43	57	55	93.5
10	64.5	123.5	167.5	224.5	301.5	314.5	367.5
11	130.5	167.5	168	212.5	273.5	292	370.5
12	52	61	74	101.5	101.5	119	173
MEAN	60.2	79.2	97.5	138.3	177.7	186.6	246.2
ST.DEV	34.3	46.8	61.6	66	89.2	94.3	99.5

TABLE 3.16: MEAN JUDGEMENT BY EACH SUBJECT TO EACH STIMULUS.

STIMULUS	3.69	4.98	5.71	7.87	9.63	10	11.93
		(pulse rate).					
LOG	1.67	1.79	1.83	2.07	2.18	2.21	2.35
ANTILOG	47.3	62.8	68.5	119.2	154.1	162.7	223.8

TABLE 3.17 : THE MEAN JUDGEMENT AFTER TRANSFORMATION.

Experiment	Exponents(Urgency /length)	Ratio of exponents	Matching Function
EXPT1	1.430/1.1	1.3	1.35
EXPT2	1.340/1.1	1.21	1.35
EXPT3	2.22/1.1	2.01	1.35

TABLE 3.18 : VALIDATION OF URGENCY EXPONENTS.

add

The slope of the matching function was used to validate the exponents measured in Experiments 1-3. Table 3.18 shows the extent to which the exponents predicted the matching function. Given that the line length exponent is well established in the literature, eg. Stevens and Galanter(1957), and that the matching function is theoretically unbiased, then the most accurate urgency exponent should produce the ratio closest to the matching function (1.35). Apparently the slope of the matching function was predicted most accurately by Experiment 1.

3.5.4. Discussion.

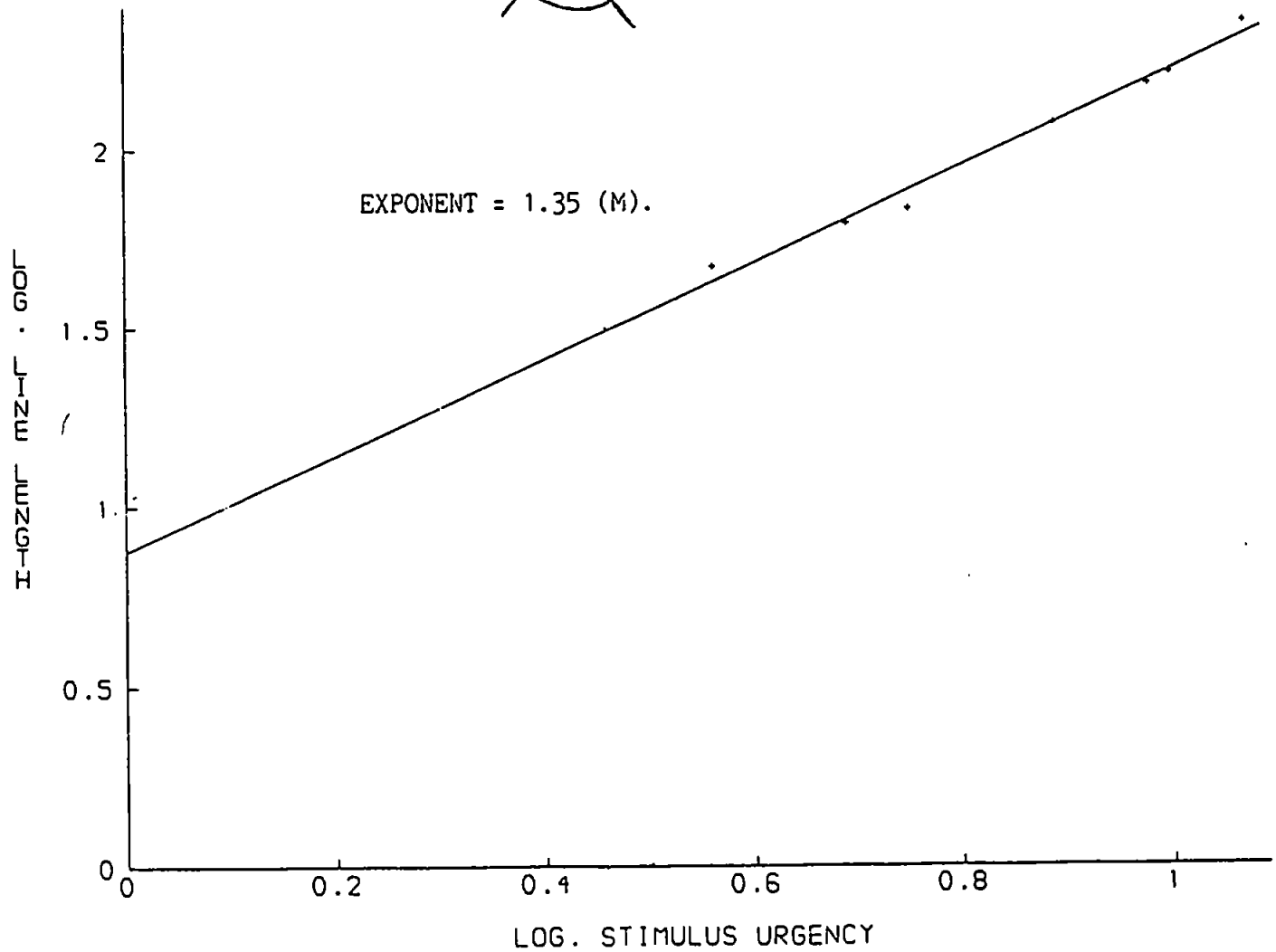
Stimuli were ranked in the order predicted from previous work, Edworthy(1988), Hellier(1988) and Edworthy(1989), with faster stimuli being perceived as more urgent. Subjects reported that they felt capable of performing the matching task, and considered the matching of length to urgency to 'make sense'.

Mean standard deviations of judgement between the group were high as a result of the large response range that was available to the subjects. As previously, (Experiments 1 & 2), the more urgent stimuli were judged least consistently. Although there was a physical limit to the responses that subjects could make, the paper size, the task was essentially a 'free modulus' one for subjects varied greatly in the amount of the available response range that they used.

The matching function related perceived urgency and line length by an slope of 1.35. As predicted by Stevens(1966), the function was a straight line, 98.2% of the variance was accounted for.

The successful matching of line length to perceived urgency allowed the validation of the exponents measured in Experiments 1-3 by the matching function. The ratio between the urgency exponent from Experiment 1 and line length was closest to the slope of the matching function. This fact indicates that Experiment 1 employed the least biased scaling technique, because its results were similar to those obtained by the unbiased cross-modality matching procedure (at least unbiased by number use).

Fig. 3.4 : Matching Function Function (Experiment 4).



3.6. General Discussion.

All of the scaling methods ranked the stimuli in the order predicted from previous work, Edworthy et al(1988), Hellier (1988) and Edworthy et al(1989), with faster stimuli being perceived as more urgent.

Although the means of subjects judgments did not differ between their first two judgments and their last six judgments in any of the experiments; in Experiments 1 and 3 a form of practice effect was exhibited. The Anova showed that individuals standard deviation after two presentations of each stimulus were significantly different from standard deviations after eight presentations of each stimulus. Examination of the data revealed that subjects were becoming less consistent with repeated presentations of the stimuli. Thus neither free modulus magnitude estimation nor category estimation would be recommended for use with more than two stimulus presentations. It is not clear whether judgement consistency is lost as a result of fatigue, boredom or confusion.

The urgency exponent measured in Experiment 1 was validated by cross modality matching. This implies that the scaling technique used in Experiment 1, free modulus magnitude estimation, was the most valid of the tested techniques, for the exponent measured by that method was closest to the slope measured by the virtually unbiased cross modality matching procedure. Furthermore, of the first three experiments Experiment 1 achieved the best fit to the Power law (99.2% of the variance accounted for), and was the only technique which was better fitted by Stevens Power Law, in a Log-Log plot than by a straight line in linear co-ordinates.

A subsidiary issue that the experiments were designed to investigate was how perceived urgency fits into Stevens Prothetic/Metathetic continua dichotomy. There is evidence from Experiments 1 and 2 that perceived urgency is a Prothetic Continuum for ratio scaling techniques revealed power function fits. However the results of Experiment 3 seem to deny this conclusion. If perceived urgency is Prothetic, the category estimation procedure would be expected to have produced a lower, not a larger, exponent than the ratio scaling methods (Stevens 1971) , and a less successful Power Function fit.

There are several possible conclusions that can be drawn from this contradictory evidence. It is possible that perceived urgency is a Prothetic Continua, and that the category rating method employed encouraged subjects to judge ratios and not differences. If the category rating method was, by methodological fault, a ratio and not an interval scaling procedure, then the close fit of the category estimation data to the power function could be explained. What could not be accounted for is the much larger exponent in Experiment 3.

A second possible conclusion is offered by Schneider and Bissett(1988). They felt that continua were not divided in terms of Prothetic/Metathetic, but in terms of their decomposability. It was said that continua that were easily divided into smaller units in the subjects mind, (their example - line length), were more easily judged in terms of ratios, whereas stimuli that were not easily conceived broken into smaller units, (their example - loudness), were more easily judged in terms of differences. It is possible therefore that perceived urgency is a 'decomposable' continuum that subjects will find it easier to judge in ratios (perhaps even if they are asked to use an interval scale as in Experiment 3).

It should be noted however that intuitively perceived urgency appears to be a Metathetic (qualitative), or nondecomposable stimulus. As such we would not have expected ratio scaling techniques to be successful. A possible explanation for the present findings is that in second order scaling such as this, it is not the descriptor of the stimulus, urgency, that determines the continuum type, but the variable, speed. If this were the case then it is plausible that what was measured was a Prothetic Continuum since speed is more qualitative and decomposable. This would explain the success of the ratio scaling techniques employed in Stevens and Schneider and Bissets terms.

Although the issue of what kind of continua perceived urgency is has not been resolved, it seems clear that it is ratio scaling techniques, in particular 'free modulus ' magnitude estimation that should be employed to scale it. In fact, providing that Stevens' Power Law exponents are used to describe data, as suggested by Weiss(1981), and the law is not used theoretically it is not particularly important that we know what type of continua perceived urgency is. Considering this, and the fact that scaling urgency involves scaling a 'second-order' continuum (about which little is known) it is proposed that exponents are only used to describe urgency data. It remains important to have an unbiased description of the data, and for this reason the 'free-modulus' method of magnitude

estimation is the technique recommended for revealing exponents. Despite Poultons (pers.comm.) claim that in most scaling cases interval scales are sufficient and ratio scales represent an unnecessary complication I remain confident in recommending the latter for scaling perceived urgency. When prioritising sounds in terms of urgency a high degree of precision is required, precision that is available through ratio scales but which would be lost by employing interval techniques.

In conclusion, three techniques for measuring perceived urgency have been compared. All scaled the stimuli in the predicted order and all produced data that fitted Stevens Power Function. It appears that perceived urgency can be successfully measured by psychophysical techniques and that Stevens Power function can be used to describe the effects of acoustic changes upon perceived urgency. On the basis of the cross modality matching validation procedure it is recommended that free modulus magnitude estimation or cross modality matching itself is used to scale perceived urgency. The exponent measured by free modulus magnitide estimation is considered the least biased exponent and thus to result from the most valid scaling procedure because it produced a ratio closest to the slope of the unbiased cross modality matching function.

CHAPTER FOUR

AN EVALUATION OF THE EFFECTS OF DIFFERENT SOUND PARAMETERS UPON PERCEIVED URGENCY.

4.1. Introduction.

In the following experiments different acoustic parameters were varied and the effects of these variations upon perceived urgency were measured by cross modality matching. The cross modality matching functions were used to show how equal amounts of perceived urgency could be communicated by each of the different parameters; that is, what change in each parameter resulted in an equal change in perceived urgency. This is important to know in practical terms so that different warnings that communicate urgency through different parameters can be set so that they can communicate the same range of urgency levels.

Chapter Three compared four different methods of measuring perceived urgency. In each experiment urgency was communicated by variations in stimulus speed. Having investigated ways of measuring perceived urgency it was possible to use the findings to consider the effects of other acoustic parameters, besides speed, upon perceived urgency. Previous research has suggested that variations in pitch (Patterson 1982, Edworthy et al 1989), repetition units (Patterson 1982, Edworthy et al 1989) and inharmonicity (Edworthy et al 1989) cause variations in perceived urgency. Experiments 5-8 investigated these relationships.

Free modulus magnitude estimation and cross modality matching were previously recommended as the best techniques to use for scaling perceived urgency. The cross modality matching procedure used in Experiment 4 was selected for use in these experiments since it is virtually unbiased, was favoured by subjects in Experiments 1-4 and is convenient to administer. Furthermore, it can be used whether the continua are Prothetic or Metathetic - a question that has not been answered for perceived urgency.

Pitch and repetition units were both easily manipulated, inharmonicity however was more problematic. Each pulse of sound has a harmonic series, a set of several

harmonics. The harmonics of a pulse exist in multiples of the fundamental frequency. Thus a pulse with a fundamental frequency of 300Hz is said to have a regular harmonic series if the harmonics are all multiples of this value, 600, 900, 1200, 1500 Hz etc. A pulse with other than regular harmonics is said to contain inharmonicity. The digital technology employed in this work numbered the harmonics in each pulse 1-15 (refer to Appendix 1A). If these values are unchanged then the harmonics are automatically created as a regular series based upon the fundamental frequency of the pulse. In order to manipulate inharmonicity each of the fifteen harmonics could be set at any level. In this way, changing the value of the 3rd harmonic, 3 by 50% to 3.5 would change its pitch from 900Hz to 1050Hz (300, the fundamental frequency multiplied by 3.5)

Preliminary work with inharmonicity however showed that it was hard to make inharmonic changes that were both quantifiable and audible. For subjects to tell whether one stimulus is more or less urgent than another, they must be able to hear differences between them, they must be audible. In the present context, it is also important that changes are readily quantifiable along one dimension so that the objective change can be plotted in logs to describe the matching function. In Experiment 7, inharmonicity was manipulated to ensure that differences between stimuli were audible although it was hard to quantify these changes along a single dimension. In Experiment 8, harmonics were manipulated in a quantifiable manner, to see if such changes were sufficiently audible. The two studies were compared to show whether or not a linear relationship exists between inharmonicity and perceived urgency when quantifiable changes are made, or whether that relationship is only demonstrated by making changes that are not as accurately quantifiable in psychophysical terms.

It was predicted that the matching functions resulting from Experiments 4-8 would demonstrate how equal increments in urgency could be indicated by manipulating either stimulus speed, pitch, units of repetition or inharmonicity. Thus different alarms could be created that communicate the same level of urgency through different parameters. Comparison of the matching functions might show that, for example a doubling in perceived urgency could be achieved by doubling stimulus speed, trebling stimulus pitch, or multiplying stimulus repetitions by four. It was predicted that comparisons of Experiments 4-8 would also show which parameter could signal changes in perceived urgency most economically i.e. which parameter requires the smallest change in itself to communicate any unit change in perceived urgency. This is important to know for alarm construction

because an economical parameter can communicate more levels of urgency than a non economical one, because it requires smaller changes in itself to do so.

Thus the present experiments aimed to assess the relationship between changes in pitch, repetition units and inharmonicity and perceived urgency. Inharmonicity is difficult to quantify and so two different ways of doing so were used. The matching functions from these experiments and from Experiment 4 were used to show how equal changes in urgency could be communicated by each of the acoustic parameters.

4.2. Experiment Five : Cross Modality Match Between Perceived Urgency (as Communicated by Pitch) and Line Length.

4.2.1. Method.

4.2.1.1. Subjects.

7 male and 7 female subjects received £1 for volunteering to participate in the study. All were undergraduate or postgraduate students in Psychology or Transport at Polytechnic South West, their ages ranged from 20-50 years. Seven subjects had previously participated in similar psychophysical studies, all reported having normal hearing.

4.2.1.2. Materials.

The laboratory arrangements were identical to those in Experiment 1, and the materials in Laboratory Two were identical to those in Experiment 4.

Bursts A, B and C, and Bursts 8-14 were stored in the Tandon. The components of these stimuli are shown in Appendix 4A. They varied in pitch from 210-680 Hz.

4.2.1.3. Procedure.

Subjects were run one at a time while seated at the desk in Laboratory Two. They were told the broad nature of the study and were asked to read the following instructions (adapted from Mashour and Hosman 1968);

"This experiment is concerned with your subjective experience of urgency. You will be presented with a series of sounds in random order. You are requested to match the length of a line to the urgency of each sound that I present by drawing a line so that its subjective length is equal to the subjective urgency of the sound. Let short lines represent low urgency and longer lines represent high urgency. Do not try to be consistent, it is only your immediate impressions that are of interest. Any questions?"

When subjects indicated that they were ready, the Experimenter sent the first stimulus from the Tandon to Laboratory Two. The next stimulus was sent when the subject indicated that he or she had finished drawing, then the next, until Bursts 8-14 had each been heard twice, in a random order. The first three Bursts that each subject heard were the practice Bursts A, B and C.

When subjects had completed the task they were asked to comment on the study and these comments were recorded. They were paid, debriefed and allowed to leave.

4.2.2. Results.

All scores are line length measurements in mm..

As shown in Table 4.1, subjects ranked the stimuli in the order predicted on the basis of previous work, Patterson(1982) and Hellier(1988). Higher pitched stimuli were perceived as being more urgent, they were represented by longer lines.

The large response range that was available to the subjects meant that their standard deviations of judgement were predictibly high.

Engen's(1971) data transformation, to eliminate inter- and intra-subject variability, was performed upon the data, (see Table 4.2). Comparison of Tables 4.1 and 4.2 showed that the data transformation produced lower mean scores for each stimulus, and that two of the stimuli were no longer ranked in the predicted order - the burst at 210Hz was ranked more urgent than the burst at 250hz

STIMULUS / SUBJECT	210	250	260	320	440	500	680
		(Pitch, in Hz).					
1	112	134.5	126.5	136	194	190	201.5
2	174	213.5	230	261	261	172.5	298
3	265	263.5	237.5	295.5	289	302.5	319
4	223.5	206	276	270	249	267.5	232
5	296	356.5	360	347	384	386	319
6	185.5	154.5	139.5	153	155.5	211	238
7	133.5	83.5	140	218.5	209	270	337
8	116.5	120	116.5	132.5	173.5	183	245
9	187	214	240.5	254.5	251	263.5	271.5
10	73	73.5	82.5	84.5	99.5	97	104.5
11	94.5	73	118	67.5	105	95	179.5
12	110.5	115.5	203	172.5	160	209.5	249.5
MEAN	164.2	167.3	189.2	199.4	210.9	220.7	249.5
ST.DEV	70.1	86.2	82.1	88	81.2	83.5	79.3

TABLE 4.1 : MEAN JUDGEMENT TO EACH STIMULUS.

STIMULUS	210	250	260	320	440	500	680
			(Pitch, in Hz).				
LOG	2.17	2.16	2.22	2.24	2.28	2.30	2.36
ANTILOG	148	146.5	169.5	175.9	194.7	201.6	233.5

TABLE4.2: MEAN JUDGEMENT AFTER TRANSFORMATION.

The method of least squares was used to fit the data to a straight line to show the slope of the matching function, see Appendix 4B. The logarithm of the pitch of each stimulus was used to represent the objective value of the urgency stimuli when the matching function was calculated. The matching function between line length and perceived urgency (as communicated by stimulus pitch) had a slope of 0.384, (Fig. 4.1)

A linear regression was performed to test the goodness of fit of the data to a straight line when it was plotted in log-log co-ordinates. 93.4% of the variance was accounted for by a straight line.

4.2.3. Discussion.

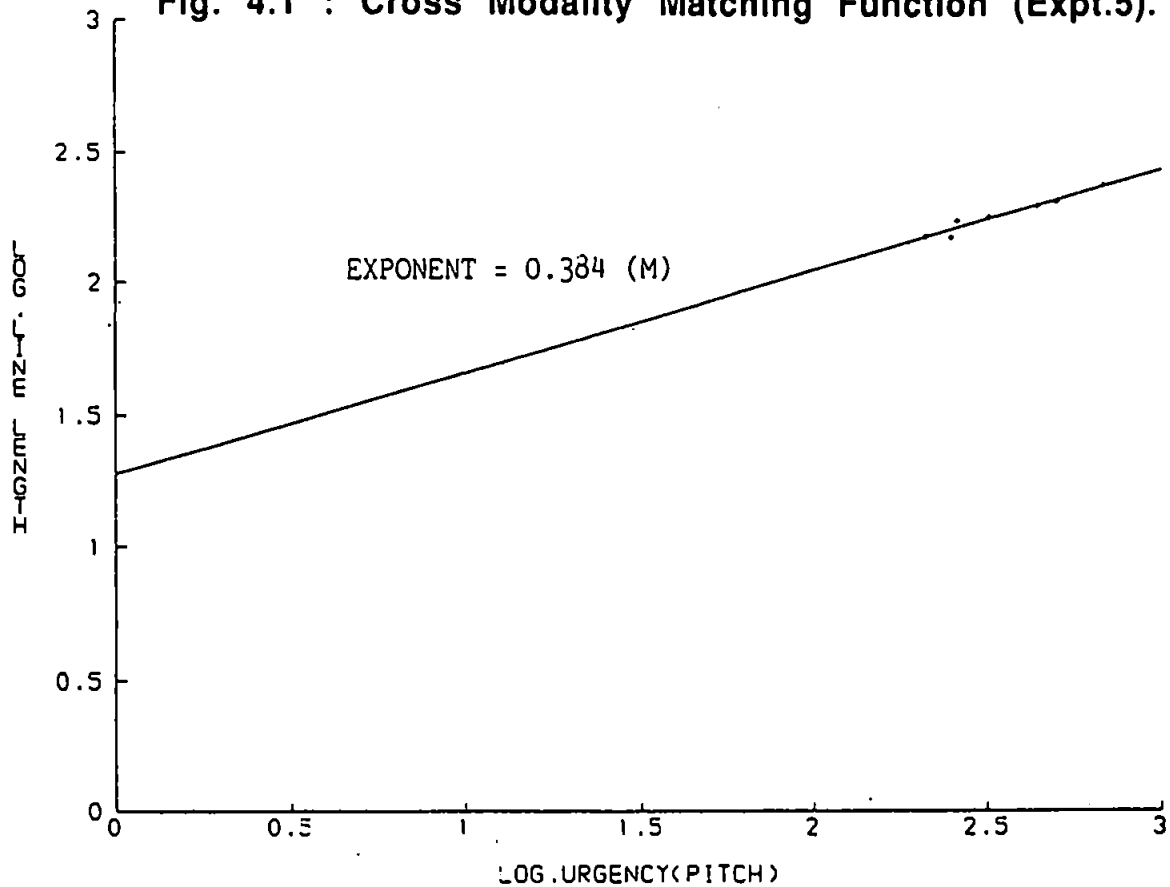
As predicted stimuli with higher pitches were judged as being more urgent (Patterson 1982, Hellier 1988, Edworthy et al 1989). Subjects stated that matching line length to pitch 'made sense'.

Standard deviations of judgement between the group were predictibly high due to the large response range that was available. Those stimuli in the middle of the urgency range were judged least consistently. As in Experiment 4 the task was essentially a 'free modulus' one, and subjects varied greatly in the amount of the available response range that they used.

After the data transformation the rank order of the two stimuli with the lowest pitches (210Hz, 250Hz) was reversed so that the latter stimuli was judged less urgent than the former. Engen (pers. comm) suggested that would typically occur when two stimuli were ranked so close, (210Hz was ranked 164.2 and stimulus 250Hz was ranked 167.3), and that it indicated that there was little real difference in their rank order.

The matching function related perceived urgency and line length by a slope of 0.384 . The fact that the slope is less than 1 implies that it takes large increments in pitch to produce relatively small increases in line length and therefore in perceived urgency. As predicted by Stevens(1966a) the matching function is a straight line in a log-log plot. A linear regression showed that 93.4% of the variance was accounted for by a straight line.

Fig. 4.1 : Cross Modality Matching Function (Expt.5).



4.3. Experiment Six : Cross Modality Match Between Perceived Urgency (as Communicated By Repetition Units) and Line length.

4.3.1. Method.

4.3.1.1. Subjects.

7 male and 7 female subjects received £1 for volunteering to participate in this study. All were undergraduate or postgraduate students in Psychology, Transport or Geography at Polytechnic South West, their ages ranged from 20-50 years. Ten of the subjects had previously participated in similar psychophysical studies. None of the participants reported having present or a history of hearing problems.

4.3.1.2. Materials.

The laboratory arrangements were identical to those used in Experiment 1.

Bursts A, B and C, and Bursts 15-21 were stored in the Tandon. The components of these stimuli are shown in Appendix 4C. They varied in repetition units from 2-6 units, each unit of repetition consisted of two 200ms pulses, the first played at 300Hz the second at 200Hz. It is unavoidable that as stimuli increase in repetition rate they also increase in length.

The materials in Laboratory Two were identical to those in Experiment 4.

4.3.1.3. Procedure.

The procedure was identical to that used in Experiment 5 except that the practice stimuli were followed by Bursts 15-21.

4.3.2. Results.

All scores are line measurements in mm..

As shown in Table 4.3 subjects ranked the stimuli in the order predicted on the basis of previous work, (Hellier 1988) - stimuli which contained more units of

repetition were perceived as being more urgent, they were represented by longer lines.

Standard deviations of judgement between subjects were exceptionally high for all of the stimuli. The highest standard deviation was for judgments to the stimulus with the least number of repetitions.

Engen's(1971) data transformation, to eliminate inter and intra subject variability was performed (see Table 4.4). Comparison of Tables 4.3 and 4.4 showed that the data transformation produced lower mean scores for each stimulus.

The method of least squares was used to fit the data to a straight line to show the slope of the matching function, (see Appendix 4D). The logarithm of the number of units of repetition of each stimulus was used to represent the objective value of the urgency stimuli when the matching function was calculated. The matching function between line length and perceived urgency (as communicated by repetition units) has an slope of 0.502, Fig. 4.2.

A linear regression was performed to test the goodness of fit of the data to a straight line when it was plotted in log-log co-ordinates. 97.6% of the variance about a straight line was accounted for.

4.3.3. Discussion.

Stimuli were ranked in the order predicted from previous work, Patterson(1982), Hellier(1988), and Edworthy et al (1989) - stimuli containing more units of repetition were perceived as being more urgent . It is unclear whether this effect was caused only by the repeating units or whether the associated increase in stimulus length also contributed. It was noted by Hellier(1988) that increases in stimulus length increase perceived urgency and so it is probable that both factors contribute to the effect. When constructing warnings by the Patterson method (1982) increasing length (by adding more pulses) always means that the repetition of that pulse is also increased. Only if continuous tone warnings were used would repetition units and stimulus length become separable. In the present investigation increases in repetition units and length will be considered synonymous since increasing stimulus length necessitates increasing repetitions.

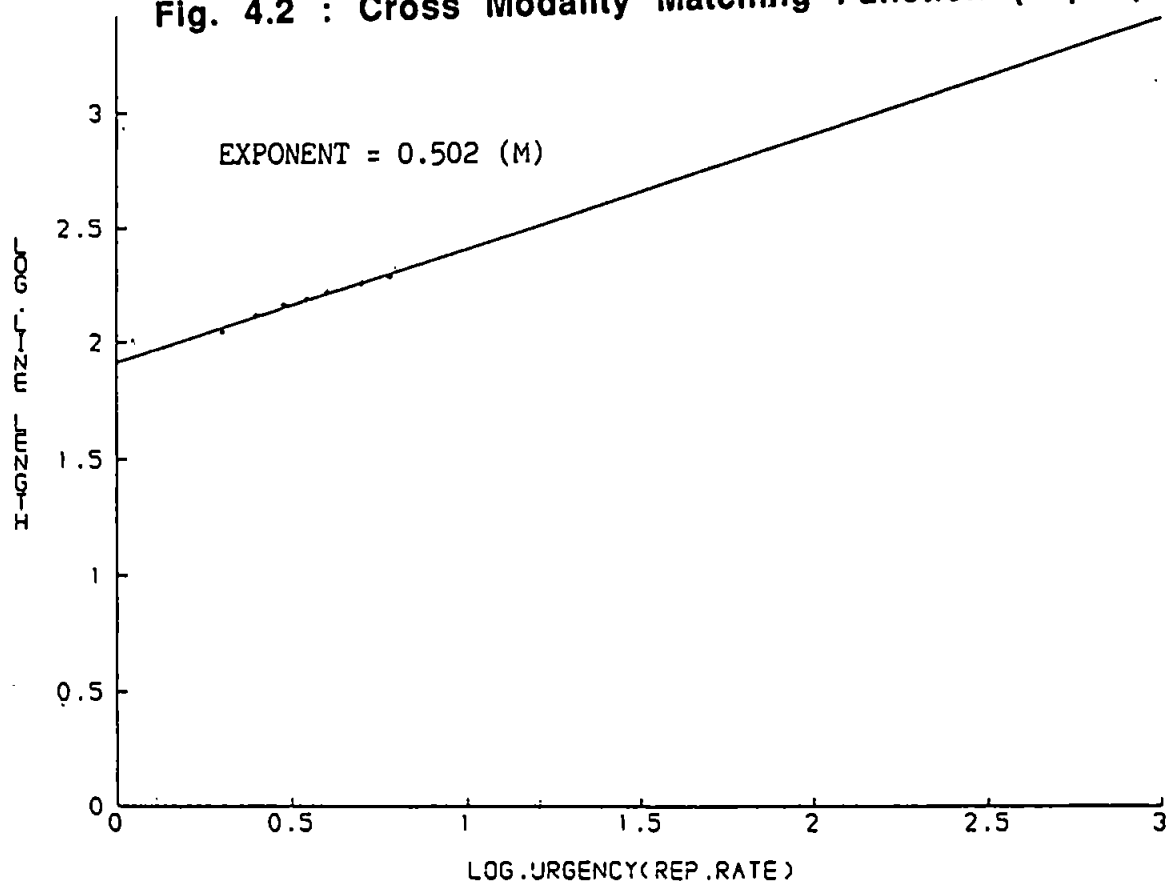
STIMULUS. SUBJECT	2	2.5	3	3.5	4	5	6
		(units of repetition).					
1	62.5	61.5	64.5	64.5	65	62	72
2	71.5	114.5	199.5	192	244	273	281.5
3	320	324	330.5	337.5	316.5	335.5	324.5
4	382.5	361	375.5	383	345.5	386	378
5	239	260	280	279.5	308	299	310
6	54.5	118.5	109	177	194.5	234	282
7	90	121	145	131.5	120.5	192	210
8	90	70.5	96.5	110.5	140.5	151.5	169.5
9	145	146	162.5	172	201	238.5	230
10	63	74	88.5	91	85	114.5	117
11	105	98	83	85.5	93.5	85	94
12	95	134	130	153.5	171.5	142.5	160
MEAN	143.1	156.9	172	181.3	190.4	209.4	219
ST.DEV	109.9	101	103.4	101.8	95.7	102.2	98.3

TABLE 4.3 : MEAN JUDGEMENT TO EACH STIMULUS.

STIMULUS	2	2.5	3	3.5	4	5	6
			(units of repetition).				
LOG	2.04	2.11	2.16	2.19	2.22	2.26	2.28
ANTILOG.	110.6	131.2	146.2	155.2	166.7	182.8	194.5

TABLE 4.4 : MEAN JUDGEMENT AFTER TRANSFORMATION.

Fig. 4.2 : Cross Modality Matching Function (Expt.6).



The standard deviations of judgement between the group were exceptionally high. This cannot be entirely accounted for by the large available response range for the standard deviations were much higher than those of Experiments 4 or 5 which had the same response range. It is possible that, despite what they said, subjects found it harder to match line length to repetition stimuli than to speed or pitch stimuli. Alternatively, the repetition stimuli may have been harder to distinguish from one another.

The matching function related perceived urgency to line length by an slope of 0.502. Again the slope is less than 1, which implies that it takes large increments in repetition rate to produce relatively small increases in line length, and thus in perceived urgency. As in Experiments 4 and 5, Stevens(1966a) claim that the matching function should be a straight line in a log.-log. plot was supported, 97.6% of the variance was accounted for.

4.4. Experiment Seven : Cross Modality Match Between Perceived Urgency (as Communicated by Number of Inharmonic Components) and Line Length.

4.4.1. Method.

4.4.1.1. Subjects.

One male and eleven female undergraduate psychology students from Polytechnic South West received one point as part of their coursework requirement for volunteering to participate in this study. Their ages ranged from 18-40 years. Five of the subjects had previously participated in similar psychophysical studies, none reported having present or a history of hearing problems.

4.4.1.2. Materials.

The Laboratory arrangements were identical to those used in Experiment 1.

The stimuli were designed so that the differences between them were as audible as was possible, and no regard was given to how easy these differences would be to quantify. The fundamental frequency was set at 300Hz, the first burst contained

regular harmonics. In the second, the middle harmonic, the 8th (2400Hz), was made irregular, so that its frequency was 8.5 times the fundamental. It was thus 50% irregular with reference to the fundamental (300*8.5=2550Hz). In each of the next four bursts two more harmonics were varied, one lower than the 8th, and one higher. Thus, in each consecutive burst, two more harmonics were made irregular. In the last burst, all but the first harmonic were altered. (Appendix 4E).

These changes were hard to quantify because the harmonics were altered by different percentages to make them maximally audible. For example, the 3rd harmonic was set to 3.1, a 10% change (930Hz), whereas the 13th harmonic was set to 13.9, a 90% change (4170hz). Having considered adding the ratios between the harmonics; scoring the change in relation to the harmonic furthest away, (so that 3.1, which was 9 points away from 4, would have a larger score than 4.5, which was 5 points away from 5); and pitch matching, it was concluded that such measures were either too arbitrary or impractical. It was decided to quantify the changes simply by counting the number of harmonics that had been altered (Shailer, pers. comm.).

This measure does not take into account the percentage change of each harmonic, it treats all changes as equal. It is an objective quantification however and provides one way of looking at stimulus inharmonicity - assuming that increasing the number of irregular harmonics increases the inharmonicity. It is possible that this broad quantification is sufficient for describing the relationship between stimulus inharmonicity and perceived urgency.

Bursts D, E, and F and Bursts 22-28 were stored in the Tandon. The practice bursts D, E and F were different to those used in previous experiments for they differed in harmonicity. Because it is unusual to listen to changes in harmonicity between sounds, these practice bursts familiarised subjects with the type of stimulus differences that they had to listen for before they heard the experimental stimuli. In the pulses which made up the practice bursts the irregular harmonics were amplitude weighted to make them more apparent. For explanation of amplitude weighting see Appendix 1A.

The materials in Laboratory Two were identical to those used in Experiment 4.

4.4.1.3. Procedure.

Subjects were run one at a time while seated at a desk in Laboratory Two. They were told the broad nature of the study and were asked to read the following instructions, (adapted from Mashour and Hosman 1968);

"This experiment is concerned with your subjective experience of urgency. You will be presented with a series of sounds in random order. You are requested to match the length of a line to the urgency of each sound that I present by drawing a line so that its subjective length is equal to the subjective urgency of the sound. Let short lines represent low urgency and longer lines represent high urgency. Do not try to be consistent, it is only your immediate impressions that are of interest. Any questions?"

When subjects indicated that they were ready, the Experimenter sent the first stimulus from the Tandon to Laboratory Two. The second stimulus was sent when the subject indicated that he or she had finished drawing, and then the next, until Bursts 22-28 had been heard twice each in a random order. The first three Bursts that each subject heard were the practice Bursts D, E and F.
When the task was completed, subjects comments were recorded. They were debriefed and allowed to leave.

4.4.2. Results.

All scores are line lengths in mm.

As shown in Table 4.5, most stimuli were ranked in the order predicted from previous work, (Edworthy et al 1989), with more inharmonic stimuli being perceived as more urgent, for they were represented by longer lines. The stimulus that was not ranked exactly in the predicted order was Burst 27, with 9 irregular harmonics.

As in previous experiments, the large available response range meant that the standard deviation of judgement between subjects was predictibly high. Standard deviations were lowest for the least urgent stimulus and highest for the most urgent stimulus.

STIMULUS / SUBJECT	0	1	3	5	7	9	14
		(Number of irregular harmonics).					
1	203.5	183.5	189	192	176.5	177.5	196.5
2	57.5	81	101.5	75	105	87.5	64.5
3	55.5	38	43.5	46	48	41	58.5
4	51.5	40.5	50	75	70	86.5	60.5
5	141.5	201.5	216	232.5	209	224	195
6	92	91	88.5	171.5	195	92.5	83.5
7	26.5	24	30.5	26	47	45	42.5
8	56	32	38.5	46	71.5	76	42
9	69.5	78	120	119.5	87.5	85.5	90.5
10	101.5	112.5	153.5	153	179	175	170
11	50.5	76.5	57.5	73	76.5	120.5	115
12	136	130	120	131	124	114	103
MEAN	86.8	90.7	100.7	111.7	115.7	110.4	118.5
ST.DEV	51	57.6	61.3	65.3	59.2	55.6	76.5

TABLE 4.5 : MEAN JUDGEMENT TO EACH STIMULUS.

STIMULUS	0	1	3	5	7	9	14
			(Number of irregular harmonics)				
LOG	1.810	1.801	1.862	1.896	1.943	1.922	1.922
ANTILOG	64.71	63.35	72.79	78.78	87.87	83.68	83.73

TABLE 4.6 : MEAN JUDGEMENT AFTER TRANSFORMATION.

Engen's(1971) data transformation to eliminate inter and intra subject variability was performed upon the data (see Table 4.6). Comparison of Tables 4.5 and 4.6 shows that the data transformation produced lower mean scores for each subject. In the transformed scores Bursts 23 (1 irregular harmonic) and 26 (7 irregular harmonics) were not ranked in the predicted order.

Although it appears that the untransformed scores reflect the predicted order of urgency of the bursts better than those in Table 4.6, it was decided to continue analyses on the transformed scores. This decision was taken to ensure continuity with previous experiments, and so that inter- and intra- subject variability would be eliminated from the data. Eliminating such variability should result in scores which more accurately reflect the perceived urgency of the sounds, as is implicitly assumed by all applications of such a transformation.

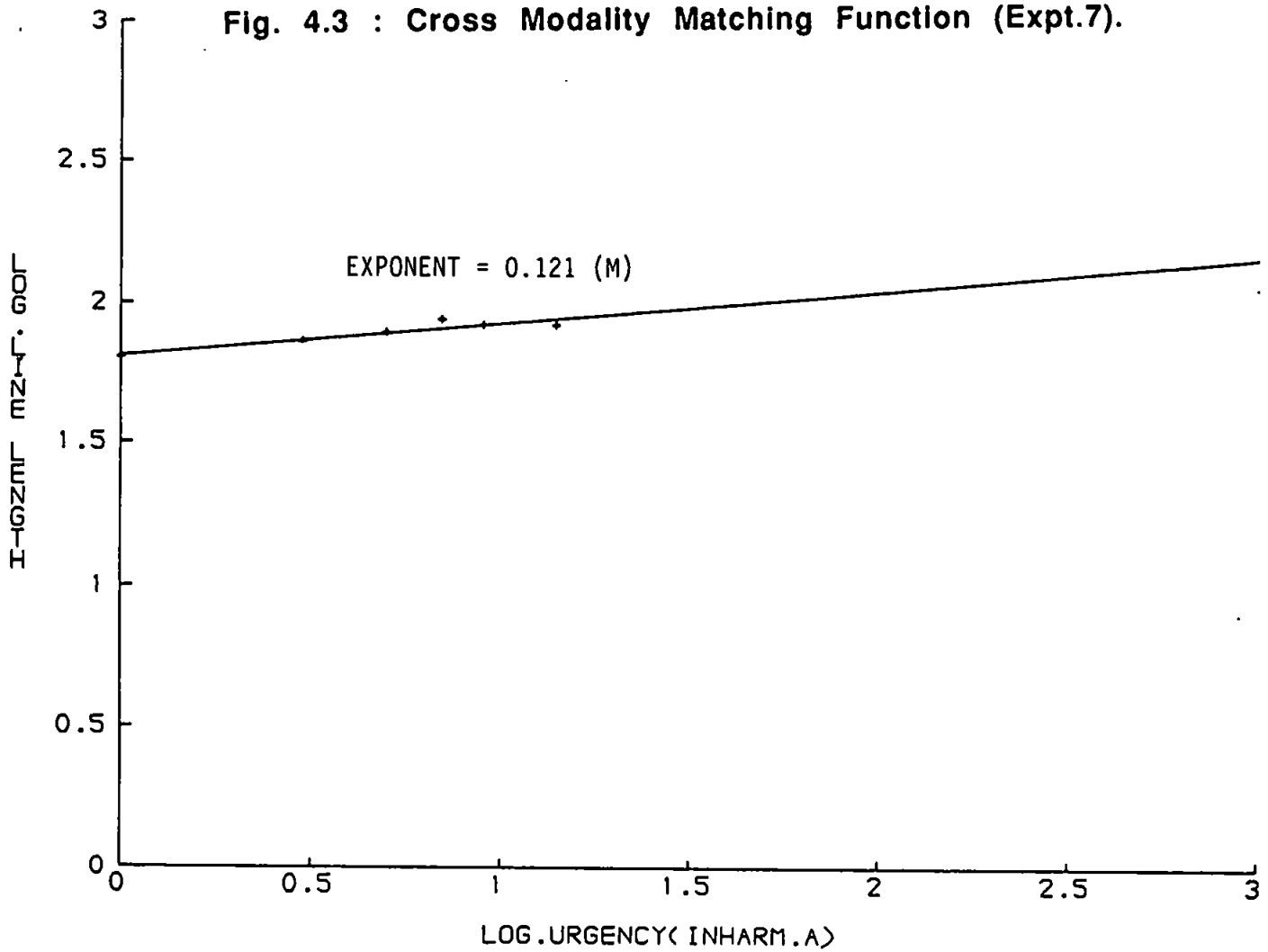
The method of least squares was used to fit the data to a straight line to show the slope of the matching function, see Appendix 4F. The logarithm of the number of irregular harmonics was used as the objective stimulus measure to calculate the matching function. Burst 22 was excluded from analyses because it contained no irregular harmonics, and there is no logarithm of 0. The matching function was therefore calculated between Bursts 23-28 and the corresponding line measurements. The matching function between line length and perceived urgency (as communicated by number of inharmonic components) had an slope of 0.121 (Fig. 4.3).

A linear regression was performed to test the goodness of fit of the data when it was plotted in log-log co-ordinates to a straight line. 84.3% of the variance about a straight line was accounted for. As Stevens(1966a) predicted, the matching function was approximately a straight line in log-log co-ordinates.

4.4.3. Discussion.

Although every score did not confirm the hypothesis, the general trend was that stimuli containing more inharmonic components were perceived as being more urgent, as predicted on the basis of previous work, e.g. Edworthy et al (1989) and Patterson(1982). According to the transformed scores , which are assumed to be closer to the 'true' scores, since sources of variability are removed, the exceptions were Bursts 27 (9 irregular harmonics) and 28 (14 irregular harmonics) which

Fig. 4.3 : Cross Modality Matching Function (Expt.7).



were rated equally urgent; Burst 26 (7 irregular harmonics) which was rated most urgent and Burst 23 (1 irregular harmonic) which was rated less urgent than the regular stimulus. Four out of the seven bursts were ranked in the predicted order, it was one burst with lowest inharmonicity and the two with highest inharmonicity that violated the predictions.

Although the standard deviations of judgement between subjects were high as a result of the large unrestricted response range that was available, they were lower than in Expts.4-6. The least urgent stimulus was judged most consistently between subjects, and the most urgent stimulus least consistently. This is not unexpected, for when judging the least urgent stimulus subjects are more likely to use similar line lengths, as the stimuli increase in urgency different subjects represent these increases differently and response variability therefore increases.

Two interpretations of this data are possible. It is possible that Burst 26 was heard as more inharmonic than the other bursts, and therefore that our quantification which equated increase inharmonicity with increased number of irregular harmonics was inadequate. This interpretation can be evaluated when inharmonicity is quantified in a different way in Experiment 8. Alternatively, there could be a discrimination asymptote beyond which further increases in irregular harmonics do not result in increases in perceived inharmonicity or perceived urgency because the increasing number of irregular harmonics do not sound perceptibly different. The latter interpretation is supported by the observation that although Bursts 27 and 28 vary greatly in the number of irregular harmonics that they contain, they are ranked almost identically.

The matching function between line length and perceived urgency as communicated by number of inharmonic components had a slope of 0.121. This is very low, and according to Stevens(1966) implies that it would take very large increases in the number of irregular harmonics to produce a unit change in perceived urgency. Increasing the number of inharmonic components is therefore not an economical way to communicate perceived urgency. A different method of quantifying stimulus inharmonicity may reveal a stronger relationship between that parameter and perceived urgency. As predicted by Stevens(1966) the matching function was a straight line with 83.4% of the variance accounted for.

4.5. Experiment Eight : Cross Modality Match Between Perceived Urgency (as Communicated by Inharmonicity) and Line Length.

4.5.1. Method.

4.5.1.1. Subjects.

Two male and ten female undergraduate psychology students from Polytechnic South West received a coursework point for volunteering to participate in this study. Their ages ranged from 18-43 years. Five of the subjects had previously participated in similar psychophysical studies, none reported present or a history of hearing problems.

4.5.1.2. Materials.

The laboratory arrangements were identical to those used in Experiment 1.

The primary objective of changes to inharmonicity in this study was that they should be easily quantified. The secondary objective was that the changes should be as audible as was possible given the primary constraint. As in Experiment 7, the fundamental frequency was set at 300Hz and the first Burst contained regular harmonics. In subsequent bursts, the value of the third harmonic was altered so that it was 3.1(930Hz), 3.3 (990Hz), 3.5 (1050Hz), 3.7 (1110Hz), 3.8 (1140Hz) and 3.9 (1170Hz) times the fundamental. Thus, in each burst the same harmonic was manipulated, but by an increasing percentage of the fundamental. These changes to stimulus inharmonicity were easily quantifiable - by assuming that increases in the amount by which the third harmonic was altered, ie. increases in the percentage change from 3.0, corresponded to increases in stimulus inharmonicity.

Several measures were employed to ensure that differences between the stimuli were as audible as possible. The pattern recognition theories of pitch perception led to the choice of the third harmonic to convey stimulus inharmonicity. Authors such as Goldstein(1973), Houtgast(1976) and Moore et al(1985) state that the lower harmonics in a complex tone, up to the fifth or sixth, were the most important for determining the pitch of that tone. It was felt that manipulating a harmonic within that 'dominant region' would have the maximum effect upon

stimulus inharmonicity for the changes would be more audible. Furthermore, the third harmonic was amplitude weighted in each burst to make it more obvious (see Appendix1A). Moore et al(1985) said that increasing the level of a single low harmonic could increase its dominance.

In this study inharmonicity is thus defined as increasing linearly with the percentage change in the third harmonic, so that the more the third harmonic was altered, the more inharmonic the stimulus . Any other dominant harmonic, the third to the sixth, could have been altered in the manner described. Other ways of manipulating stimulus inharmonicity, for example by varying the amplitude envelope, altering the odd or the even harmonics, varying the shape of the spectrum or by using the log. spectrum to scale inharmonicity are not as easy to quantify along a single dimension as the method described above.

Bursts D, E and F, and Bursts 29-35 were stored in the Tandon. All Bursts are shown in Appendix 4G.

The Laboratory arrangements were identical to those used in Experiment 4.

4.5.1.3. Procedure.

Subjects were run one at a time while seated at the desk in Laboratory Two. They were told the broad nature of the study and were asked to read the following instructions (adapted from Mashour and Hosman 1968);

"This experiment is concerned with your subjective experience of urgency. You will be presented with a series of sounds in random order. You are requested to match the length of a line to the urgency of each sound that I present by drawing a line so that its subjective length is equal to the subjective urgency of the sound. Let short lines represent low urgency and longer lines represent high urgency. Do not try to be consistent, it is only your immediate impressions that are of interest. Any questions?"

When the subject was ready, the first stimulus was send from the Tandon to Laboratory Two. The second stimulus was sent when the subject indicated that he or she had finished drawing and then the next, until Bursts 29-35 had been heard twice each in random order. The first three bursts that each subject heard were the practice Bursts D, E and F.

4.5.2. Results.

All scores are line lengths in mm.

Table 4.7 shows that the stimuli were not ranked in the order predicted by previous work, (Edworthy et al 1989), for perceived urgency did not increase with stimulus inharmonicity. The middle stimulus, with a harmonic value of 3.5 was perceived as being the most urgent.

The standard deviation of judgement between subjects was high. Standard deviation of judgement was not systematically related to perceived urgency or to the value of the altered harmonic. It was highest for the harmonic stimulus and lowest for the stimulus with the altered harmonic value of 3.1.

Engen's(1971) data transformation was performed upon the data to eliminate inter and intra subject variability (Table 4.8). Comparison of Tables 4.7 and 4.8 shows that the data transformation has resulted in lower mean scores for each subject, as in previous experiments. The transformed scores more closely reflect the predicted order of the stimuli, with stimuli 3.0 - 3.5 being ranked as increasing in urgency. The last three stimuli are less urgent than the middle stimulus 3.5.

Although the data appears to be more of an inverted u shape than the straight line function prescribed by Stevens, the method of least squares was used to test how well the data could be described by a straight line, and to calculate the matching function between stimulus inharmonicity and perceived urgency, see Appendix 4H. The value of the manipulated harmonic was used as the objective measure of stimulus inharmonicity. The matching function is shown in Fig. 4.4.

STIMULUS SUBJECT	3.0	3.1 (Value of altered harmonic).	3.3	3.5	3.7	3.8	3.9
1	94	85.5	151	174	104	77.5	195.5
2	257	109.5	228.5	229.5	166.5	216	149
3	14.5	23.5	31.5	28	27.5	29.5	23
4	22.5	57	51.5	69	68.5	87.5	70.5
5	142.5	88.5	192.5	146.5	96.5	70.5	101
6	128	175.5	134	166	158	113	133.5
7	68.5	91.5	59.5	86.5	78.5	59	55
8	28.5	29	31	33	36	30	31.5
9	33.5	23.5	34	34.5	39	50	42.5
10	7	19.5	11.5	9.5	14	13.5	15
11	71.5	86	65	84.5	90.5	57	78
12	116	107.5	120.5	91	99.5	113	117.5
MEAN	81.9	74.7	92.5	96	81.5	76.3	84.3
ST.DEV	71.9	46.5	70.9	68.8	48.3	53.9	56.6

TABLE4.7 : MEAN JUDGEMENT TO EACH STIMULUS.

STIMULUS	3.0	3.1	3.3 (value of altered harmonic).	3.5	3.7	3.8	3.9
LOG	1.581	1.674	1.734	1.753	1.731	1.701	1.717
ANTILOG	38.10	47.26	54.26	56.74	53.82	50.23	52.14

TABLE 4.8 : MEAN JUDGEMENT AFTER TRANSFORMATION

A linear regression was performed to test the goodness of fit of the logarithmic data to a straight line. Only 30.9% of the variance in the data was accounted for by a straight line. Thus the data does not fit the matching function very well. A linear regression on a linear plot of the data shows that the data was no better described by a straight line when it was plotted in linear co-ordinates, for only 24.8% of the variance was accounted for.

4.5.3. Discussion.

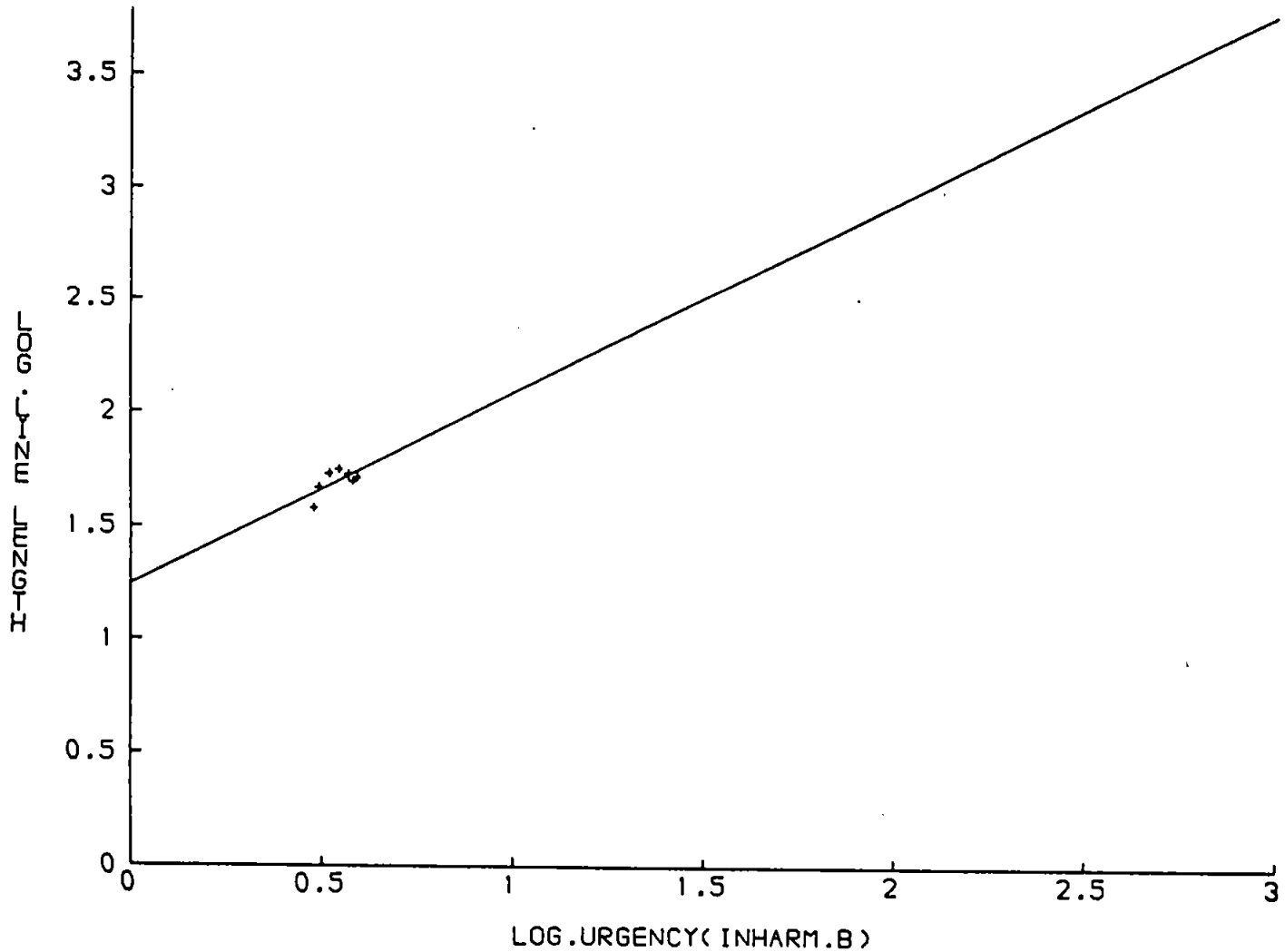
Experiment 8 showed that sound parameter changes do not always exhibit a clear or quantifiable relationship with perceived urgency. Our prediction, based on the work of for example Edworthy et al(1988), that higher percentage changes in one harmonic would be perceived as more urgent, was not supported. In fact the middle stimulus, which had a 50% change in the value of the third harmonic, was judged most urgent. The second stimulus which had a 10% alteration in the value of the third harmonic was judged least urgent, and not, as predicted, the regular Burst.

Despite efforts to make the stimulus differences as audible as possible, all but three of the subjects complained that the stimuli were either 'impossible' or 'very difficult' to tell apart. Nevertheless, they appear to have been able to do the task for their standard deviation of judgement was no higher than in previous Experiments.

As in previous studies, the standard deviation of judgement between subjects was high. This is partly the result of the large unrestricted response range that is available to them. The least urgent stimulus was judged most consistently, and the mid-urgent stimulus least consistently. Although there is no clear relationship between % alteration of the harmonic and consistency of judgement, as in previous studies lower urgency stimuli tended to be judged more consistently.

After the data transformation the burst containing the harmonic that had been altered by 50% was still judged most urgent, followed by Bursts with harmonics altered by 30%, 70%, 90%, 80%, 10% and 0%. This data indicates that a 50% change in the harmonic is the most urgent, and that % changes above or below this are less urgent. The function between % change of one harmonic and perceived

Fig. 4.4 : Cross Modality Matching Function (Expt.8).



urgency forms an inverted U shape. The only burst that did not follow this pattern was the one containing the harmonic altered by 90%.

This result is unsurprising when it is considered that a 50% change in the harmonic makes it most different from the harmonic above and below it in the harmonic series. For example setting the 3rd harmonic at 3.5 (1050Hz) means that it is as different as is possible from the 3rd harmonic (900Hz), and from the 4th (1200Hz). A change over 50% makes the value closer to the value of the 4th harmonic and a change below 50% results in a value closer to the 3rd harmonic. Thus a 50% change in the value of a harmonic makes it most different from the regular harmonic series, most inharmonic. Thus quantifying inharmonicity as the percentage change in the value of a harmonic is not suitable for our purposes since it is likely to result in a u-shaped function with a 50% change in the value of the harmonic being perceived as most inharmonic and most urgent. In order to describe the effects of parameter changes on perceived urgency in terms of the power law an objective method of quantification must be employed that reveals a linear relationship between the subjective and objective parameters.

The matching function between line length and stimulus inharmonicity had a slope of 0.850. A regression showed that only 30% of the variance was accounted for by the matching function, this indicates that the slope was not reliable. Stimulus inharmonicity, as defined by percentage increases in one dominant harmonic, was not related in a linear way to perceived urgency.

By the method of quantification presently employed an inverted U shaped relationship was revealed between perceived urgency and inharmonicity. It is possible that were a different definition of inharmonicity employed then a linear function might be revealed as in the previous Experiment. At present however it is impossible to see how inharmonicity might otherwise be quantified (Moore, B., pers. comm.).

In Experiment 7, although it was feared that the objective quantification of stimulus inharmonicity had been imprecise, a matching function relating inharmonicity and perceived urgency by a function of 0.121 was revealed. In Experiment 8 where the objective quantification of inharmonicity, as the percentage change in a single harmonic, was thought to be more precise, there was no linear relationship between perceived urgency and inharmonicity. Although in Experiment 7, the parameter changes were not precisely quantified, they

produced a linear relationship with perceived urgency which can be measured in psychophysical terms. For the present purposes inharmonicity will be defined as in Experiment Seven, as the number of inharmonic components in a stimulus. Then the power law can be used to describe the relationship between inharmonicity and urgency as it can with other acoustic parameters. Providing the same definition is used in future studies, the measurements taken will be valid. In future work on stimulus inharmonicity, the objective definition of inharmonicity should be specified. As has been shown different methods of describing inharmonicity will yield different results.

4.6. General Discussion.

It has been demonstrated in these experiments that increases in pitch, repetition units and the number of inharmonic components result in increases in perceived urgency. The previous chapter showed that increases in stimulus speed (pulse rate) also increased perceived urgency. It has also been shown that the percentage increase in the value of one harmonic is not a satisfactory way of quantifying inharmonicity for present purposes.

The data from Experiments 4-8 is summarised in Table 4.9. When line length was adjusted to match stimuli which communicated urgency through speed (pulse rate), the steepest slope(1.35) and the best fit of the matching function to a straight line resulted. The shallowest slope(0.121) resulted when line length was matched to the number of inharmonic components, and the poorest fit of the matching function to a straight line resulted when line length was matched to the percentage increase in the value of the third harmonic (30.9%).

The matching functions allow the calculation of the levels of each parameter required to communicate equal increments in perceived urgency. Stevens Power Law,

$$\log.s = m(\log.\ o) + \log.\ k, \qquad \text{(Equation 4.1)}$$

can be used to show how line lengths, and thus perceived urgency, can be altered by manipulating the values of each of the parameters. The equation allows us to predict line length, (urgency), if the parameter values of the stimuli are known. We can change it around to predict the parameter values from the line length (urgency), and can thus tell by how much we have to change the parameter values of the stimuli to obtain a set increase in line length (urgency). Thus we use,

$$\log. o = (\log. s - \log. k)/ m$$

(Equation 4.2)

We can see how the precise values are calculated by considering as an example a trebling in line length from 50mm-150mm, in Experiment 5. S is the subjective value of the parameter, the line length, k is the intercept of the matching function, and m is the slope of the matching function. When line length (s) is 50mm (log. s =1.69),

$$\log.o = (1.69 - 1.28)/0.384, \ \log.o = 1.09$$
$$o = 12.33\text{Hz}$$

(Equation 4.3)

When line length is 150mm, (log. s = 2.176),

$$\log.o = (2.176 - 1.28)/0.384, \ \log.o = 2.333$$
$$o = 215.5\text{Hz}$$

(Equation 4.4)

If we divide the value of stimulus pitch when line length is 150mm (215.5) by the value of stimulus pitch when line length is 50mm (12.33) we are left with 17.4. This is the amount by which the frequency must be multiplied by to result in a trebling in line length and so a trebling in urgency.

By taking different values of line length and using the matching function intercepts and slopes in the above manner it is possible to calculate the amount by which each of the parameters had to be increased to produce 50% increases, doublings and treblings in line length (See Table 4.10).

From the data in Table 4.10, it is apparent that the larger the matching function slope, the less parameter increase is required to communicate an equal increase in urgency. For example, to communicate a doubling in urgency, speed would have to be multiplied by 1.6, repetition units by 4, pitch by 6 and the number of inharmonic components by 307. This is because when a slope is steeper (more than 1) the subjective value of the stimulus increases quickly in relation to the objective value, so that small changes in the objective value of the stimulus can indicate large subjective changes. When the slope is shallow (less than 1) it takes large changes in the objective value of the stimulus to communicate small subjective changes. Thus when the slope is very small, as in the case of inharmonicity, it takes very large parameter changes to alter line length. These

huge changes are less surprising when the fact that the slope is close to 0, no gradient, is considered.

It should be noted that the increases in line length do not correspond exactly to increases in perceived urgency, for example a doubling in line length does not imply an exact doubling in perceived urgency. This is because line length has a slope of 1.1, since subjective and objective length do not correspond exactly. An adjustment would have to be made to the gradient of the matching function to account for this if we were to make absolute statements about the amount by which urgency had been increased, it would have to be divided by 1.1. For the present purposes it is enough to know that in each case the specific increases in line length result in equal increments in perceived urgency, without knowing the absolute values of those increments. Each of the changes in Fig. 4.10 results in a subjectively equal change in perceived urgency. Table 4.10

In conclusion, it has been demonstrated that increases in pitch, repetition units and inharmonicity result in increases in perceived urgency. It was also shown that a quantifiable relationship exists between inharmonicity and perceived urgency if the acoustic changes are described as the increase in the number of inharmonic components in a stimulus. The experiments also show that equal increments in urgency can be communicated by variations in speed, pitch, repetition units and the number of inharmonic components. The most economical parameter through which to communicate urgency is stimulus speed. Stimulus speed has to be multiplied by a smaller amount than any other parameter to communicate an equal change in perceived urgency. If practical and ergonomic factors are also considered, then stimulus speed is the only parameter that has been considered here that could be used to treble urgency - mulitplying pitch by 17.4 is likely to be aversive, mulitplying repetition units by 8.9 would make a stimulus too long to be useful, and multiplying inharmonic components by 8773 would be impossible. The data supports the work of Patterson(1982) and Hellier(1989) who felt that urgency could be communicated through stimulus speed, stimulus pitch, stimulus repetition units and stimulus inharmonicity, and demonstrates which of these changes would be most practically useful. It also reveals the various strengths of relationship between these parameters and urgency.

EXPT.	CRITERION STIM	ADJUSTED STIM	SLOPE	% var
4	URGENCY(speed)	LINE LENGTH	1.35	98.2
5	URGENCY(pitch)	LINE LENGTH	0.384	93.4
6	URGENCY(repetitions)	LINE LENGTH	0.502	97.6
7	URGENCY(inharmonicity.a.)	LINE LENGTH	0.121	84.3
8	URGENCY(inharmonicity.b.)	LINE LENGTH	0.85	30.9

TABLE4.9 : SLOPES DERIVED FROM EXPERIMENTS 4-8.

EXPT			INCREASE IN LINE LENGTH		
	PARAMETER	SLOPE	50%	DOUBLE	TREBLE
4	SPEED	1.35	*1.3	*1.6	*2.2
5	PITCH	0.384	*2.8	* 6	*17.4
6	REPETITION	0.502	*2.2	* 4	*8.9
7	INHARMONICITY	0.121	*28.5	*307	*8773

TABLE 4.10 : RELATIONSHIP BETWEEN SOUND PARAMETERS AND LINE LENGTH

CHAPTER FIVE.

EXPERIMENT NINE : A STUDY INVESTIGATING THE EFFECTS OF COMBINING DIFFERENT SOUND PARAMETERS.

5.1. INTRODUCTION.

This experiment was designed to validate and further explore the findings from Experiments 4-7. Theoretically equal units of urgency from the acoustic parameters previously studied were combined in the same experimental stimuli. The urgency equivalence of the parameters was assessed, and the effects of combining the parameters in stimuli was noted.

In Experiments 4-7 only one parameter at a time was varied while all others were held constant at ergonomic levels (see Patterson 1982). Plomp and Steencken(1969) also varied only one parameter at a time when they investigated the importance of phase relative to harmonic amplitude pattern on timbre. As Freed and Martens(1986) pointed out, examining parameters in isolation provides a way of determining the relative importance of those parameters, as has been done.

The matching functions from Experiments 4-7 enabled changes in speed, pitch, repetition units and inharmonicity to be related to changes in perceived urgency. The exponents of the matching functions quantified the changes that had to be made in the acoustic parameters to produce a unit change in perceived urgency. The changes required in each acoustic parameter to produce equivalent changes in perceived urgency were revealed. It is apparent that speed was the most economical parameter through which to communicate changes in perceived urgency (since it took the smallest changes in speed to produce a unit change in perceived urgency), and inharmonicity was the least economical. Because such huge changes are required in inharmonicity to produce change in perceived urgency the parameter is of no practical use for communicating urgency to designers of Patterson type warnings. It is therefore excluded from this investigation.

Momtahan(1990) criticised studies in which only one parameter at a time were varied. She said that the findings of such research was limited because not all parameters were co-varied together. It should be noted that the experiments were designed in this way because it was felt that before the effects of combining different parameters together were investigated it was important to see how each parameter individually affected perceived urgency and the relative strengths of the different parameters. Not only was this considered the logical first step of such an exploration, but it also provides invaluable information for use when warnings are made by only varying one parameter at a time. Having investigated each parameter in isolation it is possible to follow Momtahan's suggestion and look at them in combination.

In the present study, a high, medium and a low level of urgency was created for each of three acoustic parameters, speed, pitch and repetition units. They were combined in the stimuli so that each stimulus contained one level of urgency for each parameter. That is, each stimulus contained pitch at a high medium or low urgency level, repetition at a high medium or low urgency level and speed at a high medium or low urgency level. As was discussed in the previous chapter, the assumptions of Stevens Power Law and of cross modality matching mean that the matching functions can be used to specify equivalent levels of urgency between the three parameters. The matching functions from Chapter Four were used to create urgency levels that were equivalent between the parameters, that is, the high urgency level in pitch was as urgent as the high urgency level in speed and repetition units etc.

When the matching functions from Experiments 4, 5 and 6 were constructed, subjects made their judgements by adjusting line length. It is well documented (Stevens and Galanter 1957), that when subjects use line length to make judgements they do so in a subjective way, so that actual and subjective line length are not the same. Line length has an exponent of 1.1, that is, subjects' subjective line length is a little longer than actual line length. Before it was possible to use the matching functions to select stimulus levels it was necessary to adjust the matching function exponents so that they reflected the judgements that subjects would have made if they had been using actual, not subjective, line length. Each matching function was therefore divided by 1.1, the exponent for line length. The adjusted matching functions that were used in stimulus construction are shown below.

Urgency as communicated by speed = 7.367*speed(pulse rate)^1.227

Urgency as communicated by pitch = 18.89*pitch^0.349

Urgency as communicated by repetitions = 81.28*repetitions^0.349 456.

(Equations 5.1, 5.2, 5.3)

High, medium and low urgency levels for each of the three parameters were calculated using the above equations. The stimuli are shown in Appendix 5A. For each parameter the high, medium and low levels of urgency were the parameter values at which the urgency, in line length, was calculated from the matching functions to be 150, 113 and 85mm respectively. Thus the high, medium and low urgency levels were equal between the three parameters. The relative urgency increases from one level to another were also equal - each urgency level was 33% more urgent than the one below. These theoretically equally urgent levels of the different parameters were combined to see if they remained equal levels, in absolute or relative terms, when the parameters were co-varied in the same stimuli.

On the basis of findings in Experiments 4-6 it was predicted that higher urgency levels would result in higher judgements. It was also predicted that at each urgency level judgements would be equal across parameters so that for example the high urgency pitch levels would be judged the same as the high urgent speed and repetition levels. Furthermore it was predicted that the relative urgency differences between the levels would be preserved within and between levels. Thus for example a medium level of pitch should be perceived as being 33% more urgent than a low level of speed, pitch or repetition and 33% less urgent than high levels of any parameter. As a result of these predictions it was expected that mean judgements would be the same to all stimuli that had the same combination of levels, whatever parameters were selected. Each level of each parameter was expected to contribute equally to the urgency judgement for each stimulus. In terms of multiple regression, the regression coefficients for the three parameters would be expected to be the same.

Previous attempts have been made to assess the relative contributions of different acoustic parameters to sound judgements. These attempts have been confounded by the fact that the parameter levels selected for comparison have been arbitary - it is not very informative to say that one parameter contributes more than another to judgement when one parameter may have been at a higher level than another. What is more interesting is to set the parameters at subjectively equal levels and

to see if then one contributes more than another to judgement. Such an investigation would reveal whether one parameter was more salient or distinctive than another in terms of the particular judgement.

Freed and Martens(1986) co-varied eight acoustical parameters to find the greatest predictor of perceived hardness. They said themselves that their conclusions were limited because the parameter levels that they selected were arbitary. They recommended that stimuli were selected on the basis of psychophysical results so that parameter levels were equivalent. This recommendation is followed in the proposed investigation.

Freed and Martens(1986) recommendation was not however adhered to by Momtahan(1990). Her study was similar to the present one in that she combined different urgency levels of different parameters in the same stimuli to see how much each parameter contributed to the urgency of a sound. Unlike the present investigation however, the urgency levels that she selected were not equated. For each parameter they were selected on the basis of a paired comparisons procedure . Therefore, although the levels of urgency within each parameter were rank ordered, there was no indication of how the levels related between parameters. It is possible for example that all of the levels of loudness could have been more urgent than any of the speed levels. Because there was no equality between the different levels of the different parameters her finding that they accounted for different amounts of the variance in judgement was not very informative.

Momtahan defends the fact that she did not equate urgency levels between parameters by stating that perceived urgency is a culturally determined concept. She states that Stevens Power Law, the means by which urgency can be equated, is not usually applied to culturally determined parameters. As she points out herself it can however be argued that parameters such as speed which has been used to communicate urgency may be resistant to cultural influences. Although other parameters that have been used to communicate urgency such as pitch are more obviously cultural, it should be noted that similar psychophysical measurements have been made of parameters such as sweetness and brightness which are also to some extent culturally determined. Furthermore an exponent that is culturally determined is no less useful than any other exponent or measurement. At a particular time and for a particular culture it provides one of the most useful measurements of sensation. The historical and cultural setting is one of the factors that may determine all experimental observations, it should be no surprise

that is could be one of the factors determining an exponent. The issue of generalisability is one that touches all areas of experimentation.

Since it has been predicted that the mean judgements to some of the stimuli will be the same, and that the different levels of the different parameters will contribute equally to urgency judgements, it is predicted that there will be no effects between parameters, only between urgency levels. It was therefore necessary to consider the power of the proposed experiment to counter the possible criticism that no effect had been found, and so the predictions had been supported, only because the experiment was not powerful enough. The power was calculated by the method shown in Appendix 5B for the matched sample t Test. This procedure was chosen as the t-test could be used to compare means in this experiment. It was considered that it was a simple formula for providing an approximation of the power of the experiment. With the experimental power of 3.70, there was only a 4% chance of failing to detect an existing effect.

The proposed experiment will test the generalisability of previous results, the equal units of urgency, and will investigate the effects of combining acoustic parameters. If the predictions are not borne out then it is possible that when parameters are combined and all levels are equal one parameter emerges as the most salient or distinctive for communicating urgency. When warnings are constructed that must convey a variety of messages then that parameter should be used to communicate urgency. Similar investigations can be conducted on other applicable sound messages so that the best parameter can be used to convey each message.

5.2. Method

5.2.1. Subjects

The sample size was selected on the basis of calculations of experimental power, (Appendix 5B). Fourteen male and twenty six female subjects volunteered to participate in the study in partial fulfilment of their coursework requirements. All were first year undergraduate psychology students at Polytechnic South West, their ages ranges from 18-45 years. Two of the subjects had previously participated in similar experiments. All subjects reported having normal hearing.

5.2.2. Materials

The laboratory and hardware arrangements were as described in Experiment 1.

Three levels (high, medium and low urgency) of three acoustic parameters, speed (pulse rate), pitch and repetition units were co-varied in each experimental stimulus. The stimuli are detailed in Appendix 5A. 85, 113 and 150mm were selected as the low, medium and high urgency levels for each parameter.

The adjusted matching functions were used to determine the acoustic parameter values corresponding to each urgency level. For example, in the case of pitch;

$$\log o = \log (85/18.89)/0.349 = 74\text{Hz} \qquad \text{(Equation 5.4)}$$

Thus the pitch value that would produce an urgency of 85mm is 74Hz. This calculation can be approximately verified by reference to Fig. 4.1, remembering that the slope of the matching function has been adjusted.

Three pulses were constructed with their fundamental frequency at low (74Hz), medium (168Hz) and high (378Hz) urgency levels. These values were calculated as corresponding to urgency levels of 85, 113 and 150mm with the adjusted matching function from Experiment 5. The pulses were used to define the pitch levels of the bursts. In construction the fundamental frequency of these pulses and bursts were rounded up or down to the nearest 10Hz by the Tandon.

The three levels of speed were determined by the adjusted matching function from Experiment 4. The calculated low (7.338), medium (9.225) and high (11.65) urgency values are expressed in terms of pulse rates. The pulse rates were divided into 2500 ms., the maximum stimulus length of the stimuli from which the calculations were made. This gave pulse to pulse times, (ms from the start of one pulse to the start of the next), which can be used as an objective measure of speed when stimuli are not all the same length, as in this study. Larger pulse to pulse times represented slower stimuli. The pulse to pulse times were 340ms, 270ms and 214ms for the low, medium and high urgency stimuli respectively.

The low (1 unit), medium (2 units) and high (3.5 units) levels of repetition were calculated from the adjusted matching function from Experiment 6. In Experiment 6 repetition was communicated by a 33% drop in pitch of the second pulse of the repetition unit, from 300Hz to 200Hz. In the present study, the units

were defined in the same way. The pitch of the first unit was determined by the pitch level of the stimulus, 70Hz, 170Hz or 380Hz. The second pulse of the unit was at a pitch one third lower than that, 46.5Hz, 113Hz or 253Hz. In construction the Tandon rounded these pitches to the nearest 2.5Hz. The units are therefore constructed in the same way as in previous experiments.

It could be argued that the pitch level of the bursts is changed by the introduction of the lower pitches, so that the pitch no longer corresponds to urgency levels of 150,113 and 85 mm from the original graphs, but to lower levels. The pitch level of the burst is thought to be preserved because the level is defined by the first and the highest pitch that subjects hear and is thus salient on two counts. Furthermore, working through with the matching function for pitch shows that the urgency of the lower pitched pulses are 33% increases of each other - the relationship between the pitch levels is preserved in the low components of the repetition units. Even if the argument that the absolute pitch level of the stimuli has been changed holds, the stimuli are all still 33% more urgent than each other and so the size of the urgency changes they communicate is unchanged.

All possible combinations of the high medium and low levels of the different parameters were combined to make 27 experimental bursts. Three practice bursts were also constructed.

5.2.3. Procedure

Subjects were run one at a time while seated at a desk in Laboratory Two, approximately 1 metre from the speaker. They were told the broad nature of the study and were asked to read the following instructions (adapted from Engen, T(1971);

"I am going to present you, in irregular order, a series of sounds. Your task is to tell me how urgent they are by assigning numbers to them. When you have heard the first sound, give its urgency a number - any number that you think appropriate. I will then present another to which you will also give a number, and a third etc. Let high numbers represent high urgency and low numbers represent low urgency. Try to make the ratios between the numbers that you assign to the different sounds correspond to the ratios between the urgency of the sounds. In other words try to make the numbers proportional to the urgency of the sound as you hear it. Remember that you can assign any number. There is no limit to the

number that you assign. There is no right or wrong answer. I want to know how you judge the urgency of the sounds.
Any questions?"

In this study the free modulus magnitude estimation procedure previously recommended, (Section 3.6), as the least biased method of scaling urgency was employed. It was used in preference to cross modality matching because in this study it was predicted that no, or very small differences between urgency levels would be perceived. It was felt that subjects would be better able to make fine discriminations using a familiar medium, numbers, rather than with line length.

When subjects were ready to begin, the experimenter sent the first stimulus from the Tandon to the speaker in Laboratory Two. When the subject indicated that he or she was ready the next stimulus was sent and so on. The first three stimuli that each subject heard were the practice Bursts G, H and I, the responses to these bursts were excluded from analysis. After the practice stimuli, the 27 experimental bursts, 36-62, were played twice each in a different random order to each subject. Each subject made 57 judgements.

When subjects had completed the task they were thanked, debriefed and allowed to leave. Their comments on the study were recorded.

5.3. Results

It was predicted that the stimuli would be ranked from most to least urgent - HHH, HHM, HHL HMM, HML MMM, HLL MML, MLL, LLL, irrespective of which parameter was at which level in each of the combinations. These predictions were calculated by considering the high (150), medium (113) and low (85) urgency values of the parameters. These values were obtained when subjects were responding using ratios and were thus from a logarithmic scale. The logs of the urgency values were therefore taken, which were, high (2.176), medium (2.053) and low (1.919). These values were equally spaced on a logarithmic graph and could be added up according to the different combination of levels to derive the predicted urgency order if the different levels are equal across the different parameters. Thus a HHL stimulus, (2.176+2.176+1.919 = 6.27), is predicted to be as urgent as a HMM stimulus, (2.176+2.053+2.053 = 6.28).

The untransformed means and standard deviations of the subjects responses to each stimulus are shown in Appendix 5C. As in previous experiments, Engen's(1971) data transformation to eliminate inter and intra subject variability was applied to the data. The transformed mean scores by each subject to each stimulus are shown in Table 5.1. The rankings in this table are from most (1) to least (7) urgent.

17/27 stimuli were ranked in the predicted order. A Spearmans' rank correlation coefficient between the predicted and obtained mean values for each stimulus was 0.901, this was significant (p=0.01, see Appendix 5D).This provides preliminary support for the hypothesis that the three urgency levels were equally urgent in all three parameters and that the same combination of levels were judged the same regardless of which parameters were at each level.

This data is further described by Fig. 5.1. Each point on the graph represents the mean urgency level for one parameter, collapsed across the other parameter levels. For example, the top left hand point on the graph represents the mean judgement to all of the stimuli in which pitch was high. Fig. 5.1 indicates that there were the predicted urgency trends for the three levels of each parameter. Furthermore it shows that while judgements to the different levels of speed and repetition were similar, high levels of pitch were

judged higher than high levels of other parameters and low levels of pitch were judged lower than low levels of the other parameters. This suggests that pitch levels are contributing more than speed or repetition levels to urgency judgements.

Fig. 5.1 can be used to investigate the relationship between the high, medium and low urgency values between and within parameters. For pitch the medium level is judged 34% less urgent than the high level, and the low level is judged 39% less urgent than the middle level. For speed, the medium level is judged 18%less urgent than the high level and the low level is judged 17% less urgent than the medium level. For repetition, the medium level is judged 16% less urgent than the high level and the low level is judged 29% less urgent than the medium level. Examination of Figs. 5.2a-5.2c shows mean responses to each level of each parameter, collapsed across the other parameters.

PREDICTED RANK ORDER.	PARAMETER LEVELS.	LOG MEAN.	ANTI LOG MEAN.
1	PHSHRH	1.339	21.87
2	PHSHRM	1.277	18.95
2	PMSHRH	1.246	17.63
2	PHSMRH	1.232	17.07
5	PHSLRH	1.186	15.34
5	PHSMRM	1.184	15.29
11 *h	PHSLRM	1.138	13.75
5	PHSHRL	1.109	12.87
5	PMSHRM	1.099	12.57
11 *h	PHSMRL	1.093	12.40
5 *l	PMSMRH	1.088	12.26
11	PMSMRM	1.031	10.75
18 *h	PHSLRL	1.018	10.44
11	PMSLRH	0.976	9.47
5 *l	PLSHRH	0.954	9
11	PMSHRL	0.935	8.62
18 *h	PMSLRM	0.905	8.03
11 *l	PLSHRM	0.901	7.97
11 *l	PLSMRH	0.892	7.80
18	PMSMRL	0.847	7.03
18	PLSMRM	0.818	6.58
18	PLSLRH	0.786	6.10
24 *h	PMSLRL	0.774	5.94
24	PLSLRM	0.706	5.08
18 *l	PLSHR	0.686	4.86
24	PLSMRL	0.686	4.82
27	PLSLRL	0.623	4.20

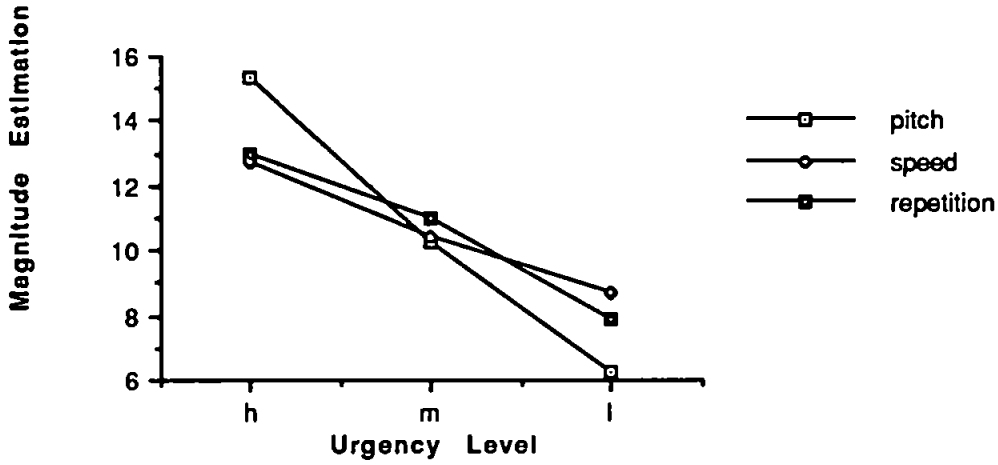
P= PITCH, S= SPEED, R=REPETITION.

H=HIGH URGENCY, M=MEDIUM URGENCY, L=LOW I=URGENCY.

*= STIMULUS RANKED HIGHER (h) OR LOWER (l) THAN PREDICTED

TABLE 5.1 : MEAN JUDGEMENT TO EACH STIMULUS AFTER TRANSFORMATION.

Fig.5.1: The mean judgement when each parameter is at each level



A three way (Speed h,m,l; Pitch h,m,l; repetition h,m,l) repeated measures Anova was conducted upon subjects untransformed urgency judgements. The purpose of ths was to see, by comparing the F ratios for each parameter, whether each parameter was contributing equally to the urgency judgements. As is shown in Table 5.2, and as predicted, there were significant main effects of urgency level on all three parameters; speed ($F(2,78)=25.06$, $p=0.00$), pitch ($F(2,78)=16.87, p=0.00$), and repetition ($F(2,78)=19.53$, $P=0.00$). Parameters were judged more urgent as their urgency levels increased. Three two-way significant interactions were found, between pitch and speed ($F(4,156)=2.995, p=0.02$), speed and repetition ($F(4,156)=6.727, p=0.00$) and pitch and repetition ($F(4,156)=2.553, p=0.04$). There was no significant speed*pitch*repetition interaction ($F(8,312)=1.218, p=0.287$). The significant interactions are plotted in Figs. 5.3a-5.3c in order that they might be more fully investigated.

It should be noted that the full degrees of freedom as reported in the anova assume that each independent variable is totally independent. This was not so in our design because in a within subjects factorial design responses to different levels of the same factor will always be correlated. If conservative degrees of freedom (where one digit of the ratio remains at 1) are employed then only the main effects and the interaction between speed and repetition remain significant at $p=0.01$. The 'truth' in terms of significance lies somewhere between the two since the conservative degrees of freedom assume a perfect correlation between responses which is obviously not obtained.

A multiple regression was employed to see how well urgency could be predicted on the basis of speed, pitch and repetition, according to the formula,

$$\text{Urgency} = b0 + b1\text{speed} + b2\text{pitch} + b3\text{repetition} + b4\text{interactions} \qquad \text{(Equation 5.5)}$$

Each regression coefficient, (b1,b2,b3), represents the change in urgency resulting from each parameter. The model is linear and so the logarithms of the urgency judgements were used to fit into the equation because the urgency levels were equally spaced on a log. graph.

The urgency levels for each parameter were coded 1(low), 2(medium) and 3(high). The codings represent the fact that the log stimulus levels were equally spaced. Some of the parameter levels had to be rounded up and down in stimulus

Fig. 5.2a:
Subjects Mean Responses
for each Level of Pitch.

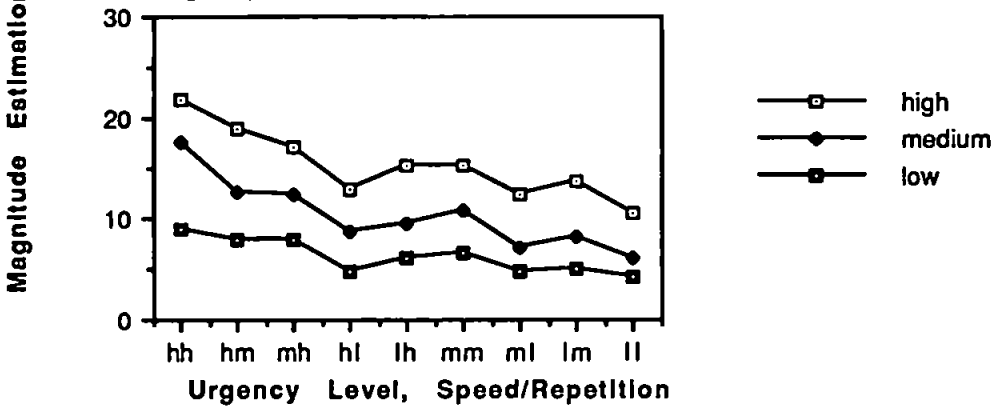


Fig. 5.2b :
Subjects Mean Responses
for each Level of Speed

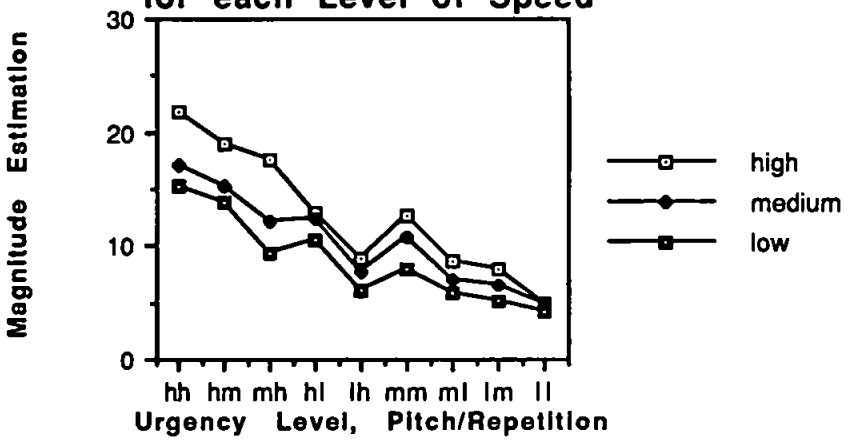
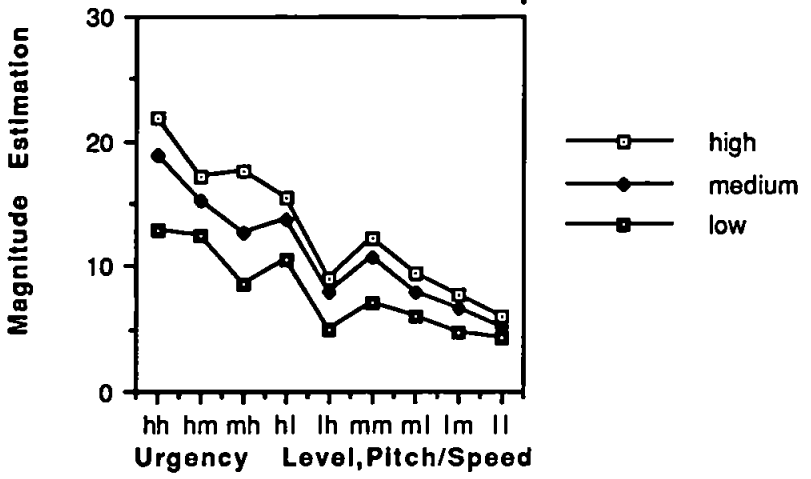


Fig. 5.2c:
Subjects Mean Responses
for each level of Repetition



SOURCE	SUM OF SQU	df	MEAN SQ	F	P
Subjects	758001.1	39	19435.9		
Speed	8528.4	2	4264.2	25.06	0.00
Speed*Subject	13272.5	78	170.16		
Pitch	51822.6	2	25911.3	16.87	0.00
Pitch*Subject	119759.3	78	1535.3		
Repetition	18060.6	2	9030.3	19.53	0.00
Repetition*Subj	36056.4	78	462.2		
Speed*Pitch	710.6	4	177.6	2.99	0.02
Speed*Pitch*Subj.	9252.0	156	59.30		
Speed*Rep.	2379.4	4	594.8	6.72	0.00
Speed*Rep.*Subj.	13793.1	156	88.41		
Pitch*Rep.	1092.7	4	273.1	2.55	0.04
Pitch*Rep.*Subj.	16689.4	156	106.9		
Speed*Pitch*Rep.	738.4	8	92.30	1.21	0.28
Speed*Pitch*Rep* Subj.	23628.0	312	75.73		

TABLE 5.2 : THREE WAY ANOVA, SPEED*PITCH*REPETITION.

PARAMETER	COEFFICIENT	P=
SPEED	0.03666(b1)	0.043
PITCH	0.18917(b2)	0.00
REPETITION	0.05863(b3)	0.003
S*R	0.02235(b4)	0.012

TABLE 5.3 : REGRESSION COEFFICIENTS FOR SIGNIFICANT PREDICTORS OF MEAN LOG. URGENCY.

Fig. 5.3A : Pitch * Speed Interaction

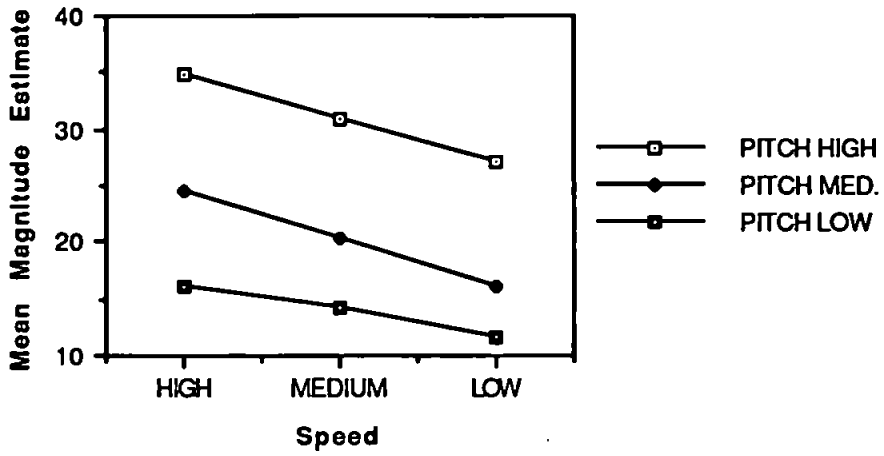


Fig. 5.3B : Speed * Repetition Interaction

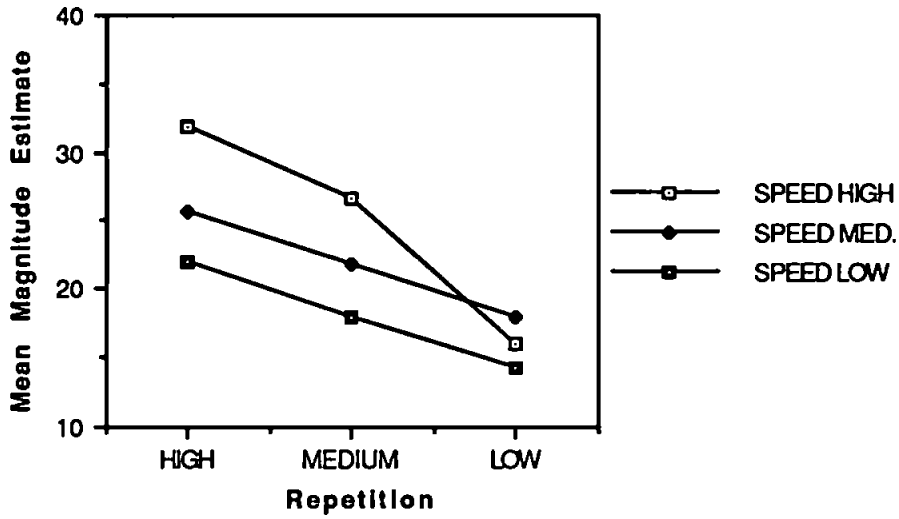
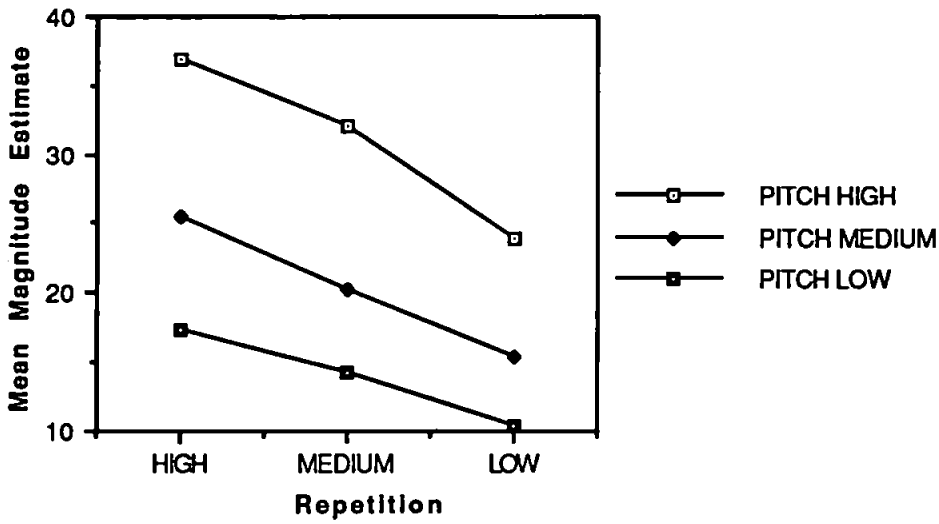


Fig. 5.3C : Pitch * Repetition Intercation



construction, and thus they were no longer exactly equally spaced. The coded values for these parameter levels were appropriately adjusted. The exact parameter value was fitted to a linear equation with the appropriate urgency coding as the known value of y. This was done for two levels of urgency and the equations solved simultaneously. Thus for example in the case of pitch, where the equally spaced parameter values were 74 and 169Hz for the low(1) and medium (2)levels of urgency respectively, the equations took the form,

$$1 = k1\ 74Hz + k2$$
$$2 = k1\ 168Hz + k2$$

(Equations 5.6, 5.7)

Solving the equations allowed the rounded parameter values to be substituted so that the appropriate urgency codings, instead of 1, 2 or 3 could be substituted.

The coded parameter values were regressed as predictors of mean log urgency (Appendix 5E, Full Model). All the interaction terms that were significant in the Anova were included in the equation. Interactions that were not significant after this 'first run' were excluded and mean log urgency was regressed against speed, pitch, repetition and the significant speed*repetition interaction, (Appendix 5Ea). Table 5.3 shows the regression coefficients for these significant predictors of log. urgency.

As predicted repetition and speed had similar sized coefficients, as did the repetition*speed interaction. Pitch had a larger regression coefficient. This implies that pitch levels were contributing more to urgency judgements than levels of the other three parameter combinations.

The method described by Neter(1985) was used to see if the regression coefficients were the same. The data was fitted to three reduced regression models in which speed and pitch, speed and repetition and pitch and repetition were combined together as one predictor (Appendix 5E, Reduced Models). The method described in Appendix 5Eb computes whether error is increased by reducing the model. It tests Ho : b1=b2, b1=b3, b2=b3. As is shown, the coefficients for speed and repetition were not statistically different from each other at p=0.01, but they were at p=0.05. Contrary to the predictions, the other coefficients were significantly different from each other, (p=0.01). The prediction that there was no difference between the regression coefficients was rejected. This indicates that

the different parameters were influencing urgency judgements by different amounts.

5.4. Discussion.

There was preliminary support for the hypothesis that the urgency levels of the three acoustic parameters were equal because most of the stimuli were ranked in the predicted order. However, a visual examination of the ten stimuli that were not ranked in the predicted order, Table 5.1, shows that it was the value of the pitch parameter that determined whether the stimulus was ranked higher of lower than predicted. This indicates that pitch may have been contributing to judgements more than speed or repetition. This notion was further supported by examination of Figs. 5.1 and 5.2a-5.2c.

Fig.5.1 showed that the differences between urgency levels is larger for pitch than for speed or repetition. For pitch the high medium and low levels are almost equidistant, 34% and 39% apart. This is close to the predicted 33% separation. For speed and repetition the high medium and low levels are equidistant as predicted, except for the medium to low levels of repetition, but are separated by less than the predicted 33% difference in urgency. Figs. 5.2a-5.2c show that when either speed or repetition are combined with low levels of the other parameters, judgements are greatly lowered. For pitch however having the other two parameters at a low level has far less of a lowering effect. Similarly in Figs. 5.2b and 5.2c there is a noticeable dip in the middle of the graphs when pitch is low, regardless of the level of either of the other parameters. These figures add weight to the suggestion that pitch is adding more to the urgency judgements than the other two parameters, which appear to be approximately equal.

Although a visual examination of the data had suggested that pitch was having a larger effect than the other parameters on urgency judgements, the Anova, (Table 5.2), showed that the largest F ratio was for speed. Examination of the raw data however showed that the largest effect in terms of absolute mean differences was for pitch. Pitch also had a larger error term than the other factors and this reduced the F ratio. In an Anova the error term is partitioned out between factors so that each factor has its own error term which reflects the treatment*subject interaction. Although pitch had a large absolute effect it was highly variable between subjects. This variability increased the error term and reduced the F ratio. The magnitude of the F ratio therefore takes into account the variablity of

the effect. Under these terms speed is shown to have the largest effect on judgement as it has the largest F ratio. Pitch and repetition have F ratios of a similar size. This indicates that these two parameters have a similar effect on urgency judgements when variability is taken into account.

A multiple regression was employed to further test the hypothesis that the urgency levels of the three acoustic parameters were equal. When the full model was regressed with all the regression coefficients the coefficient for speed was not significant, Appendix 5E. This is because the speed* repetition interaction (to be discussed) was so significant. For the effect of speed to be significant it would have had to have a consistent effect across all levels of the other parameters. When the non-significant interactions were omitted from the regression the speed effect was significant.

The results of the multiple regression did not support the prediction that the regression coefficients would be the same for all of the parameters. Speed and repetition had similar sized coefficients that were not significantly different at p= 0.01. This finding supports the visual inspection of the data which indicated that pitch was having the largest effect on urgency while speed and repetition were having effects that were a similar size. However the coefficients were significantly different at p=0.05.

In order to investigate further the regression coefficients an extended regression model was fitted to the data which included subjects factors. This model was fitted to see if very much information had been lost by taking the means of subjects' data. Since the results of the extended model were no different from the results of the full model it appears that no information was lost by regressing using the means. Since the extra power of the extended model is not required (the significant effects are shown by the full model) the full model is considered to be sufficient.

The significant interactions between the parameters were plotted in Figs. 5.3a-5.3c. The interaction in Fig. 5.3a, between speed and pitch, indicates that when pitch was low, speed was having less of an effect than when pitch was medium or high. This implies that low pitch was having a stronger effect than other parameters and levels and was not allowing the high medium or low levels of speed to have an influence. This is supported by the previous observation (Fig. 5.1) that low pitch was judged lower than the low levels of the other parameters, and by Fig. 5.2a which shows that low pitch has the straightest line, thus other parameters

are having less of an effect. This visual inspection of the interaction adds further weight to the suggestion that pitch has a larger affect upon judgement than other parameters.

The largest interaction was between speed and repetition, Fig. 5.3b. In this case the low level of repetition is having an especially large effect of the high level of speed, greatly reducing its urgency, so that it becomes less urgent than the medium level of speed combined with the low level of repetition.

An explanation of the interaction is implied by the observation that the stimulus with the high level of speed and low level of repetitions is shorter than any other stimulus (428 ms). It is possible that the stimulus was too short to convey the impression of fast speed and so much of the urgency of the stimulus was lost. The fast stimuli could be heard as such in the stimuli with more units of repetition since they were longer. In the other short stimuli with few units of repetition the speed was slower, there were thus larger interpulse intervals which would make it easier to hear out the separate pulses and thus preserve the impression of speed. This effect could have been compounded by that fact that shorter stimuli are perceived as being less urgent (Hellier and Edworthy 1989).

Further support for this idea is found in the perceived duration literature. Eisler(1974) averaged the exponents for duration from many different stimuli and many different methodologies and stated that it was 0.9. This means that perceived duration is consistently underestimated. The exponent of less than 1 results in a negatively accelerated function in a linear plot, so that durations at the bottom end of the scale are more underestimated than those at the top. This means that the very shortest stimulus in this study would have been more underestimated than any others.

To summarise, it seems that the stimulus with a high level of speed and a low level of repetition may have been too short to adequately convey an impression of speed. This resulted in the urgency of the stimulus being reduced. This reduction in urgency may have been accentuated by the low actual and perceived duration of the stimulus which would have further reduced its urgency.

In the interaction between pitch and repetition, Fig. 5.3c it seems that the high level of pitch is being made less urgent by the low level of repetition. In this interaction there is also a slight reflection of the first interaction where a low

level of pitch prevented the message from the other parameter, here repetition, being communicated.

Although all of the regression coefficients were found to be significantly different in size, the study has provided some general support for the equal units of urgency that were calculated on the basis of Experiments 4-6, and thus for the assumptions of the Power Law and cross modality matching. Firstly, the main effects for each acoustic parameter were in the predicted order. Secondly, most of the stimuli were ranked in the predicted order. Thirdly, plots of the data and regression coefficients appeared to be similar for speed and for repetition. This indicated that speed and repetition were contributing equally to the urgency judgements ie that at each level, urgency was equal. What is indisputable is the finding that pitch was contributing more than the other parameters to urgency.

To conclude, although the stimuli employed equal units of urgency between the three parameters, pitch as been shown to have a greater influence than the other parameters upon urgency judgements. This implies that pitch is the most salient or discriminable parameter to use for communicating urgency. This idea is supported by two observations. Firstly, of the parameters employed, pitch had the smallest exponent. It thus took larger changes in pitch than in any other parameter to communicate the increases in urgency. These large changes may have drawn attention to pitch and made the difference between the levels more discriminable. Secondly, according to Stevens and Galanter(1957), pitch is a Metathetic continuum. Changes in pitch thus alter the quality of the stimulus, and such changes may be more discriminable than changes to Prothetic continua where the quantity of the stimulation is altered.

CHAPTER SIX.

DETERMINANTS OF PERCEIVED URGENCY.

6.1. Introduction.

It has already been demonstrated that variations in some acoustic parameters result in variations in perceived urgency. In this Chapter, an attempt will be made to highlight factors that may determine whether or not a parameter variation will affect perceived urgency. That is, to show what it is about a stimulus that makes it urgent or non urgent, what the determinants of urgency are. In particular, the time perception literature will be reviewed to investigate the idea that the concepts of urgency and time are related. Ways in which variations in perceived duration might be expected to affect perceived urgency will be discussed. A glossary of time perception terms is presented in Appendix 6A.

There are several possible non acoustic determinants of perceived urgency that could be investigated. The predictability of a stimulus may have an effect, with more unpredictable stimuli being perceived as more urgent. This idea receives support from the finding that unresolved stimuli are judged more urgent that resolved ones (Edworthy et al 1989), and from the finding that more inharmonic stimuli are judged more urgent. Predictability cannot account for all perceived urgency effects however because stimuli that vary in speed, pitch or repetition do not seem to vary in predictability and yet they can alter perceived urgency. Furthermore, Edworthy et al(1989) found that regular, more predictable, rhythms were perceived as more urgent than irregular and thus less predictable ones.

Another possibility is that perceived urgency is determined by evolutionary factors. Warren(1981) proposed that physical correlates exist for subjective sensations, and that these physical correlates are often innate or evolutionary in origin. It is easy to see that increases in pitch which result in increases in urgency may be correlated to animal distress cries that rise in pitch with increasing distress, and thus communicate increasing urgency in evolutionary terms. It is also possible to see how a stimulus getting faster could be correlated to the sound a stimulus moving closer to the subject and thus becoming more immediate or threatening. For other parameters such as repetition, resolution,

inharmonicity and length however it is more difficult to imagine what their physical correlates might be, and difficult to imagine the evolutionary basis for the changes in urgency that correspond to changes in the parameter level.

A third possible determinant of perceived urgency, and one that may also be evolutionary in origin, is arousal. In evolutionary terms arousal is associated with the flight or fight response of a threatening situation. It is possible that stimuli that are arousing are perceived as being more urgent for that reason. This is a contentious point however and one that is very difficult to investigate. It is not clear why certain parameter changes would increase arousal, whether the stimulus would be expected to be emotionally or physiologically arousing in order to affect perceived urgency, or how this stimulus -related arousal might interact with the situational arousal derived from the working environment in which an alarm sounds.

Perceptions of urgency must also to some extent be determined by the subject or operators' previous experience of alarm sounds. For example a nurse who has worked in an ICU where many of the alarms are high in pitch will have learnt to associate high pitches with increasing urgency. This factor, while being an important consideration when alarms are modified or introduced into a workplace, is not especially relevant to the present research. By revealing the natural determinants of urgency, it will be possible to design alarms that can be used to train the new generation of operators in which the urgency of the alarm is communicated through the sound itself and does not have to be learnt in the same way.

A final possible non acoustic determinant of perceived urgency, which will be investigated fully in this Chapter, is the apparent passage of time or perceived duration. Besides the fact that duration is an ongoing stimulus attribute and so actual and perceived duration must affect urgency judgements to a stimulus, there are several other reasons why time was selected for further investigation in preference to the other factors as a possible non acoustic determinant of urgency.

Firstly, it was thought that perceived duration might be an important aspect of urgency because the concept of urgency involves the amount of time available in which to act. Something that is urgent must be attended to faster, and thus duration is a direct analogue of the response time to an urgent stimulus.

Secondly, the experimental work previously conducted on stimuli that vary in speed supports this intuitive notion. It was shown that speed is an important determinant and communicator of urgency, and that something that was faster was perceived as being more urgent. Although pitch was shown in Chapter Five to be the most discriminable or salient communicator of urgency, speed is also considered to be an important parameter to investigate since it is the most economical parameter through which to communicate urgency and because responses to speed stimuli are less variable than those to pitch stimuli (Chapter Five) . Furthermore, if pitch changes are discriminable because the quality of the stimulus changes due to pitch being a Metathetic continua, then this calls into question the way in which pitch was scaled in Chapter Four because ratio scaling techniques are inappropriate for Metathetic continua. This issue is not in debate for the speed stimuli which are, in these terms, obviously Prothetic and have thus been appropriately scaled.

Speed is a concept that is closely linked to the concept of time for to perceive something as being 'fast' requires an understanding of the item's relative duration. The fact that speed influences perceptions of urgency may again indicate that the notion of time is involved in those perceptions.

If perceived duration is a determinant of perceived urgency then it is possible that factors which influence perceived duration may influence urgency perception by affecting the perceived length or the perceived speed of a stimulus. For example, a factor which increases perceived duration, so that a stimulus appears to be longer than it is, may make that stimulus appear to be slower than it is because the events within it would be appear to occur over a longer period of time. This may either make the stimulus appear less urgent by making it slower, or more urgent by making it longer.

A review of the time perception literature was conducted to further investigate the possible link between time and urgency perception. Fraisse(1978) pointed out that time perception studies have looked at many different aspects of perceived duration. They have looked at the order of events (judgements of which events come after which other events); the duration of events (judgements of how long events are); and at the interval between events (judgements of whether an inter stimulus interval exists or whether events are simultaneous). Time perception thus involves the concepts of succession (the notion that one event follows another) and duration (the length of an event or an isi). Although Luce(1985)

said that mean time judgements for humans and animals are relatively accurate, actual and estimated duration are rarely exactly the same because systematic errors occur as a result of different tasks, attitudes, methodology etc. These differences between objective and subjective time have prompted many psychophysical investigations into time perception using duration scaling and duration discrimination tasks. According to Hogan(1978) the literature remains unclear as to whether different people have different time perception or whether different stimuli during the judgement interval lead to differences.

The aforementioned factors will be investigated in the following chapter where they are relevant to the construction of auditory warnings and perceived urgency. Only the findings that relate to intervals of the same approximate duration as warning sounds will be included, research pertaining to very short or very long intervals will not be discussed. A summary of the findings of time perception studies will be followed by a look at the different theories that attempt to explain those findings. By examining the theories it will be possible to see which, if any, of them could explain or predict urgency perception in terms of changes in perceived duration .

6.2. Background Findings.

6.2.1. Methodology.

One factor that may influence the outcome of time perception studies is the methodology employed and phenomena associated with the methodology.

6.2.1.1. Scaling Technique.

Estimation, production, comparison and reproduction are the measures usually employed in time estimation studies. Measures yielding binary/tertiary choices eg forced choice or single stimuli are also sometimes used, as is category rating. The many different methodologies make it hard to compare studies, and attempts have been made to see if methodology systematically influences the results of time perception studies.

Fraisse et al(1962) and Treisman(1963) are two of the many authors who have compared methods. Generally estimation and production are found to have negatively correlated errors and reproduction is found to have the smallest inter-

subject variability. There is little consensus among findings as to the direction and magnitude of constant errors in different methods. Kruup(1971) found that the highest duration estimates resulted from production, then reproduction, then estimation. The smallest inter-subject variability was associated with reproduction. In 1971 McChoncie et al found that the percentage constant error (an accuracy measure, the difference between judgement and the standard measured as a percentage of the standard) was smallest for reproduction, had a mid value for estimation and was largest for production. The same results were found for intra - subject variability, and inter- subject variability was smallest for production. Estimation and production judgements were negatively correlated, and reliability was about the same between the methods. Allan(1979) reviewed the literature on methodological differences and concluded that no one methodology could claim superiority in terms of inter subject variability or accuracy.

6.2.1.2. Time Order Error.

Another factor to be considered in the interpretation of time estimation studies is time order error. Fechner (1860) first noted that judgements could be affected by the order of presentation of the stimuli. This phenomena was noted in reproduction and in comparison tasks, where the stimulus to be reproduced or compared follows another one. It is referred to as time order error and must be taken into account when time perception studies are interpreted. Fraisse(1978) explained that the error was usually negative, that is the second stimulus that was presented to the subject was overestimated. He claimed that the error could be avoided by randomising the order of stimulus presentation between subjects in estimation tasks.

6.2.1.3. Stimulus Factors.

There are two main aspects of the stimulus that may influence perceived duration.

The first important stimulus attribute is familiarity. Avant et al(1975) found that perceived duration was decreased when the stimulus was familiar. They proposed that stimulus familiarity facilitated automatic contact between the stimulus and its memory representation, so reducing perceived duration. McCray(1969) on the other hand found that familiarity increased perceived duration, as did Kowal(1987) for melodic sequences. Kowal interpreted these

effects in terms of apparent numerosity, he said that familiar melodies allowed the regeneration of more notes and thus perceived duration increased.

Stimulus intensity has also been shown to influence perceived duration. Ekman et al(1966,1969), Bergland et al(1969), Zelkind(1973), Goldstone et al(1978) and Fraisse(1978) all found that perceived duration increased when the stimuli were more intense. A model by Hawkins et al(1979) suggested that the nature of judgement interacted with stimulus intensity to affect perceived duration. Nisly et al(1989) reviewed the literature and supported this idea. In 1988 Bringer found that higher pitched stimuli increased perceived duration relative to lower pitched ones.

6.2.1.4. Summary

To summarise, there is little conclusive evidence on the superiority of one scaling technique over another. There are several phenomena that could have implications for warning design if a link between perceived duration and perceived urgency is established. However, in the main these appear to be methodological effects and so, providing the normal experimental controls are employed, ie. randomisation of stimulus presentations and random sampling of subjects, they should have little impact on the relationship between duration and urgency or on warning design.

6.2.2. Effects.

It is important to be aware of some of the common effects in time perception studies to provide a background against which the theories can be evaluated.

6.2.2.1. Indifference Interval.

Fraisse(1975) looked at the relationship between perceived duration and real duration and found that many authors eg Horing(1864), Woodrow(1934), reported that although short intervals were overestimated and long ones were underestimated, in the middle was an 'indifference zone' in which durations were estimated accurately. Wundt claimed that is was the interval around 0.7sec. that was most accurately reproduced, the indifference interval. If changes in perceived duration do affect perceived urgency then the indifference interval may

be an important consideration in warning design, affecting perceived duration and thus the urgency of the warning.

Fraisse(1975) noted that the indifference interval corresponded to the time taken to swing a leg when walking fast, and hypothesised that the most accurately reproduced intervals could have developed from body movements. Although Fraisse supported approximately Wundt's indifference zone, claiming that 0.75 sec. corresponded to the complete process of perception, Woodrow(1966) pointed out that people have rarely found the same indifference interval, citing between 0.36ms-5sec. Woodrow claimed that the indifference interval could be influenced by, for example, markers, attitude and instructions. He also pointed out that some drugs eg Mescalin can sometimes produce huge overestimations and sometimes huge underestimations.

Other authors such as Hollingworth(1910) claimed that the indifference interval was a methodological effect, resulting from central tendency. He felt that central tendency occurred when long intervals were underestimated and short ones overestimated because judgements were shifted towards the median of the stimulus range.

Treisman(1963) disputed the suggestion that underestimation of long intervals and overestimation of short ones occurred as a result of central tendency. He cited Woodrow(1934) who used one subject per judgement and still got the characteristic under and over estimations and the indifference interval. This refuted Hollingworth's suggestion that the effects were caused by central tendency because when only one judgement per subject was used there was no median of the stimulus range for judgements to be shifted towards. Treismans' own experiments further refuted Hollingworth for he found that the indifference interval did not settle at the median of the range, nor did it decrease when the stimulus range was shifted down. Both of these effects would have been predicted to occur by Hollingworth. In fact Treisman found that the indifference interval increased as sessions progressed, he termed this increase in judgements over the session 'lengthening', and demonstrated it in six out of seven of his experiments. Lengthening was also demonstrated in reproductions by Brown et al(1965) and by Von Sturmer(1966).

6.2.2.2. Filled Interval Effect.

The existence of the filled interval effect implies that nontemporal information is important in determining time perception. This supports the notion that perceived duration and perceived urgency are related because in auditory warnings, non temporal information is varied to communicate changes in urgency. The filled interval effect implies that that information may also communicate changes in perceived duration.

There is little agreement as to the nature of the filled interval effect. Although Gavin(1959) claimed that it was an individual difference, with some people overestimating filled intervals and some overestimating empty ones, most authors prefer to claim that the effect lies in one direction. Rai(1973) and Lordahl(1973) both said that filled intervals usually produced shorter estimations than empty ones, perhaps by as much as 33%, whereas Schiffman et al(1977) found that filled intervals were perceived as being longer. Buffardi(1971) showed that the filled interval effect was not modality specific. Although the precise nature of the effect on time perception has not been established, there are only a few authors, for example Deehring(1961) who claim that there are no significant differences between temporal judgements of filled and unfilled intervals.

6.2.2.3. Psychophysics.

People have tried to see if time perception obeys Webers' Law (JND between two durations a constant proportion of the shorter duration). Allan et al(1974) found little evidence to support Webers Law, but Treisman(1963) supported it with the methods of estimation,production reproduction and single stimuli. Allan(1979) felt that if Weber's Law held at all for duration, it was Getty's(1976) more generalised form of the law that had been supported. Weber's Law predicts that the standard deviation of duration estimates in a scaling task is proportional to the mean duration estimate. The more general form of the law allows variability in estimates to arise from various sources, only some of which are dependant on stimulus magnitude.

More common is the claim of Stevens and Galanter(1957) that objective and subjective duration fit the power function. They claimed that the exponent for the

duration of white noise was 1.1. Since then many different exponents have been proposed. Bjokman et al(1960) said the exponent was 1.1 when they used adjustment, and closer to Ekmans'(1957) figure of 1.37 with reproduction. Michon(1967) found the exponent for short filled intervals was 0.6 whereas for longer ones it was 1.1. Schiffman et al(1977) also supported the power function with an exponent close to unity, as did Bobko et al(1977). Eisler(1976) arrived at 0.9 as the average exponent, in 1981 he found that 0.74 was the average exponent for duration discrimination data.

In 1988 Nakajima et al proposed a similar idea to the power law to describe duration experience. Their supplement hypothesis said that subjective duration of an empty time interval was proportional to the physical duration plus a constant (80ms). This model predicted the findings that subjective ratios between two durations were less extreme than the physical ones and that the j.n.d of an empty duration was proportional to the standard deviation plus 80ms.

Although Eisler(1984) was adamant that the power function was the internal representation of duration in humans and animals authors like Anderson(1971), Allan and Kristofferson(1974) and Allan(1979) said subjective and objective time were related by a linear function. Allan(1979) felt the power function was only exhibited when ratio setting data was used. She said exponents could not be directly estimated from ratio setting data and were thus dependant on the model employed by the experimenter. Divenyi et al(1987) claimed that if intervals were bounded by speech or speech-like sounds then no psychophysical law at all was obeyed by time perception data.

6.2.2.4. Summary

The preceding sections have considered some general findings. These findings provide a background that can be considered when theoretical accounts of time perception are evaluated as possible explanations for the effects of different sound parameters upon perceived urgency.

6.3. Time Perception Models.

There are four main categories of traditional time perception model. They are reviewed below.

6.3.1. The Internal Clock.

The idea that an internal clock controls time perception has been proposed by many authors. As early as 1934 Hoagland found that time estimation/body temperature and body temperature/ alpha rhythm were related, and suggested that the pacemaker for subjective time and for alpha frequency was the speed of molecular motion in cellular metabolism. This hypothesis was supported by, for example, Fischer et al(1962) and Cahoon(1969). Landis'(1925) finding that metabolic rate increased with arousal prompted the idea that arousal should increase cellular metabolic rate and thus subjective time rate, (see Glossary, Appendix 6A). This would lead to overestimations. Research since then, for example, Mundy et al(1953), Werboff(1957) and Cahoon(1969) has suggested that high arousal is accompanied by faster subjective time rate and results in overestimations. This has given weight to the idea that time perception could be governed by an internal clock.

Could

One of the first and most complete internal clock theories was proposed by Treisman(1963). He said that physiological arousal resulted in specific arousal of the internal clock, which increased subjective time rate. It was said that a pacemaker produced a constant stream of pulses, the rate of the pacemaker was affected by specific arousal. The pulses could be read into a store for later retrieval by the comparator(decision mechanism) or directly read into the comparator(which compared retrieved measures with current counts and selected a response). He said that specific arousal was like general arousal in nature and could be affected by meaningful aspects of the experimental situation. Treisman predicted that drugs that increased arousal speeded up the internal clock and resulted in overestimations and those that produced sedation would slow the internal clock and produce underestimations. These predictions were supported by Goldstone et al(1959).

Many authors have supported the notion of an internal clock, proposing their own models. Zelkind(1973) said the internal clock was based on the functioning of the

brainstem reticular formation, the brf, (which was said to alter the rate of the internal clock by regulating cortical activity). He supported his work with the finding that more intense auditory stimuli resulted in overestimations. This suggested that intense stimuli activated the brf, increasing cortical arousal so subjective time rate increased. He said the increases in stimulus intensity were too small to have caused such effects and said that arousal must amplify them. He said the effect of more intense stimuli was like the effect of increased metabolic rate (which speeds up the internal clock and leads to overestimations). Zelkinds' ideas were supported by Delay and Richardson(1981).

A derivative of the internal clock idea was Eislers(1975) Parallel Clock Model. This was devised with ratio setting tasks, especially reproduction, in mind. It assumes that subjective total duration (the first and second durations) are accumulated in separate sensory registers. The contents of the registers were thought to be compared until the moment of response. In reproduction for example the second interval would be terminated when the contents of the registers were equal. The model thus assumes that parallel processing occurs and is especially relevant to the emergence of parallel distributed processing, pdp, as an important concept in cognitive psychology. Eisler(1981) supported the model with Allans(1977) duration discrimination data, and with the behaviour of rats.

Church(1984) supported the idea of an internal clock in animals. He said that pulses were produced by a pacemaker, the rate of which could be varied by diet, drugs or stress, and that the accumulator held the sum of the pulses. A comparator was said to use the ratio between the value in the accumulator and the value in reference memory (information about past trials and their consequences) to determine a response on the basis of a decision rule.

Work in music perception led Povel and Essens(1985) to the conclusion that sequences of temporal patterns were assessed by an internal clock with a periodic pulse and a counter. They assumed that listeners tried to generate an internal clock while listening to a temporal pattern. The distribution of accented events within the pattern was said to determine whether a clock would, and which clock would, be generated. They said that the clocks were hierarchical, with medium length units determined by equally spaced rhythmic pulsing of simple music. The units could be divided or joined. The authors proposed that many different clocks, differing in units and location, could be associated with one musical pattern. The clocks were said to be used as measuring devices to specify the temporal structure

of patterns. The work was supported by their finding that musical sequences which were not rhythmic and which therefore did not induce a clock very strongly were harder to form representations for. They said that better reproductions resulted if the sequence strongly invoked a clock.

Clynes and Walker(1986) also felt that musical concepts interact with psychobiologic clocks. They said musical thought and memory were indications of subjective time rate, and that long term temporal stability of musical performances over time was a demonstration of the long term stability of the clock rate within and between performers. Furthermore, they found that musical performance deviations were quantized. This meant that preferred values of timings existed which corresponded to quantized differences in the tempo of the music. They went on to find evidence for a preferred quantum step in the tempo of musical thought because the percentage of deviation in the timing of different composers performances was similar. It was suggested that a quantized rate in the main clock governing tempo selection could manifest itself in the percentage deviation of timing duration. They also said that the ability to imagine and space out musical tones in time implied that an internal clock existed. A programmed signal was said to show the performer when it was time to do the next thing - like setting an alarm. They wanted to know how the musical concept was converted to the right alarm setting and suggested time form printing by the cns, or subconscious mental agents.

Many authors have looked for correlations between physiological activity and time perception to cite as evidence for internal clocks. Hawkes et al(1962) found productions were correlated to variations in heart and respiration rate, they concluded that as autonomic nervous system activity speeded up, subjective time rate increased so that the objective duration of an interval was overestimated and productions got shorter. Latour(1967) said that if an internal clock existed it would play a role in reflex-like activities. He found evidence for this role in periodicities of the visual threshold, eeg and the reaction time of the eye.

As White(1963) pointed out the conclusion that a psychological unit of duration exists has been reached by philosophers, physiologists and cyberneticists. However, Schiffman et al(1974) insist that the idea of an internal clock has not met wide acceptance.

Ornstein(1969) reflects the views of many when he complains that no biological identification of the clock has been made, and no process to relate its functioning to time experience has been proposed. Furthermore, he states that the 0.7sec indifference interval at which time is accurately judged, which has been suggested as the time base for an internal clock is an experimental artifact caused by central tendency. Curton et al(1974) found that arousal by exercise was negatively correlated with perceived duration, a finding that is the opposite of what internal clock theorists predict. They tried to account for this by saying that the exercise had tired the subjects out, and as they recovered they became more alert and that this is what increased perceived duration. The effects of increased alertness were thus said to be overriding the effects of decreasing arousal.

Although it is tempting to try and relate perceived duration and perceived urgency in terms of arousal and the internal clock theories, there are problems with those theories, some of which have been mentioned above. An additional problem is that the arguments presented by some of the models are somewhat circular and untestable. Povel and Essens(1985) for example claim that musical sequences that are better represented have invoked a clock yet the evidence that a clock has been invoked is the fact that the sequence is better represented. It seems that there are so many unknowns relating to the internal clock, such as its biological basis and which factors may or may not result in arousal of the clock, that it would be very difficult to speculate as to how such a mechanism could influence perceived urgency by altering perceived duration.

6.3.2. Storage Size.

Ornstein(1969) proposed the storage size hypothesis based on his observation that estimates were a positive function of the number of stimuli in the interval and their complexity. He rejected the ideas of theorists who said it was the information registered in consciousness that determined perceived duration. Instead he said that perceived duration was a function of the amount of storage space required by the information stored during the interval, and was dependant not only upon the amount of information but also on how it was stored. He said that if a subject could organise the information into 'chunks' then it required less storage space and perceived duration would be reduced. Ornstein explained his finding that when subjects could respond automatically perceived duration decreased relative to when they could not. He said that in automatic responding less information entered consciousness and thus less was stored. Thus, Ornstein felt

that perceived duration increased with the number and complexity of the stimuli in an interval because these factors increased the amount of storage required by the interval.

Ornstein(1969) attempted to demonstrate that it was storage size and not just increased information input that influenced perceived duration. He altered storage size without altering the input. He said that time order error was due to items dropping from storage which decreased perceived duration. He did two experiments altering the way information was stored after the interval was over so that the only thing that could have an effect was altered storage size. When there were no manipulations during the interval (so input was the same) perceived duration was varied by the way the information was coded. Ornstein also maintained that storage size could account for increases in perceived duration under psychedelic drugs, sensory deprivation results and the 'watched pot phenomenon'. He said that the latter was caused by an increase in vigilance which resulted in increased awareness of input and thus an increase in perceived duration.

Ornstein's claim that more complex stimuli result in increases in perceived duration has been supported by Hogan(1975) and Schiffman et al(1974). Block(1978) however did not find the predicted increase in perceived duration as individual stimulus items got more complex. He said that the concept of complexity was not well enough defined by Ornstein, and also pointed out that he had demonstrated positive time order error which was unaccountable for in terms of the storage size hypothesis.

The assertion that the degree of relatedness between items in storage is important in determining storage size has received more widespread support. Harton(1939) noted that the more organised an experience was, the more perceived duration decreased. He said that on a holiday it seems to be lasting a long time, but when you get back it seems to have been short. He said that this was because on return, the holiday was chunked in memory as a 'vacation', not in detail. Berg(1979) found that a film that was given organising labels was judged as being shorter than an unlabelled one (for intervals longer than 1.6 sec). Mulligan et al(1979) showed that providing a simplifying code for remembering line drawings reduced perceived duration, and Achamanda(1988) found that cognitive efficiency was related to perceived duration. Hawkins et al(1979) failed to support the prediction that related items would result in decreases in perceived duration. They

found a tendency for unrelated tapes to be judged shorter than related ones. They did however find that interesting tapes (which according to Ornstein would have established relationships) reduced perceived duration. Block(1974) supported Ornsteins finding that perceived duration was influenced by the number of events that a subject could remember, as did Poynter(1979) and Buffardi(1971). Buffardi investigated the filled duration illusion and found that the number of intervening elements was the most important factor in increasing the illusion that filled intervals are longer than empty ones.

Ornstein ascribes attention effects to an increase in vigilance resulting in more information reaching storage. Underwood and Swain(1973) find this assumption questionable, and state that the storage size hypothesis is unable to account for the effects of attention on time perception. They varied attention independently of information by varying intensity of noise (high intensity noise increases selectivity of attention) and found that increased attention resulted in increased perceived duration. This contradicted Ornstein's hypothesis because it was attention alone, not increased information, that was having the effect. In fact less information was able to reach storage in the high noise condition and thus Ornstein would have predicted a decrease in perceived duration. Many theories have been developed to account for attention effects in duration judgements. They are discussed below.

If the storage size hypothesis were accepted as an account of perceived duration and if perceived duration effects perceived urgency then there are several aspects of the model that would be relevant to the design of auditory warnings. Warnings vary in the amount and complexity of the information that they contain and may thus require differing amounts of storage size, this could affect perceived duration and thus perceived urgency. Furthermore , the relatedness of warnings may also vary if different sounds vary along similar or the same parameters. However the main problem with the storage size hypothesis is that the concepts of complexity and the determinants of storage size are not well enough defined to enable specific predictions to be made about the effects of different parameter manipulations upon perceived duration.

6.3.3. Attention Theories.

Frankenhauser's(1959) early attention model stated that perceived duration is a function of the amount of attention allocated to the passage of time. He

suggested that a cognitive timer utilises attention to process temporal information, if non temporal load is reduced then more attention is paid to the timer and thus perceived duration increases. This model is supported by findings that increases in task difficulty result in decreases in perceived duration and by findings that empty intervals increase perceived duration relative to filled ones (a finding that refutes the storage size hypothesis). The model also predicts that number of responses and perceived duration should be positively correlated because many responses are thought to be associated with an easy task and thus higher estimates.

Underwood et al(1973) proposed an attentional effort model which said that complex stimuli increased perceived duration because they require more attention. In 1975 Underwood said that increased selectivity of attention to the passage of time could also increase perceived duration, (as in the watched pot scenario). Thus the attentional model predicts that perceived duration will increase if a task is dull and has low attentional demand because then the subject can attend to the passage of time, or if the task has a high attentional demand. This attentional effort model was supported by Curton et al(1974) and Thomas et al(1978). Considering the finding of Martin(1972) that more processing was demanded of retrieval than encoding, Underwood(1975) assumed that more attention would be required of the former and predicted that retrieval would be subjectively longer than encoding. He found that retrieval intervals were judged longer than encoding ones, and that this difference increased as the meaningfulness of the material decreased. Underwood interpreted these results as evidence that retrieval requires more attention and more attention results in increased perceived duration. Fraisse(1984) states however that subjects may have been given too long for retrieval and so they were paying more attention to the passage of time and that is what increased perceived duration.

The attention accounts of time perception again suffer from lack of precise definition since it is hard to say how complex a stimulus must be in order require additional attention and thus to increase perceived duration or to know when the passage of time is being attended to. These issues are critical to auditory warnings research. If perceived duration and thus to a certain extent perceived urgency depended in part on attention then these terms must be very clearly defined so that the effects of the operators attention to the task and to the warning could be assessed.

6.3.4. Information Processing Theories.

Michon(1965) and Vroon(1970) both distinguished between presented and processed information, stating that perceived duration was related to the amount of information actively coped with. They found that perceived duration was an inverse function of the amount of information that was processed during an interval, so that perceived duration decreased as the information processing in the interval increased. Their findings were supported by Hicks et al(1974) who also found that perceived duration decreased when information processing was involved, they said that this was because processing load led subjects to neglect temporal cues.

In 1976 Hicks et al reviewed the literature and found that in the prospective paradigm perceived duration increased with the number and complexity of stimuli, when no processing was required. When processing was required perceived duration decreased with the number and complexity of stimuli. Hicks et al(1977) conducted prospective experiments where the processing demands of the tasks were systematically varied. They found that perceived duration decreased linearly as processing time increased, so supporting their hypothesis that time perception requires processing capacity. The authors extended the ideas of Frankenhauser(1959) to propose a time base responsible for subjective temporal units. They said that more presented information added to the counter thus increasing perceived duration while more information processing prevents the storage of events in the counter and thus decreases perceived duration.

To summarise, the work of Hicks et al(1976) on prospective judgement implied that when time itself was attended to perceived duration increased, and that when the interval was filled with other tasks perceived duration decreased. These findings have implications for the performance of specific tasks in high workload environments. In retrospective judgement they said that perceived duration increased with the number and complexity of events stored about the interval. In this paradigm no subjective temporal units were made and thus judgement was based on what was remembered about the interval. Their work was supported by Underwood et al(1973) and by McClain(1983).

Thomas and Brown(1974) proposed a reversible encoding model. They said that stimulus input was encoded as a vector with an encoded and a decoded duration.

They said that a 'filled' interval caused the interval to be encoded in chunks that were decoded serially. Perceived duration was said to be the sum of the decodings. They accounted for the filled interval effect by saying that the length of the decoding was a function of the length of the encoded chunks.

Thomas and Weaver(1975) extended these ideas, when they proposed that a direct relationship existed between the time spent processing the non temporal aspects of stimuli and perceived duration. They said that duration was analysed by a timer and an information processor. The output of the timer was determined by stimulus duration and the output of the information processor was related to encodings of non temporal information and encodings of the time spent processing non temporal information. It was said that attention was shared between the timer and the information processor , the processor with most attention had the most influence on judgement. The authors said that in prospective paradigms, if the subject knew that the interval would be empty, only the timer would be used; similarly in a retrospective paradigm using filled intervals duration judgements would be based only on the output of the information processor. More usually perceived duration would be the weighted average of the timer output and the encoding of time spent processing information in the interval. According to their model, when the information content of the stimulus is large, the timer gets less attention so its output is smaller and/or more variable. Perceived duration then depends more on the output of the information processor, and increases with increased information processing in the interval. Brown's(1985) finding that perceived duration decreased in prospective judgements as nontemporal task demands increased supported these assertions. Zakay et al(1983) suggested an elaboration to include importance weightings for the information from the two processors, with the weightings influenced by paradigm, attention and duration.

Fraisse complained that it was hard to define the quantity of information to be processed in models like Thomas's. His own work on time perception claimed that the conditions that affect it do so by affecting attention and adaptation level. The first idea predicts a contrast effect when unexpected stimuli are introduced, this was found by Mo(1971). In 1975 Fraisse proposed his own model which said that perceived duration was based on the number of perceived changes during the interval. He said that attention increased the number of perceived changes and thus made the interval seem longer as in the 'watched pot' scenario, whereas motivation made subjects absorbed in the task so that it took on unity of significance and fewer changes were perceived so the perceived duration decreased. In short, the

more unified that tasks were, the shorter their perceived duration, whereas more divided tasks were said to increase perceived duration. His observation that perceived duration increased under the influence of hashish and during dreams was cited as evidence for his theory, he said that in both conditions the succession of un-unified images resulted in more perceived changes and thus increased perceived duration. Fraisse(1984) claimed that the easier the information processing during an interval, the more perceived duration increased because the subject was more attentive to the duration, and could thus perceive more changes. Matsuda's(1965) finding that the more subdivisions there were in an interval the longer it was judged, supported Fraisses' ideas.

Block et al(1978) also felt that perceived duration was mediated by the remembered amount of contextual change during an interval. They said that perceived duration would increase with more complex stimuli and with more stimulus events, so that perceived duration was a positive function of processing load. In 1985 Block explained that contextual changes could come from the environment in terms of, for example, task demands; or from the organism in terms of for example mnemonic activity. They said that changes were monitored by an internal cognitive device that output a complexity index based on changes per unit time. The contextual change hypotheses were supported by Poynter(1983) who found that segmented word lists produced increased perceived duration relative to unsegmented ones.

In 1977 Bobko et al found that the main effects of stimulus complexity were not statistically significant, although they were more pronounced at shorter intervals. The authors felt that duration judgements could be determined by complexity only at briefer intervals, but conceded that methodological problems may have meant that they did not find an effect.

The information approach to time perception appears to offer the least contentious means by which the changes known to affect perceived urgency might effect perceived duration. These accounts focus on the information that is presented or processed in the stimulus interval, or in the warning.

6.3.5. Comparison Of Models.

In Table 6.1 the focus and predictions of the different models can be examined, as well as areas of compatibility.

It is clear that many of the factors which increase perceived duration according to the information processing and attention theories could increase storage size and perhaps arousal and so increase perceived duration. This would mean that these factors are compatible with the predictions of the storage size and internal clock models. The contextual change hypothesis is particularly compatible with the storage size hypothesis, for it can be suggested that the more change that is perceived, the more storage space is required. Those models that predict increases in perceived duration with increases in presented information, or with increased attention to complex stimuli are also compatible here if it is assumed that these factors result in more perceived changes. All of the models would predict increased perceived duration in these conditions. The prediction of the attention model that more attention to an empty interval results in increased perceived duration (as a result of accumulating more temporal cues) is compatible with the information processing prediction that more information processing decreases perceived duration (due to neglect of temporal cues). However if more temporal cues result in more perceived change then this model is compatible with the contextual change and storage size hypotheses. To summarise, the theories assume that more 'cognitive activity' results in increased perceived duration. Some say that this activity is due to the accumulation of temporal cues, others that it is due to accumulation or processing of nontemporal information. Both forms of activity could result in more perceived change and the requirement of more storage space. In order to include the internal clock hypothesis it would have to be assumed that the cognitive activity is arousing.

There have been a few explicit comparisons of time perception models. Block et al(1980) studied how different models accounted for the 'watched pot' phenomena. They found a task by paradigm interaction. In prospective designs, perceived duration increased when the pot never boiled and an interruption decreased perceived duration if the liquid did not boil but had no effect if it did. In retrospective designs perceived duration increased if the pot boiled or there was an interruption or both. These results were supported by Hicks et al(1976) and Miller et al(1978) both of whom also found that the variable produced opposite

INTERNAL CLOCK Focus: Determinants of arousal,biological base Predict : More arousal = inc. perceived duration Less arousal = dec. perceived duration
STORAGE SIZE Focus : How information stored Predict : More storage space required = inc. perceived duration Less storage space required = dec. perceived duration
ATTENTION Focus : Where attention directed Predict : Attention to empty interval = inc.perceived duration Attention away from interval = dec.perceived duration Attention to complex stimuli (effort) = inc.perceived duration
INFORMATION PROCESSING Focus : Amount of processing Predict : More info. presented = inc. perceived duration (prospective) More processing = dec.perceived duration (prospective) More stimuli = inc. perceived duration (retrospective) Focus : Contextual change Predict : More change = inc. perceived duration Less change = dec. perceived duration Focus : Encoding Predict : More processing = less reliable/dec. perceived duration Large stimuli = more processing thus inc. perceived duration

TABLE 6.1: SUMMARY OF THE TIME PERCEPTION MODELS.

effects in prospective and retrospective paradigms. They concluded that different theories were needed to explain experienced and remembered duration. In prospective designs where an interruption had not affected perceived duration when there were changes in the task related content (boiling), it was claimed that people based such judgements on task related content. The authors said that if there were no changes in task related content then an interruption could shift attention from the time to the interruption thus reducing perceived duration. In retrospective designs changes in task related or unrelated (interruption) content both increased perceived duration. This was as explained in terms of the contextual change hypothesis, with more change resulting in increased perceived duration. The authors claimed that their findings supported theories that considered the role of attention. They also supported authors such as Miller et al(1978) who said that prospective judgements were based on the number of subjective temporal units that were created and stored while retrospective judgements were based on the amount of content of an interval that was remembered. Vigilance and selectivity of attention explanations for the 'watched pot' phenomena were not supported.

Gomez and Robertson(1979) tested the assumption of Ornstein and of Thomas, that the filled interval effect was a function of the nominal properties of each stimulus event. Ornstein said that it was a result of the load placed by non temporal information on memory, whilst Thomas and Weaver said that less attention was allocated to the timer when there was more non temporal information. Their model predicts that if the non temporal information is encoded before the interval ends then attention should be allocated back to the timer, thus the illusion should be eliminated for longer intervals. The authors looked at the influence of processing strategy, and found that variations in the pattern size of stimuli presented in the interval only affected the illusion if varied as a within subject variable, there was thus nothing inherent in the pattern size alone that created the illusion. They also found that the range of durations in the stimulus set influenced the perceived duration, whereas the time available for processing nontemporal information did not. Neither model predicted that the environment in which judgements were made would influence the illusion, and neither predicted that range would have an effect. Ornstein's model also had problems with the finding that the illusion increased as temporal discriminability got harder. Thomas and Weaver's prediction that the effect would be eliminated at longer intervals was not demonstrated. The assumption of both models, that the illusion was a function of the absolute parameters of the stimulus, was not supported by Gomez and

Robertson. They said that the nontemporal properties of the stimulus alone did not produce the illusion.

Poynter and Homer(1983) compared the storage size, contextual change and processing effort (eg Thomas and Weaver) explanations of time perception. They found that increased memorisation (which required more processing effort and more storage space), increased pattern uncertainty (which required more processing effort) and number of events (which should cause more contextual change) all affected estimates. They found that for short intervals more stimulus changes increased perceived duration whereas for longer intervals less stimulus changes increased perceived duration. The idea that filled time increases perceived duration thus seemed to apply only for short intervals. This was supported by the finding that memorisation increased perceived duration for short intervals and decreased it for long ones. The authors work helps to reconcile previous experiments that have reported both positive (eg Thomas and Weaver) and negative (eg Hicks et al 1976) relationships between processing effort and perceived duration. It suggests that the clock length of the interval may determine whether processing load increases or decreases perceived duration. The authors conclude that change is the most consistent factor that affects time judgements, and state that if change is the unit on which time perception is based then organistic changes account for the perception of long empty intervals.

Zakay et al(1983) compared the storage size model with attention models that postulate a cognitive timer. Using a prospective design they supported timer models by finding that perceived duration was a negative function of task difficulty and that empty intervals were judged longer than filled ones. They concluded that in their particular experimental setting the timer models had been supported over the storage size hypothesis.

It is apparent that tests of the models have been unable to recommend conclusively one over another. The emphasis of the research does appear to have been focussing more recently upon change during a temporal interval, eg Poynter et al(1983). For this reason attention will turn to the temporal patterning approach to time perception to see if it can offer a unified perspective.

6.4. Temporal Patterning.

More recent theories of time perception have looked at it in terms of temporal patterning. Since the basis of these ideas lie in the field of music and rhythm perception, it is necessary to look at that area before the theories are discussed in detail. We might expect rhythm and time perception to be linked for as Fraisse(1978) pointed out, rhythm is an ordering of temporal succession, a patterning of time in time.

6.4.1. Rhythm and Music Perception.

One aspect of rhythm perception is the 'grouping' of elements. This will be reviewed in detail because it is an important aspect of research into rhythmic behaviour and because this grouping may have important implications for auditory warnings design.

Fraisse(1978) said that rhythm was organised by pause (between elements and groups of elements), accent (an element standing out relative to other ones) and run (the grouping of identical elements). When his subjects had to tap their own rhythms 92% of their intervals were less than 1 sec, and only 2% were more than 1.8 sec. He said that this was the limit above which there was no longer wholistic perception of two consecutive elements in a pattern. Fraisse also claimed that his subjects produced intervals in the ratio 2:1 (he said that this represented a preferred tempo that could be related to an internal clock), and that relative mean produced durations increased with increases in the length and complexity of the pattern. When subjects had to reproduce intervals the author found that high ratios between intervals were overestimated and low ones were underestimated. The pauses produced between patterns were found to be at least as long as the longest interval in the group, this was thought to be essential for unambiguous pattern perception. The author concluded that rhythm perception involved the wholistic grouping of a pattern as well as linking it to what follows.

According to Fraisse(1978) Wundt's 'subjective rhythmization' (identical sounds separated by equal intervals are spontaneously heard in groups) occurs when stimuli follow each other so that they are distinct but not independent. He said that the maximum interval for this to occur was 1.5 - 2 sec., and 400ms was the best interval. Fraisse required subjects to make subjective groupings and found

that intervals between two successive groups seemed longer than intervals between elements, and that groups were perceived as ending with a longer stimulus or interval. This work has important applied implications for auditory warnings design. Manipulation of the intervals between warnings could encourage subjective rhythmization so that successive warnings were heard as part of a group when they signalled the same event. Similarly the interval could be manipulated to ensure that different warnings remained distinct from each other. Further research could be conducted to confirm the optimal intervals for the above functions for different warning sounds.

Other explanations for rhythmic behaviour also looked at grouping. In 1982 Martin said that rhythms were hierarchical and that their perception involved complex laws of subjective grouping. This idea was continued by Longuet Higgins et al(1982) said that listeners inferred rhythm by comparing note lengths and constructing metrical hypotheses on the basis of what they heard. The most important assumption of their model is that at any time after hearing the beginning of a sequence the listener has in mind a hypothetical grouping that is accepted or rejected. As expectations are confirmed the listener tries to move up the hierarchy by combining the confirmed units into a larger one. Deutsch(1980) and Butler(1979) also thought that grouping is important in rhythm perception. The former claimed that temporal relationships between tones were important in determining grouping when they were from different spatial locations, whereas the latter said that tones could either be grouped in terms of frequency range or spatial location.

Povel(1979) proposed a model of rhythm perception that did not require grouping. He said that instead the sequence was segmented into beats and coded onto a temporal grid which fixated on part of the sequence creating a framework that allowed the specification of the remaining elements. Intervals shorter than the beats were coded as subdivisions of the beat. The grid allowing the most economical description of the sequence was then selected. The model thus saw rhythm perception as determined by the internal structure on to which listeners tried to map the presented sequence, he said that distortions occurred unless the sequence exactly matched the mental structure; it was also to a certain extent hierarchical, with beats at higher levels and subdivisions of them at lower levels.

In 1968 Garner suggested that rhythm perception was hierarchical. He said that at fast rates of presentation patterns were perceived in an integrated manner,

whilst at slow rates of presentation the unintegrated elements had to be organised by an active observer. Garner also introduced the finding that listeners were left feeling incomplete if patterns did not end at their natural ending. Other studies on rhythm perception such as Deutsch(1980) found that subjects distorted temporal patterns so that they were closer to simple metrical descriptions. She said that such distortions occurred because earlier durations in a pattern were updated in accordance with later ones, and concluded that patterns were characterised in terms of metric hierarchies that consisted of successive divisions of time spans into units of equal length. This was supported by the work of Gabrielsson(1974) who had observed systematic deviations from the norm in recordings of musical performances. The idea that rhythm is hierarchical was also expressed by Todd(1985) and by Shaffer(1984). Shaffer said that the timing of musical performances were organised at three levels, global, intermediate, and local, and that rhythm was accentuated by slowing at structural endings where the degree of slowing reflected the hierarchical structure. He felt that the slowing points were like parsing devices allowing the listener to perceive the hierarchical structure.

In 1985 Povel and Essens suggested that a hierarchical internal clock might have a role in rhythm perception. They extended the previous grid idea to say that listeners tried to generate an internal clock and use it to specify the temporal structure of a pattern. It was said that the ticks of different clocks were matched against the pattern elements and the clock that best matched the pattern was induced. They claimed that patterns with more strongly induced internal clocks were better reproduced. The notion of an internal clock controlling rhythmic behaviour also receives support from the animal world where rhythmic behaviour such as flying is controlled by the central nervous system. Clynes et al(1983) see human rhythm perception as evidence that there is a stable psychobiologic clock.

It is apparent that rhythm, and thus music perception, are closely linked to our ability to perceive time. We have seen that there are invariants in musical perception, preferred ratios for reproduction, and we have also seen that there is evidence to suggest that rhythm perception is hierarchical. Furthermore it appears that music generates expectancies in the listener because apparently we are left feeling incomplete if patterns end before when is natural. These concepts are incorporated into one of the most recent time perception models, by Reiss Jones and Boltz. This may have particular relevance to auditory warnings research because many new auditory warnings are musical in nature.

6.4.2. Background to the Model.

Reiss Jones and Boltz(1989) said that time perception theories such as those discussed in preceding sections, were inadequate because they concentrated too much on the processing of nontemporal information. They felt that it was not the processing but the temporal patterning of information that was important in time perception. They attempted to create a wide theoretical framework for time perception by proposing an alternative hypothesis based on the idea that events are temporal and that the structure of events in time is crucial in determining perceived duration. It was felt that time could not be evaluated independently of the events that were used to signal duration. These ideas were supported by environmental observation and by work in auditory pattern recognition.

Evidence that events might be temporal in nature was provided by Reiss Jones(1976). She noted that when understanding, for example, speech, people have to retain the temporal order of the sounds. She said that subjects could recognise the order of vowels only 30ms long in natural sound whereas in repeating sequences of synthetic sound vowels had to exceed 168ms in length for order to be recognised. This was cited as evidence that the detection of temporal order involves the sequence in which the sound is embedded. She tried to see what it was about the sequence that facilitated the detection of order, and concluded that the sequence was not just represented by changes in auditory dimensions such as pitch but also by changes in the temporal dimension. It was said that context facilitated order retention both in speech and non speech. She also asserted that organisms were rhythmical, for example animals produce music via co-ordinated body gestures; that our representation of the auditory environment was hierarchical; and that the structure of events led to temporal predictibility .

In the same year Reiss Jones said that listeners generate expectancies along simplified schemas which activate graded rhythms appropriate to the expected time periods. As the real world sound begins the activated rhythm is thought to lock on to it (entrainment). She said that when a pattern is perceived as being easy, that is because it verifies expectancies. When expectancies were violated she said that patterns are perceived as being hard. In 1978 she said that expectancies were important in the perception of rhythm. Schmuckler(1989) also supports the idea that music generated expectancies.

In Reiss Jones et al(1981) the idea was introduced that attending might be a rhythmic, time -dependant process based on nested subjective rhythms. She said that attention was dynamic so that small attentional rhythms helped the selective pickup of serial relations between adjacent events and larger temporal rhythms locked onto higher order relations that held between non adjacent events. She conducted several experiments in which rhythmic context was varied to see if attention was dependent on temporal and spatial structure. Evidence to suggest that attention is rhythmic was provided by Reiss Jones et al(1982) when they found that pattern regularities encouraged the rhythmic attender to focus attention within an unfolding sequence over anticipated temporal interval. They varied rhythmic and melodic context within a pattern recognition task to see if rhythmic context could direct attention to or from tones that instigated higher order melodic rules. Subjects were better at detecting violations in higher order melodic rules when the rhythmic context induced tones that instigated these rules. Thus rhythmic context was guiding attention and temporal predictibility was enhancing the detection of higher order melodic structure. The authors concluded that attending was rhythmic because it was guided by temporal patterning.

The findings that auditory events are temporal, hierarchical and able to induce expectancies, and that organisms are rhythmical, formed the basis of the Reiss Jones and Boltz explanation of time perception.

6.4.3. The Model.

Reiss Jones and Boltz employed the relativistic approach to time perception which saw absolute time intervals and points as less important than time periods relative to other ones (rhythmic structure) and time periods relative to spatial extents (velocity structure). It was felt that time was only one dimension of serial pattern structure. Pitch and loudness were said to be the other dimensions. All three dimensions were thought to be interdependent. Because time relations were inseparable from the event they felt that subjects may be unreliable when judging intervals in isolation.

The Reiss Jones and Boltz(1989) model assumes that the environment is filled with temporal events on a continuum of temporal coherence. Coherent events were said to have hierarchical time structures (consistent time transformations), objective accent regularities such as the cats' locomotion sequence, distinctive non temporal markings and lawful temporal nestings. The

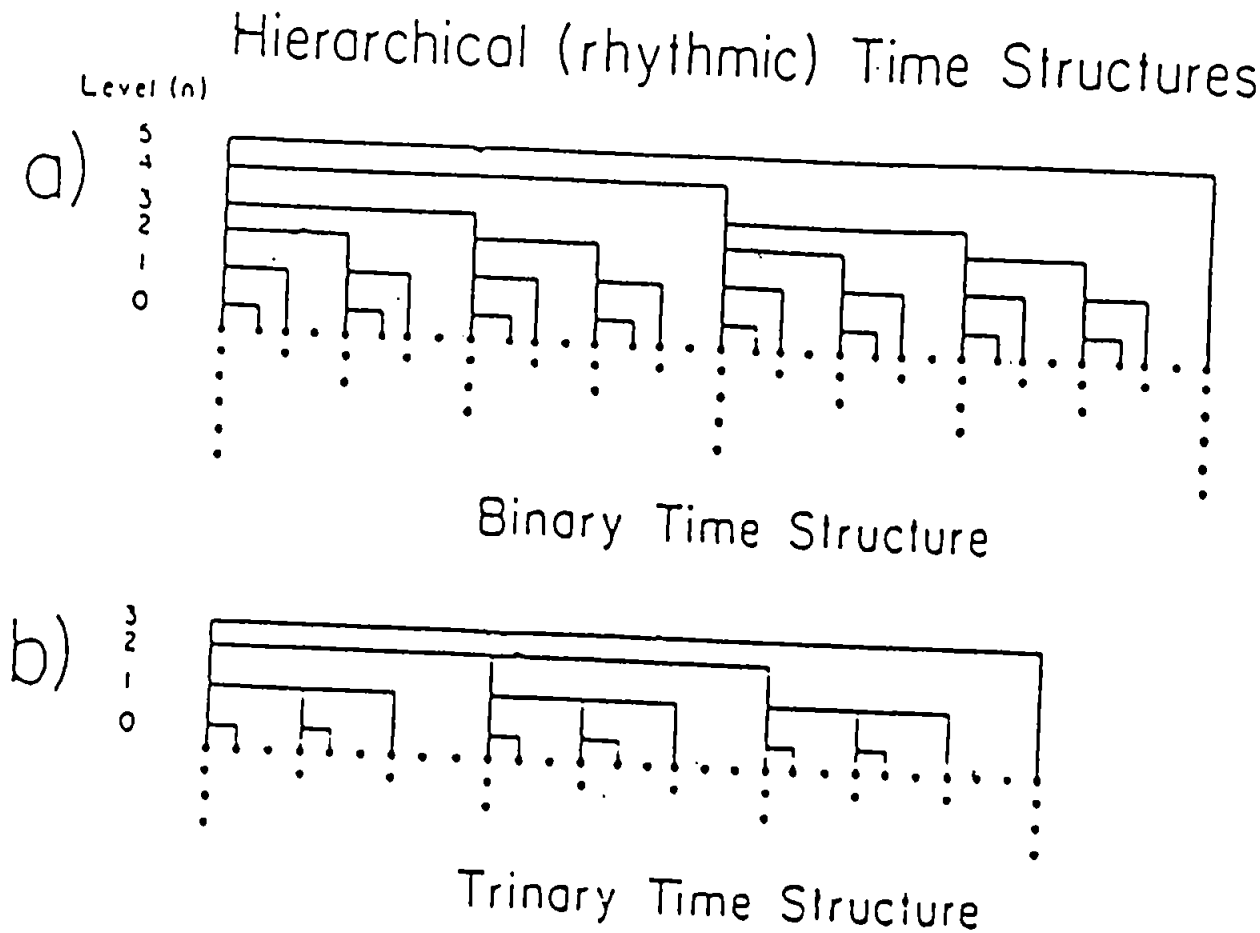
beginnings and endings of these nested time periods are marked by structural change. Non coherent events were said to be non hierarchical (obscure connections between internal structure and time span).

The authors said that in hierarchical events each nested level was associated with a recurrent time period. If the structure was binary then the nested levels were related by a ratio time transformation of two, trinary structure is less coherent because the time ratio involves a larger integer. These relationships are shown in Fig.6.1. Non hierarchical structures do not have simple temporal recursivity, and the transformation rule is inconsistently applied with the result that relations between embedded levels and total duration are obscured.

The model suggested that attenders interact with the environment (attunement) and that attending is sensitive to environmental rhythmicities. Attunement was said to be most likely to occur to temporally coherent events and to result in entrainment (the locking of a biological rhythm onto a regular well marked time period in an event which then functions as an anchor/referent time level). Attunement thus creates an event-determined time scale that calibrates time spans. The referent is influenced by age and biology. It was said that attending was dynamic and that it could be future orientated (global focal attending over time periods higher than the referent) or analytical (attending to periods lower than the referent in the temporal hierarchy). The attender could synchronise with nested time levels in different ways to achieve different goals. It is possible that an internal clock could facilitate these processes.

According to the model, time perception is biased by the style of attunement failure, which is caused by asynchrony between the attending rhythm and the referent/focal level of the event. In coherent events attunement failure results in temporal contrast which biases time judgement. In non coherent events time perception is biased by subjects' structuring strategies in analytic attending. Thus duration estimates of the same event can vary with attending mode,

Fig. 6.1 : The Hierarchical Nature of Time Structure.



From Reiss Jones and Boltz(1989)

prospective judgements are more likely to induce future orientated attending, whilst retrospective judgements favour analytic attending. The model predicts that judgement variability would decrease with increases in the predominance of one attending mode.

To summarise, Reiss Jones and Boltz(1989) said that the temporal patterning of non -temporal information within an event determines how it is attended to. Events with high and low structural coherence are attended to differently - in the former eg speech and tonal music, attending is said to be future orientated and to exploit global time structure (here estimations are longer if one of two sounds violates its expected ending by ending later, and are shorter if it appears to end too soon). Thus in coherent events, estimates depended on temporal contrast - the disparity between an events' actual and expected ending. Events with low structural coherence are said to promote analytical attending. In this instance time judgement is said to be biased by attention to local detail(in an attempt to organise information), thus events with more items are judged longer. The authors thus suggested that the interpretations of responses to time should consider what the time means.

6.4.4. Support for The Model.

Boltz(1989) studied the relationship between musical resolution and time perception. She conducted an investigation to see whether violations of musical resolution affected duration judgements. The model predicted that unresolved melodies would appear shorter because they would end before their expected ending. It was found that perceived duration was affected by expectancies about musical resolution - when the expected final tonic was missing the melody appeared to end too early and was underestimated. The study went on to investigate the cues that subjects use to derive musical completion. It was revealed that violating rhythmic structure enhanced the time distortions achieved by manipulating tonal endings. Again the model was supported because time perception was shown to be also affected by the rhythmic structure of the event, which Jones(1976) had said could guide attending.

In the same year Boltz looked further into how rhythmic structure generates expectancies about the future of an event. It was noted that the tone-tonic progression and resolution marked the end of much Western music. Preliminary

work indicated that subjects used structural markers and contextual factors to guess endings. The author felt that internal representations of tonality included temporal order relationships which provided information on completion and musical embelishment. It was also suggested that rhythm creates a predictable pattern of temporal accents that highlight nested levels of temporal organisation within a melody. An experiment was conducted to see if variation in the timing of rhythmic accentuation could affect perceived completion. Melodies ending on a leading tone-tonic were judged complete, whereas the least complete melodies left the listener hanging by ending on a tone. Ratings were also influenced by accent structure - accent structures leading to endings which occurred earlier or later than expected had lower resolution ratings, melodies that ended 'on time' were judged most resolved. As predicted by the model, markers of musical completion include the temporal ordering of tonal relationships and preceding rhythmic context.

The Reiss Jones model claims that time perception is affected by factors such as musicality, expectancy and grouping of the information within a stimulus interval. This is relevant to warnings design because the variations of the information within the stimulus interval that affect perceived urgency may also convey changes in time perception through such factors.

6.5. Conclusions.

Having reviewed the literature, it is apparent that the information processing and Reiss Jones models of perceived duration are the most relevant in relation to perceived urgency. If perceived urgency and perceived duration are related then it is those models that may explain the effect because they see variations in time perception as resulting from factors that are known to cause variations in perceived urgency. The information processing models concentrate on the content of the interval which is the variable in auditory warnings design. The Reiss Jones theory will be studied in more detail not only because it is the most global time perception theory but also because it employs the relative approach to time perception. It is felt that this approach is particularly relevant to auditory warnings work for many auditory warnings are 'musical' in nature. The perception of these warnings may thus be influenced by relative acoustic and temporal values as is music perception. Furthermore auditory warnings are not heard in isolation like the stimuli in 'absolute' time perception experiments, it is

often the perception of them relative to each other and in context that is of interest and important.

Thus both types of model take into account many factors that could affect perceived urgency, such as information content, attention, rhythm perception, pattern perception, expectancy and event structure. By manipulating these factors in specific ways it may be possible to see how they affect the perceived duration, and so, possibly, the perceived urgency. Such investigations will directly test the assumptions of both models, will provide further practical evidence on perceived urgency and may relate the concepts of urgency and duration theoretically.

CHAPTER SEVEN

INVESTIGATIONS INTO THE RELATIONSHIP BETWEEN PERCEIVED DURATION AND PERCEIVED URGENCY.

In this chapter the previous findings on perceived urgency are evaluated in the light of the information processing and Reiss Jones accounts of perceived duration. In the first experiment the information processing accounts of perceived duration are evaluated, in the second the Reiss Jones account is evaluated. The experiments described investigate the relationship between perceived urgency and perceived duration to see whether changes in perceived duration contribute to the effect of perceived urgency and whether perceived duration might be a non acoustic determinant of urgency. In the first experiment parameter changes known to cause increases in perceived urgency are manipulated and the effects of these changes upon perceived duration is noted. The relationship between perceived duration and perceived urgency, and the information processing theories of time perception are evaluated in the light of the findings. In the second experiment, the Reiss Jones approach to time perception is used as a framework to investigate the idea that there are two different types of parameter that affect perceived urgency - one that is culturally determined and one that is more fundamental or innate.

7.1. Experiment Ten : A Study of the Effects of Changes in Speed, Pitch, Units of Repetition and Inharmonicity Upon Perceived Duration.

7.1.1. Introduction.

In this experiment acoustic parameters were varied in ways known to affect perceived urgency. The effect of these variations upon perceived duration was noted.

Having reviewed the literature on perceived duration it is apparent that the 'information processing' view of perceived duration, which concentrates upon information processing, memory and attention, is an appropriate area on which to concentrate. In the previous experimental settings (Experiments 4-9) the only manipulated variable has been the content of the stimuli. If perceived duration

has an effect upon urgency judgements to those stimuli then the content of the stimuli must have altered perceived duration. An explanation for the relationship between perceived urgency and perceived duration might thus be found in the information processing accounts of perceived duration for they concentrate on the content of the stimuli. It is important to know how the stimuli used in previous experiments might have affected perceived duration.

The information processing theories of perceived duration which focus on the amount of information that is presented or processed during the stimulus interval are particularly relevant to previous studies in which speed or repetition units were used to vary urgency. When speed or repetition units increase so does the information that is presented and/or processed in the stimulus. If perceived duration did influence urgency judgements, it may have been as a result of that variation.

In retrospective time perception research paradigms, increases in information presentation (Vroon 1970, Ornstein 1969) and information processing (Hicks et al 1974, Ornstein 1969) are said to increase perceived duration. More specifically, Underwood et al (1973) stated that increases in the selectivity of attention required by an information processing task increased perceived duration and Block et al (1980) said that more changes in cognitive context increased perceived duration. These changes are thought to occur because retrospective judgements of duration were based on the memory of the amount of 'events' in the interval, with more memory for theses events increasing perceived duration. (Ornstein is included in the information presentation and processing categories since it is not clear whether he felt that stimuli had to be processed, or merely presented, to reach storage).

Vroon(1970) said that increases in information processing decrease perceived duration. This claim is usually made for prospective paradigms (Michon 1967, Thomas et al 1978, Hicks 1974 and Block et al 1980). The effect is said to occur because subjects are prevented from accumulating temporal cues. When information is just presented, rather than processed, in the prospective paradigm it is thought to be added into the temporal counter (Hicks et al 1974), added into the nontemporal information processor (Thomas et al 1978) or to increase mental content (Frankenhauser 1959). All of these changes are said to increase perceived duration.

Apparently there are competing predictions about the way in which the speed and repetition stimuli used in previous experiments might have influenced perceived duration. These predictions depend upon whether the information in the stimuli was processed by, or merely presented to, the subjects, and possibly on whether subjects were making prospective or retrospective style judgements.

It is hard to speculate as to whether the information in speed and repetition stimuli is processed or presented, and the problem is made worse by the poor definitions of the terms in the literature. Subjects were requested to judge stimulus urgency and so must have been attending to the stimulus content in order to derive urgency cues. The task of making urgency judgements does not however meet Brown's(1985) criteria for ensuring processing. Subjects did not have to respond explicitly to different aspects of the stimuli, nor did they have to perform a demanding task. It is felt therefore that the stimulus content in previous experiments was probably presented rather than processed.

Before any prediction could be made as to the contribution of perceived duration to urgency judgements, authors, such as Hicks, would need to know whether subjects' judgement style was prospective or retrospective . In order to answer this question we turn to Block(1990). He stated that prospective judgements represented the experience of time in passing, whereas retrospective judgements represented remembered duration. In previous experiments, urgency judgements were prospective. It is hard to believe that when subjects judged urgency prospectively they waited until the end of the stimulus to consider its length, and then added that factor to their judgement. If length or perceived length is part of what makes a stimulus urgent it must be judged with the urgency, not as a separate consideration afterwards. It is therefore felt that time would have been experienced 'in passing' in the previous studies, and thus that the judgement style was prospective.

It should be noted that some authors, such as Brown et al 1988, do not consider the issue of prospective versus retrospective judgement important. They feel that the same fundamental judgement process underlies both paradigms, and recommend prospective designs because they result in more accurate time judgements and are methodologically superior for they allow subjects to make more than one judgement each. Brown (1985) said it is attention to temporal cues such as contextual change that determines duration judgements in both paradigms. Similarly, Fraisse (1981) said that perceived duration depends partly on

remembered (retrospective) or experienced (prospective) changes, and partly on temporal cues. He said that in prospective judgements subjects purposefully memorise change, whereas in retrospective judgements they rely on incidental memory. Thus the same fundamental procedure underlies both paradigms. Both Brown and Fraisse predict that increases in presented information results in increased perceived duration.

Given the aforementioned assumptions of presented information and prospective time judgement, authors such as Hicks, Thomas and Frankenhauser would predict that increases in speed and repetition would result in increases in perceived duration, and that changes in pitch and inharmonicity would have no effect upon perceived duration. Fraisse and Brown would make the same prediction regardless of whether or not the assumptions that the task involves prospective judgement and presented information hold true. Experiment Ten was designed to test these predictions. By doing so it will be possible to evaluate the information processing theories of time perception, to investigate the relationship between perceived urgency and perceived duration and so to examine the concept of urgency.

All of the parameters that have been shown to increase perceived urgency were included in the study since it is important to know whether a finding relating one of them to perceived duration is applicable to them all. Subjects were required to estimate the duration of stimuli that varied in these parameters and also in their actual duration.

Stimuli that varied in resolution were also included in this study because Edworthy et al (1991) showed that resolution was another parameter that could be used to communicate perceived urgency - unresolved stimuli were perceived as being more urgent than resoved ones. The stimuli varied from resolved (whereby the stimulus ends sounding complete), to unresolved (whereby the stimulus appears to end too soon, it violates its expected ending), to atonal (whereby no expectancy is generated at all). Boltz(1989) found that more resolved stimuli were judged longer, in accordance with the predictions of the Reiss Jones contrast model of perceived duration. The contrast model states that resolved stimuli are coherent and thus encourage future orientated attending. This means that expectancies are generated as to the expected ending of the stimulus, these expectancies are violated by unresolved stimuli which appear to end too soon. Unresolved stimuli are thus underestimated relative to resolved ones . The information processing theories however predict no effect of resolution upon

perceived duration because altering resolution does not affect the information content of the stimulus. The inclusion of the resolution stimuli in this experiment enables the competing predictions regarding the effect of resolution upon perceived duration of the information processing and Reiss Jones accounts of perceived duration to be tested.

Thus five parameters that have been shown to affect perceived urgency were varied and the effect of changes in them upon perceived duration was noted. The information processing accounts of time perception predict that only changes in speed and repetition should affect perceived duration.

7.1.2. Method.

7.1.2.1. Subjects.

Four male and eleven female subjects volunteered to participate in this study in partial fulfilment of their coursework requirements. They were first year undergraduate Psychology students from Polytechnic South West, aged 18-29 years. None of the subjects had previously participated in similar psychophysical studies, and all reported having normal hearing.

7.1.2.2. Materials.

The laboratory and hardware arrangements were as reported in Experiment One.

The stimuli employed in this experiment are described in Appendix 7A. There were 39 stimuli in total, 9 varied in pitch, 9 in inharmonicity, 9 in speed, 9 in resolution and 3 in repetition. Three repetition stimuli were constructed at a low, medium and high level of urgency, they were not presented at different durations because it is impossible to co-vary repetition and duration without altering the speed of the stimulus. For each of the parameters pitch, inharmonicity and speed there were three levels of urgency (low, medium and high), presented at three different lengths (2400, 1600 and 800 ms). Three different actual durations were presented so that it would be possible to see whether any effects of parameter level applied across several durations or not. The resolution stimuli varied from resolved to unresolved to atonal, each was presented at each of the three durations . Where possible the parameter levels were within the range of those levels employed in previous experiments.

Inharmonicity was defined as the number of inharmonic components in the pulse, as in Experiment Seven.

For the speed stimuli, the actual lengths were as close as possible to, but not precisely the 2400,1600 and 800 ms reported for the other parameters. The mean lengths of the speed stimuli were 2399, 1560, and 874 ms. Speed was defined as the pulse to pulse time of the stimulus, as in Experiment 9.

7.1.2.3. Procedure.

Subjects were run one at a time while seated at a desk in laboratory Two, approximately 1 metre from the speaker. They were told the broad nature of the study, their wrist watches were removed and they were asked to read the following instructions (adapted from Bobko et al 1977);

"This is an experiment on time perception. Different sounds will be presented for different periods of time. I want you to estimate how long you think the sound is on for during each period of time, to the nearest fraction of a second. Try to estimate the first sound as accurately as you can in seconds and fractions of a second. Thereafter try to keep your judgements proportional. For example, if you think the first sound is 1.25 seconds long, any sound that you think is twice as long should be judged 2.5 seconds, and any sound that you think is half as long should be judged 0.625 seconds.
Please do not count or tap during the experiment.
Any questions?"

When subjects' questions had been answered and they were ready to begin, the experimenter sent the first stimulus from the Tandon to the speaker. When the subject indicated that he or she was ready the next stimulus was sent and so on. The first three stimuli that each subject heard were practice trials, the responses to them were excluded from analyses. After the practice trials, the 39 experimental bursts were played twice each in a different random order to each subject. Each subject thus made 81 judgements in all.

When the subjects had completed the task they were thanked, debriefed and allowed to leave. Their comments on the study were recorded.

7.1.3. Results.

Subjects' mean duration estimates for the stimuli are shown in Table 7.1.

All of the actual durations were overestimated except for the 2400 ms stimulus when speed was slowest. Clear trends exist for inharmonicity, where increases in inharmonicity decrease perceived duration at all of the actual durations, and for speed, where increases in speed increased perceived duration at all of the actual durations. For pitch, increases in pitch to resulted in a decrease in perceived duration at 1600 ms and an increase at 800 ms. For resolution, the only clear trend was for more resolved stimuli to be perceived as being longer at 2400 ms. These trends are shown more clearly in Figs. 7.1-7.4.

Five two way within subjects analyses of variance (stimulus duration by parameter level) were conducted upon the mean duration judgements for all of the parameters except repetition (Table 7.2). For each parameter there was a significant effect of actual stimulus length on duration judgement. For the pitch stimuli, there was no significant effect of pitch level ($F(2,28)=1.076$, $p=0.354$) and no interaction. For the inharmonicity stimuli there was no significant effect of level of inharmonicity upon duration judgements ($F(2,28)=2.804$, $p=0.078$) and no interaction. A significant effect of speed upon the mean duration judgements was found ($F(2,28)=8.544$, $p=0.001$) and a significant interaction ($F(4,56)=4.039$, $p=0.006$). The interaction is shown in Fig. 7.3. There was also a significant effect of resolution upon duration judgements ($F(2,28)=4.024$, $p=0.029$) but no significant interaction ($F(4,56)=2.219$, $p=0.079$).

As suggested by Brown(1985) duration estimates were converted to directional and absolute error scores. For the directional measure of error, each duration estimate was divided by the actual duration of the stimulus. A score of less than 1 therefore represented an under estimation and a score of more than 1 represented an overestimation. Absolute error scores were calculated by subtracting the estimated from the actual duration and ignoring the sign of the difference. These scores were converted to percentages of the actual duration.

		ACTUAL LENGTH					
		2400		1600		800	
PARAMETER	LEVEL	Estimated duration (sec)					
		MEAN	ST.DEV	MEAN	ST.DEV	MEAN	ST.DEV
	200	3.331	1.289	2.185	0.919	1.293	0.650
PITCH	450	2.962	1.184	2.152	0.946	1.332	0.683
	700	3.098	1.438	2.130	0.951	1.378	0.695
	0	3.148	1.256	2.195	0.947	1.269	0.653
INHARMONICITY	5	2.870	1.097	2.161	0.859	1.227	0.573
	12	2.845	1.168	2.042	0.869	1.226	0.603
	220	3.185	1.133	2.040	0.957	1.339	0.628
SPEED	314	2.801	0.987	1.895	0.729	1.069	0.498
	733	2.342	0.971	1.679	0.872	1.065	0.514
	A	2.873	1.079	2.009	0.841	1.186	0.489
RESOLUTION	U	2.967	1.004	2.257	0.850	1.167	0.509
	R	3.264	1.203	2.148	0.853	1.226	0.534
	6	3.157	1.098				
REPETITION	4			2.016	0.845		
	2					1.154	0.570

(*A=atonal, U=unresolved, R=resolved).

TABLE 7.1 : SUBJECTS MEAN DURATION ESTIMATES

Fig. 7.1 : Mean Duration Estimates for Pitch Stimuli

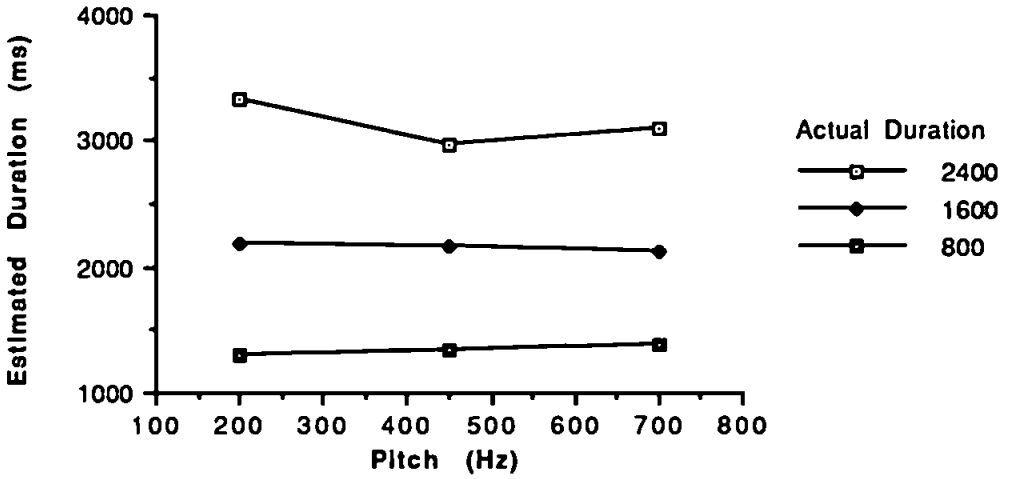


Fig.7.2 : Mean Duration Estimates for Inharmonicity Stimuli

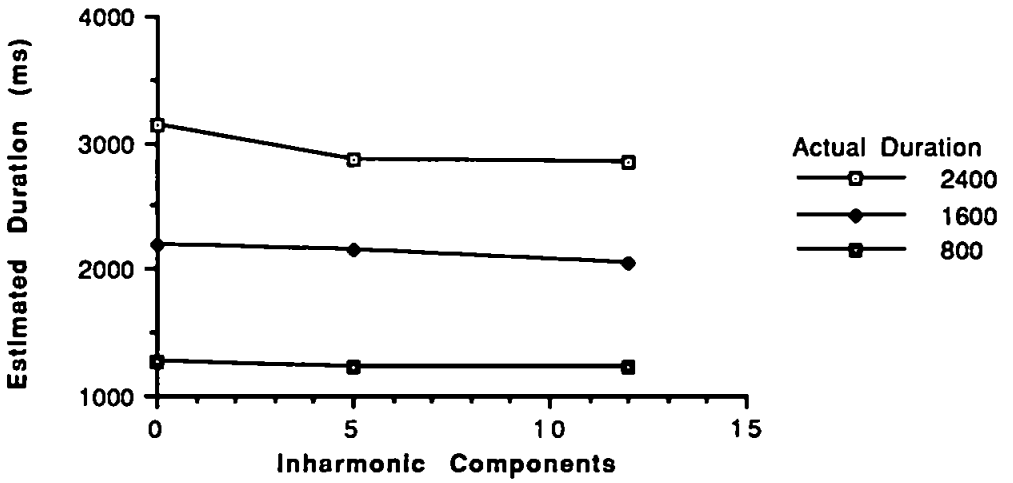


Fig.7.3: Mean duration Estimates for Speed Stimuli

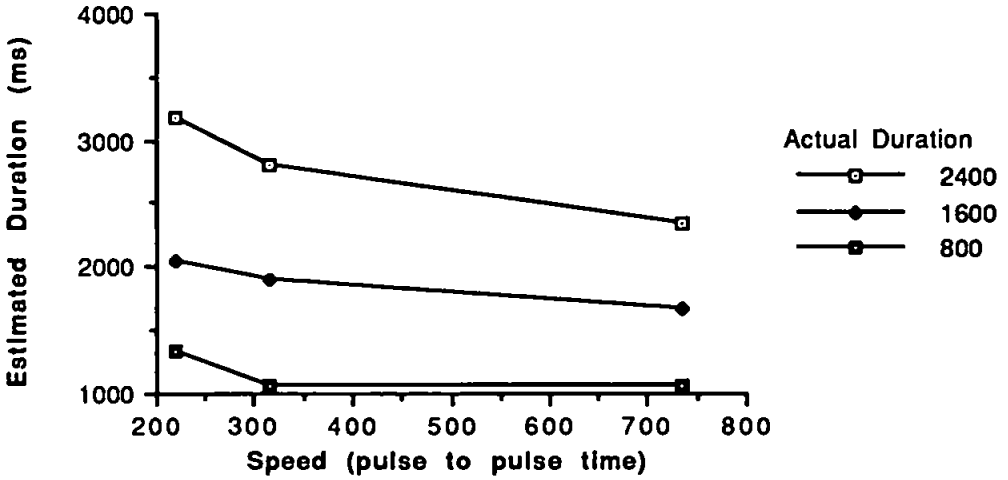
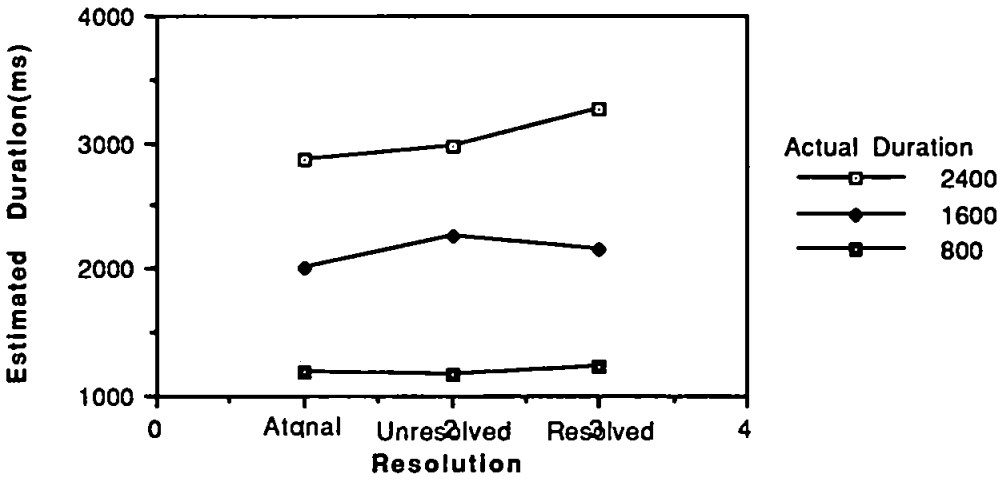


Fig.7.4: Mean Duration Estimates for Resolution Stimuli



SOURCE	SUM OF SQ	df	MEAN SQ	F	P
PITCH					
Within Cells	100.77	14	7.197		
Duration	72.75	2	36.37	72.72	0.000
Error(Duration)	14.005	28	0.500		
Level	0.332	2	0.166	1.07	0.354
Error(level)	4.325	28	0.154		
Dur*level.	0.792	4	0.198	1.233	0.307
Error(D*L)	8.993	56	0.160		
INHARMONICITY					
Within Cells	86.68	14	6.196		
Duration	66.12	2	33.06	85.92	0.000
Error(Duration)	10.77	28	0.384		
Level	0.659	2	0.329	2.80	0.078
Error(Level)	3.292	28	0.117		
Dur.*level.	0.401	4	0.10	0.877	0.483
Error(D*L)	6.410	56	0.114		
SPEED					
Within Cells	66.141	14	4.724		
Duration	59.19	2	29.59	107.2	0.000
Error(Duration)	7.73	28	0.276		
Level	5.480	2	2.740	8.544	0.001
Error(level)	8.979	28	0.320		
Dur.*Level.	1.602	4	0.400	4.039	0.006
Error(D*L)	5.553	56	0.099		
RESOLUTION					
Within Cells	73.35	14	5.239		
Duration	76.34	2	38.17	108.5	0.000
Error(Duration)	9.845	28	0.351		
Level	0.814	2	0.407	4.024	0.029
Error(Level)	2.834	28	0.101		
Dur.*Level.	0.921	4	0.230	2.219	0.076
Error(D*L)	5.815	56	0.103		

TABLE 7.2 : TWO WAY ANOVA, (DURATION * URGENCY LEVEL) FOR EACH PARAMETER

Thus, judgements were expressed as proportions of the actual duration (ratios) and absolute error in estimates of different durations were represented on the same scale.

The mean error and ratio scores for each parameter are shown in Figs. 7.5-7.12. Generally increases in actual stimulus length decreased overestimation (ratio scores) and absolute error. The figures also show that the only clear effect of the parameter level upon the ratio and error scores was for increases in speed to increase both the amount of overestimation and the absolute error of judgements.

For the repetition stimuli, the mean absolute error scores were 61.74, 50.99 and 44.29 for the low, medium and high repetition stimuli respectively. Stimuli containing more units of repetition were thus less overestimated.

7.1.4. Discussion.

The results for speed, pitch, and inharmonicity support the predictions of the information processing accounts of perceived duration. The finding that increases in resolution result in increases in perceived duration cannot be accounted for by the information processing theories. This finding supports the Reiss Jones contrast model.

For all the parameters in the Anova there was a significant effect of actual length on judgement. This was expected and implies that subjects could differentiate the stimulus lengths.

The finding that there was no significant effect of pitch upon perceived duration supports the information accounts of perceived duration. Increases in pitch do not increase the information content of the stimulus and were thus not expected to increase perceived duration. Bringer(1988) found that increases in pitch resulted in increases in perceived duration. The pitch values that he

Fig. 7.5 : Pitch - Effect of Stim. Length on Actual(A)/Estimated(E) Duration Ratio

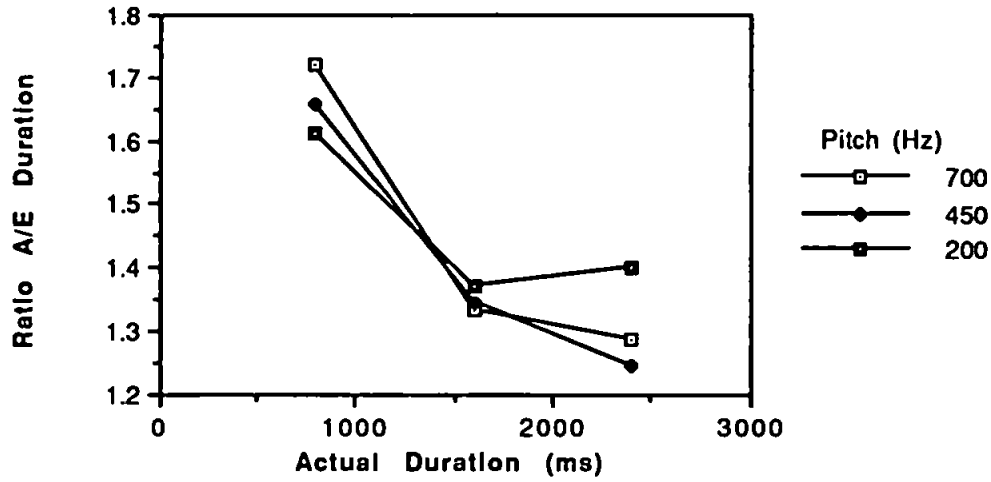


Fig. 7.6 : Pitch - Effect of Stim. Length on Absolute Error

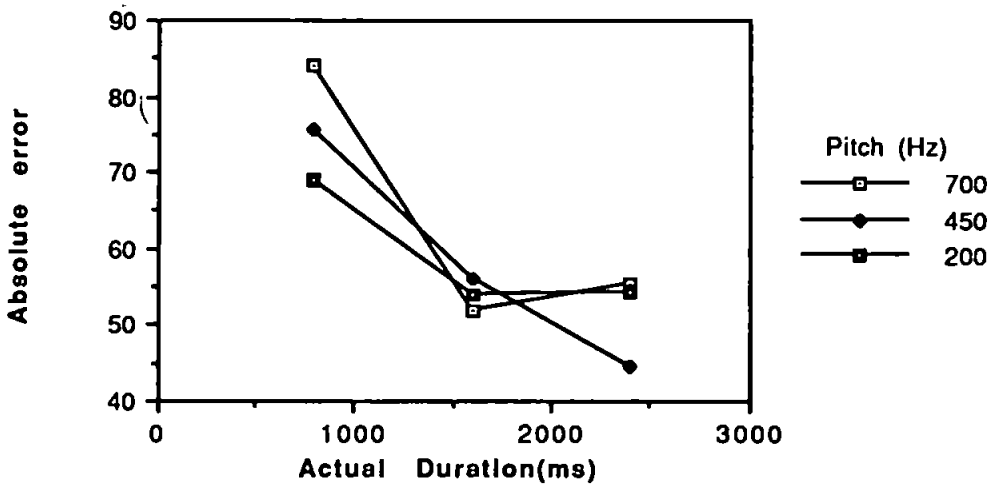


Fig.7.7 : Inharmonicity - Effect of Stim. Length on A/E Duration Ratio

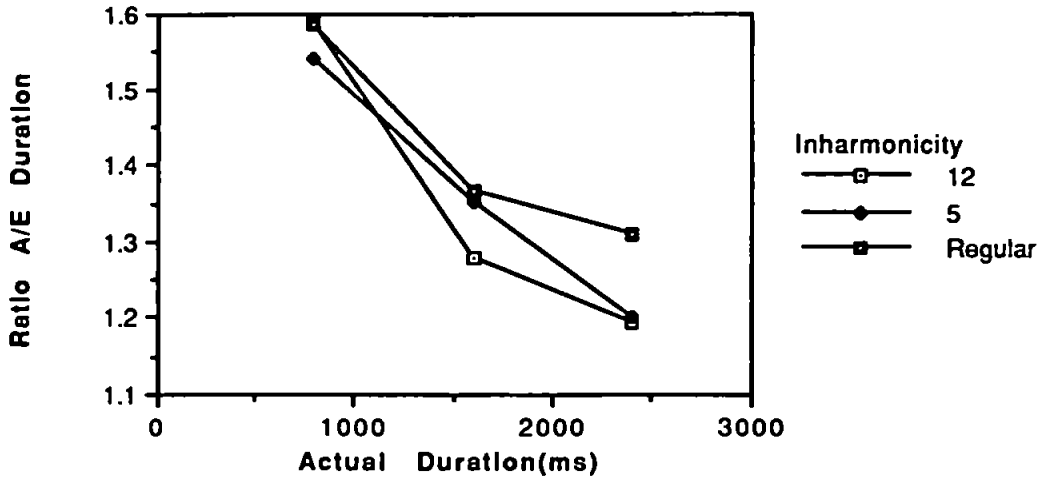


Fig. 7.8 : Inharmonicity - Effect of Stim. Length on Absolute Error

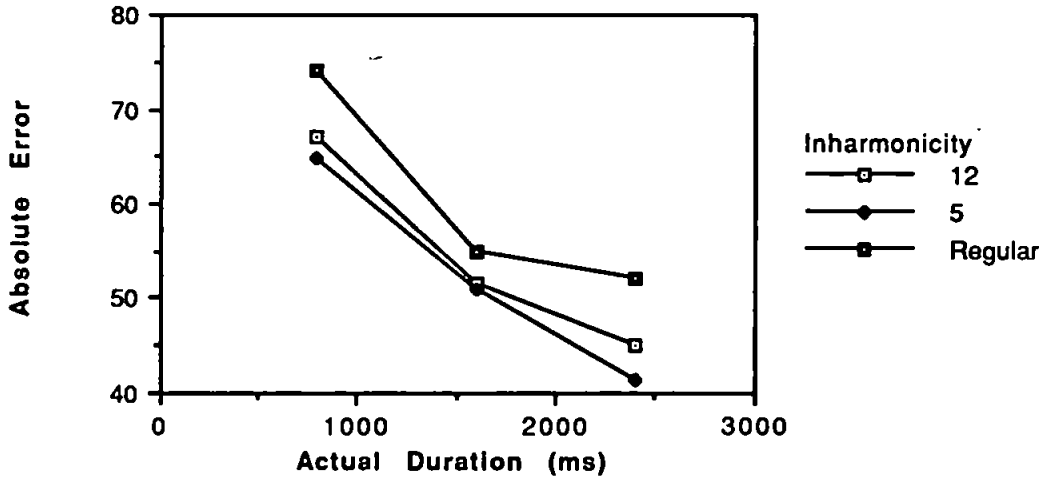


Fig. 7.9 : Speed - Effect of Stim. Length on A/E Duration Ratio

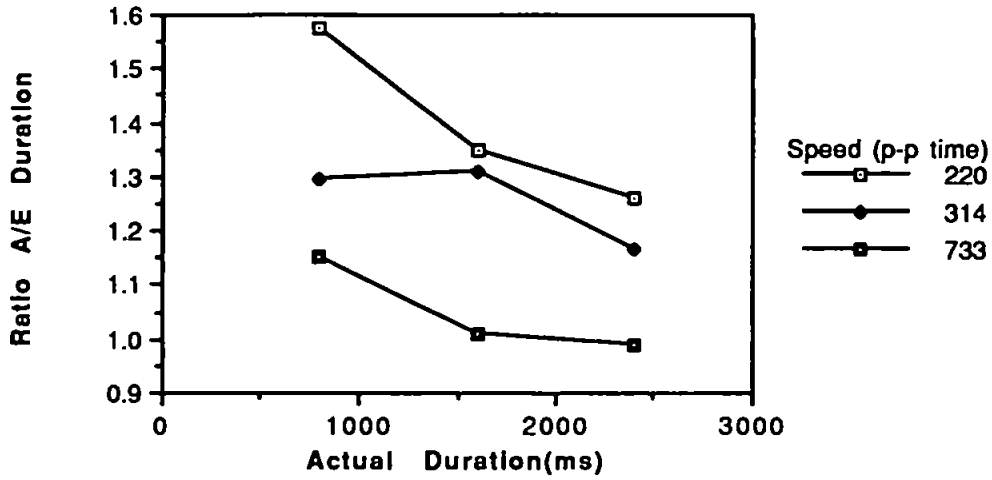


Fig. 7.10 : Speed - Effect of Stim. Length on Absolute Error

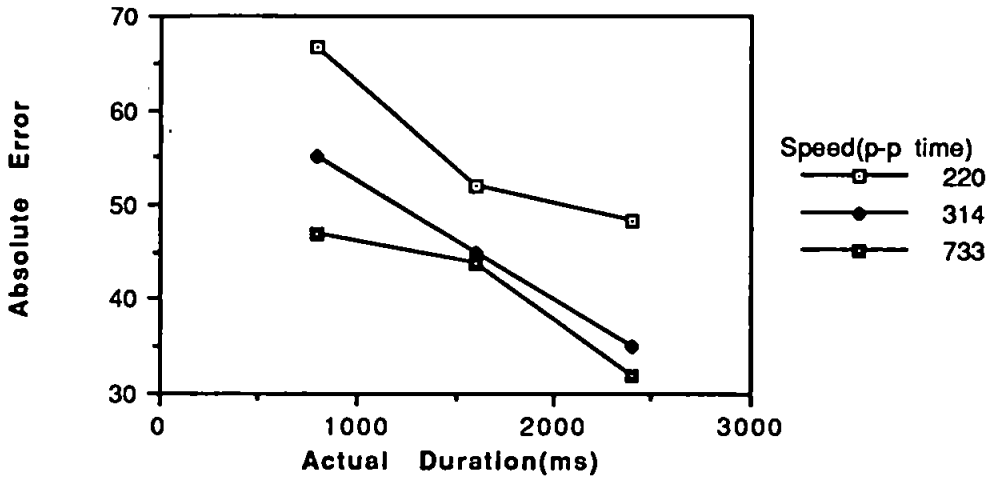


Fig. 7.11 : Resolution - Effect of Length on A/E Duration Ratio

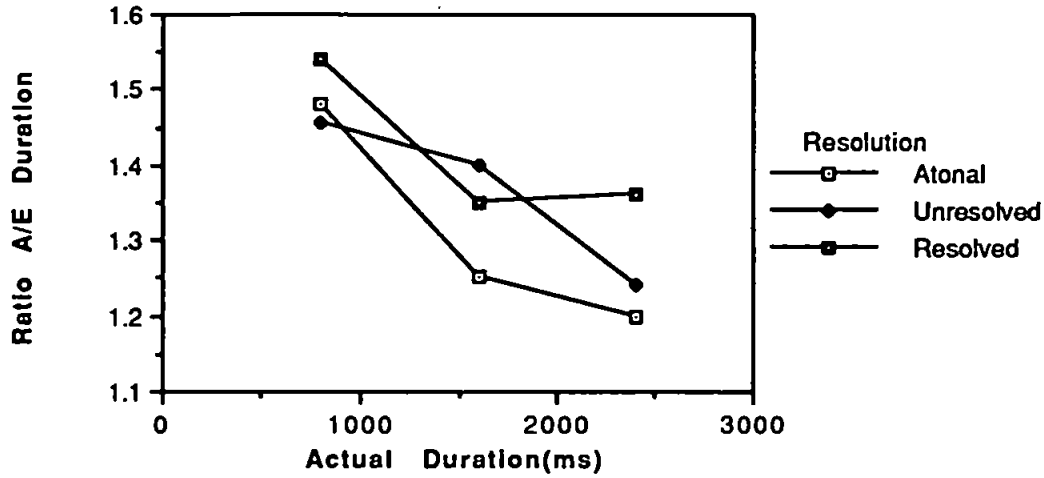
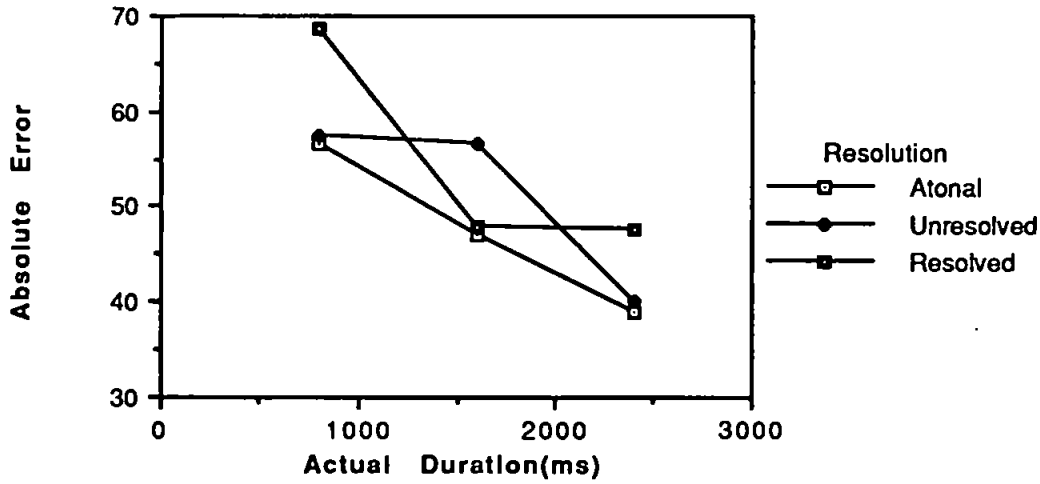


Fig. 7.12 : Resolution- Effect of Stim. Length on Absolute Error



used (4000 and 500Hz) were much higher and more widely separated than those employed in this study. It is possible that differences in perceived duration exist across such a large range or for high pitches, but not across the smaller range and at the lower levels employed here. Since the present study explored the effect of perceived duration upon the range of pitches employed in Experiment Five, and on pitches that could be used in auditory warning design, the present findings were accepted - for the range of pitches relevant to auditory warnings work there appears to be no significant effect of pitch on perceived duration. It should also be noted that Bringer used a different procedure, adjustment, and a shorter actual duration, 385ms.

The finding that there is no significant effect of inharmonicity upon perceived duration also supports the information processing models. Increases in inharmonicity do not increase the information content of the stimulus and were thus not expected to increase perceived duration. However a clear, near significant ($p=0.07$), trend exists for increases in the number of inharmonic components to decrease perceived duration.

The significant effect of speed upon perceived duration was in the opposite direction to tthe trend shown by inharmonicity - increases in stimulus speed were found to increase perceived duration. This supports the information processing accounts of perceived duration because increases in speed do increase the information content of the stimulus and are thus expected to increase perceived duration. The significant interaction between the level of speed and actual duration (Fig. 7.3) showed that when the stimulus was very short and very fast there was an increase in overestimation, ie that stimulus increased perceived duration especially.

One aspect, shown in Fig 7.9

It is interesting to note the similarity between this interaction and the interaction in Experiment 9. In Experiment 9 when the stimulus was very short and fast it was perceived as being especially non urgent. Here it has been shown that such stimuli result in an extra large increase in perceived duration. It is tempting to conclude from this that increases in perceived duration correspond with decreases in perceived urgency. However the present experiment has shown that increases in speed increase both perceived duration and perceived urgency. Thus a contradiction exists, Experiment 9 demonstrated that a very short fast stimulus is perceived as very non urgent, Experiment 10 showed that such a stimulus

increased perceived duration. It cannot be concluded however that decreases in perceived urgency correspond to increases in perceived duration because speed has been shown to increase both factors.

A possible explanation for the interaction between speed and duration can be found by examining the ratio and error scores. These scores showed that for all parameters subjects' estimates became more accurate as the actual duration of stimuli increased. (This could be because subjects were more used to judging longer durations and were therefore more accurate). Any effect of increased perceived duration upon urgency would be most pronounced at the shorter durations, where most overestimation occurs. At shorter durations stimuli were judged particularly non urgent as shown in Experiment 9. The ratio and error scores show that increases in speed also increase overestimation and error . These findings are reflected in the interaction shown in Fig. 7.3 in which the shortest fastest stimulus was most overestimated.

It is possible that usually increases in speed result in increases in perceived duration, as has been demonstrated here. Length has been shown to increase perceived urgency (Hellier and Edworthy 1990), and so the increase in apparent length could contribute to the effect of increasing urgency as speed increased. The ratio and error scores have shown that overestimation occurs especially in very short or very fast stimuli. A short fast stimulus is thus overestimated on two counts. If it is assumed that the effect of these overestimations are additive, then is possible that the overestimation due to speed could increase urgency by makng the stimulus appear longer, but if the stimulus was also short and so further overestimated it could appear so long that perceived information per unit time (speed) would be decreased and so urgency would be reduced, as in Experiment 9.

This explanation is superior to that proposed in Experiment 9 which saw the decrease in urgency of the very short fast stimulus as being caused by a decrease in perceived duration. This study shows that this is not a viable explanation because the perceived duration of a similar stimulus is increased not decreased. The explanation can also account for the fact that the interaction at the very short actual duration was only demonstrated for the speed parameter. This is the only parameter for which an increase in perceived duration would decrease the urgency of the stimulus by decreasing information per unit time. Until more is known about the interaction is it recommended that short fast stimuli such as these are

avoided in auditory warning design since they distort what we know about the effects of parameter changes upon perceived urgency.

Increases in resolution were also found to increase the perceived duration. This supports the work of Boltz(1989) who found that unresolved stimuli were perceived as being shorter that resolved ones, but cannot be accounted for by the information processing theories since increases in resolution are not associated with increases in information content. The apparent, but insignificant, interaction (Fig. 7.4) appears to be the result of the resolved stimulus being perceived as shorter than the unresolved one at the mid duration. In general terms though another parameter that has been shown to affect perceived urgency (Edworthy et al 1991) has been shown to increase perceived duration. In this case however, increases in perceived duration are associated with decreases in perceived urgency for more resolved stimuli.

For the stimuli that varied in units of repetition the findings are very ambiguous. It is not clear whether the effect of decreasing perceived duration as the repetitions increase is caused by the parameter change itself , or by the fact that, as shown by the ratio and error scores for all parameters, at longer durations subject's judgements are more accurate, ie they overestimate the stimuli by less and thus perceived duration is decreased. The latter explanation is favored since the effect was demonstrated with all parameters as actual stimulus length increased.

The findings that increases in pitch had no effect on perceived duration and that increases in speed increased perceived duration supported the information processing theories of perceived duration such as those posed by Hicks et al(1974), Thomas et al (1977), Frankenhauser(1959), Brown et al (1988) and Fraisse(1981). In the former instance there were no increases in memory requirement, cognitive change or mental content, therefore no increase or decrease in perceived duration was predicted. In the case of speed, increases in speed corresponded to increases in stimulus information and thus increases in cognitive change/mental content/ memory requirement. The information processing theories predicted that these changes would have increased perceived duration.

Results from the other two parameters, resolution and inharmonicity, pose problems for the information processing account of perceived duration. The

finding that more resolved stimuli were perceived as being longer is problematic for such stimuli contained no more information and did not pose any greater 'processing load'; they therefore should not have been perceived as being any longer than unresolved stimuli. This finding can however be accounted for by the Reiss Jones model of perceived duration which says that unresolved stimuli violate an expected ending and are thus underestimated relative to resolved stimuli which communicate the expected ending. The non significant trend for increases in inharmonicity to result in decreases in perceived duration cannot be explained by the information processing theories because increasing inharmonicity is not associated with any decrease in information content etc.

To summarise, the results from the speed data imply that increases in perceived duration correspond to increases in perceived urgency, whereas the resolution data implies that increases in perceived duration are associated with decreases in perceived urgency. Other parameters that can communicate increases in perceived urgency, pitch and inharmonicity, do so without an associated change in perceived duration. The information processing account of perceived duration cannot fully account for these findings. It seems that changes in perceived duration may be associated with increases in perceived urgency, but that changes in perceived duration are not necessary components of urgency. The Reiss Jones account of perceived duration can account for the findings of the resolution stimuli. In the next study an attempt is made to explain the findings for the other parameters also in terms of her contrast model.

7.2. Experiment Eleven : An Investigation into the Types of Parameter that Affect Perceived Urgency.

7.2.1. Introduction.

This study was designed to explore the Reiss Jones approach to perceived duration to see if it can offer a satisfactory to explanation for the findings in Experiment Ten, that only some of the parameters which communicate perceived urgency are associated with changes in perceived duration, and that for speed an increase in perceived duration is associated with increased urgency whereas for resolution it is associated with decreased perceived urgency. These findings could not be fully accounted for by the information processing approach to time perception. The possibility that there are two different types of acoustic parameter is explored.

The results of Experiment Ten show that four parameters that have been shown to affect perceived urgency have different effects on perceived duration. For speed, increases in perceived duration contribute to the effect of increased urgency. For resolution, increases in perceived duration are associated with decreases in perceived urgency. For pitch and inharmonicity perceived duration does not appear to contribute to urgency. The different results for the different parameters could perhaps be explained if the parameters could be shown to be qualitatively different.

There are two parameters for which changes that decrease perceived urgency have been shown to increase perceived duration. For resolution this was a significant effect, for inharmonicity this was a non-significant trend. For one parameter (speed) changes that increase perceived urgency have been shown to increase perceived duration. One way in which these two groups of parameter might differ is in the extent to which responses to them are culturally determined, with resolution and perhaps inharmonicity being more culturally defined than speed.

Pitch will not be considered at this stage because no relationship between pitch and perceived duration was demonstrated in the previous experiment. It is also not easy to see the extent to which responses to pitch are culturally determined. Furthermore, pitch may be a special case, or a different type of parameter altogether. Of the parameters originally scaled it was the most salient or discriminable, and was the only one that would be classified in Stevens terms as Metathetic.

If the distinction between cultural and non cultural parameters holds for the parameters apart from pitch, then for culturally determined parameters, decreases in perceived urgency are associated with increases in perceived duration. Thus as perceived urgency increases, perceived duration decreases and the stimulus appears to be shorter. This supports an intuitive notion of urgency whereby something that is more urgent is rushed (for it takes place in an apparently shorter period of time), it is apparently faster (and we know anyway that faster things appear more urgent). For more 'fundamental' parameters such as speed increases in perceived urgency are accompanied by increases in perceived duration . In this case part of what makes the sound urgent could be the increase in perceived length - we know that increases in perceived length result in increases in perceived urgency. Thus it is possible to see how both increases and decreases in perceived duration could contribute to the impression of urgency.

Theoretical support for the idea that there might be two different types of parameter comes from the work of Riess Jones and Boltz(1989). Their contrast model suggests that stimuli vary on a continuum of structural coherence. It was said that more coherent stimuli preserved objective accent regularities and involve simpler structural hierarchies. They said that two forms of attending were possible to any stimulus. Future orientated attending was said to involve global attending over periods higher than the referent(the basic time span of the stimulus eg. one beat), and analytic attending was said to occur at low levels of a stimulus' hierarchy. It was said that future orientated attending was more likely to occur to coherent stimuli, and that analytic attending was more likely to occur to non coherent, non hierarchical stimuli.

The two different forms of attending were said to have different implications for duration judgements of a stimulus event. During analytical attending the authors said that subjects employ mnemonic activities such as monitoring information content/ change to aid duration judgements. During future orientated attending however the authors said that expectancies are established as to the ending of the event, and temporal contrast occurs when these expectancies are violated. Perceived duration is thus influenced by temporal contrast, events that appear to end too soon are judged as shorter and those that appeared to end too late are judged as longer.

If the experimental parameters are placed along the 'continuum of structural coherence' it is possible to account for the contradictory findings in Experiment 10 concerning perceived duration. It is possible to see that, as stated by Boltz(1989), stimulus events varying in resolution are coherent and hierarchical, for they communicate an expected ending, whereas those varying in speed are less hierarchical and noncoherent . Thus future orientated attending will occur to resolution stimuli, and this will lead to negative temporal contrast when subjects are required to judge the perceived duration of the unresolved stimuli and so to them being judged shorter. Mnemonic strategies applied as a result of analytic attending to the speed stimulus will mean that increases in speed increase perceived duration as the information content of the stimulus increases, as was found and as the information processing theorists predicted.

It is not clear where stimuli varying in inharmonicity would lie along this continuum, although it is thought that the pitch changes caused by variations inharmonicity would make them more coherent than speed stimuli. Because of this uncertainty and because the effects of inharmonicity upon perceived duration did not reach significance in the previous experiment, the parameter is excluded from consideration for the time. At this early stage in the development of the present ideas only speed and resolution will be considered. Their effect on perceived duration is known, and is contrasting, and it is easy to suppose where they might lie on the continuum of structural coherence.

It is possible that Reiss Jones' 'coherent and noncoherent' events may correspond to our own 'cultural and fundamental parameters. The cultural parameters that have been discussed might be similar to Reiss Jones coherent events, with her coherence and expectation being generated by cultural norms, and cultural parameters being attended in a future orientated manner. The fundamental parameters might correspond to Reiss Jones noncoherent events, they would thus be attended to analytically. This would explain the different results for the parameters in terms of perceived duration. It appears that changes in perceived duration contribute to perceived urgency. For cultural parameters decreases in perceived urgency make the parameter faster and thus more urgent whereas for fundamental parameters changes in perceived duration make the parameter longer and thus more urgent.

Experiment 11 was designed to test the hypothesis that there are cultural and fundamental parameters that contribute to perceived urgency. Subjects were

varied in terms of their musical experience, and required to judged the duration of sounds. Non musical subjects were expected to be less attuned to cultural cues than musical subjects. For musical subjects, musical training would have taught them to attend to the temporal patterns in sound. For fundamental parameters the two groups were expected to be equal in their ability to attend to the sounds. If the two parameters were different it is predicted that the results for musical and non musical subjects will be the same for speed stimuli since it is a fundamental parameter. For resolution stimuli however it is predicted that the effect of increased perceived duration and decreased urgency as stimuli become more resolved will be larger for the musical than the non musical subjects. The musical subjects were expected to be more attuned to generating expected endings on the basis of temporal cues and thus temporal contrast should be greater for them, that is the resolved stimuli should be more overestimated relative to the unresolved ones. Subjects' urgency judgements were also collected to see if the urgency/duration relationship exhibited in previous experiments was again demonstrated, and to see if it was affected by musical competence.

7.2.2. Method.

7.2.2.1. Subjects.

Twenty five non musical and fifteen musical subjects took part in the study. The non musical subjects had no musical training (music lessons etc) whatsoever. The musical subjects had all passed Grade 7 or 8 in any instrument or demonstrated a high degree of musical sophistication such as performing in an advanced orchestra or choir. The sample sizes were unequal because it was difficult to obtain the highly trained musical subjects. Of the non musical subjects, there were 20 females and 5 males, there ages ranged from 18-35 years. In the musical sample there were 8 females and 7 males, their ages ranged from 18 to 60 years.

The nonmusical subjects and four of the musical subjects were undergraduate Psychology students from Polytechnic South West. They volunteered to participate in the study in partial fulfilment of their coursework requirement. The remaining eleven musical subjects were contacted through personal contacts and were paid one pound for volunteering to participate in the study.

Four of the subjects had previously taken part in similar psychophysical studies.

7.2.2.2. Materials.

The laboratory and hardware arrangements were as reported in Experiment One, except that subjects sat in a sound attenuated booth instead of in Laboratory Two.

The stimuli employed in this Experiment are detailed in Appendix 7B. There were nine stimuli in all, three practice stimuli, three stimuli that varied in speed from fast to slow and three stimuli that varied in resolution from resolved to unresolved to atonal. The speed stimuli were described in terms of their pulse to pulse times (220, 314, 733 ms) as in Experiment 10. All of the stimuli were 2400ms in length, they were taken from the speed and resolution stimulus sets used in Experiment 10. The longest stimulus durations were used here because Experiment 10 showed that they were the most accurately judged.

The actual duration of the stimuli was not varied in this experiment since it was the possible difference between musical and non musical subjects that was of interest. Experiment 10 demonstrated the effects held for different stimulus durations.

7.2.2.3. Procedure.

Subjects were run one at a time while seated at a desk in the sound attenuated booth. The speaker was approximately 0.5m away from them in the booth. They were told the broad nature of the study and their wrist watches were removed. Subjects then read one of the following sets of instructions adapted from Engen(1977) and from Bobko et al (1977);

"I am going to present you in irregular order a series of sounds. Your task is to tell me how urgent they are by assigning numbers to them. When you have heard the first sound give its urgency a number - any number that you think appropriate. I will then present another to which you will also give a number and a third etc. Let high numbers represent high urgency and let low numbers represent low urgency. Try to make the ratios between the numbers that you assign to the different sounds correspond to the ratios between the urgency of the sounds. In other words try to make the numbers proportional to the urgency of the sound as you hear it. Remember that you can assign any number. There is no limit to the number that you assign. There is no right or wrong answer. I want to know how you judge the urgency of the sounds.

Any questions?"

" I am going to present you in irregular order a series of sounds. Your task is to estimate, to the nearest fraction of a second, how long you think each sound is on for. Try to estimate the first sound as accurately as you can in seconds and fractions of a second. Thereafter try to keep your judgements proportional. For example, if you think that the first sound is 1.25 seconds long, any sound that you think is twice as long should be judged 2.5 seconds and any sound that you think is half as long should be judged 0.625 seconds. Please do not count or tap during the experiment.
Any questions?"

Half of the musical subjects and half of the nonmusical subjects read the duration instructions first and the other half read the urgency judgements first. Thus half of the subjects made urgency judgements first and half made duration judgements first.

The first three stimuli that each subject heard were the practice stimuli, (the responses to these were not analysed). Thereafter the six experimental stimuli were played twice each, in a different random order to each subject. When subjects had completed these fifteen judgements they read the instructions that they had not previously seen, either duration or urgency. Six experimental stimuli were then played in a different random order to each subject. Subjects thus made twenty seven judgements in all, three to practice stimuli, twelve urgency judgements and twelve duration judgements.

When subjects had finished their watches were returned, they were thanked, debriefed and allowed to leave. Their comments on the study were recorded.

7.2.3. Results.

Subjects mean duration and urgency judgements to the speed and resolution stimuli are shown below in Table 7.3.

	Speed Stimuli			Resolution Stimuli		
	Fast	Med	Slow	Atonal	Unres	Res
	Duration					
Musical(Mean)	2.78	2.66	2.24	2.73	2.66	2.96
(St.Dev)	0.644	0.814	0.826	0.924	0.794	0.842
Non Musical	3.72	3.70	2.94	3.75	3.79	3.86
(St.Dev)	1.32	1.53	1.06	1.79	1.66	1.55
	Urgency					
Musical	28.13	24.32	16.70	23.95	20.27	17.47
(St.Dev)	25.1	20.1	15.1	24.4	19.9	19.2
Non Musical	18.22	16.49	9.98	12.43	11.21	10.55
(St.Dev)	24.4	22.0	14.1	16.7	14.5	11.9

TABLE 7.3 : SUBJECTS MEAN JUDGEMENTS TO SPEED AND RESOLUTION STIMULI

As predicted, increases in speed result in increases in perceived urgency and duration for musicians and non musicians. Similarly, increases in resolution resulted in increases in perceived duration and decreases in perceived urgency for both musicians and non musicians. This visual inspection of the data indicated that the only outcome that was not in the predicted direction was the finding that musicians judged the atonal stimulus as being longer than the unresolved one. The unresolved one was still judged shorter than the resolved one which was the most important part of the prediction. These means are represented in Figs. 7.13-7.16.

Examination of the distribution of responses to each stimulus showed that duration judgements were normally distributed. The urgency judgements however were positively skewed. To overcome this potential bias, subjects median judgements are also presented (Figs. 7.17-7.20). Where trends differed between the means and the medians, these were for non musical subjects when judging the unresolved stimulus as less urgent and shorter than the atonal one (Fig. 7.18, 7.20) and for nonmusical subjects when judging the medium speed stimulus slightly more urgent than the fast one (Fig. 7.19).

Although an assumption of Anova is that responses are normally distributed, Howell(1982) states that providing the skew is all in one direction and providing judgement variance is homogenous, then the procedure is robust enough to cope with the violation of that assumption. Since these criteria were met, a mixed two way analysis of variance (musicality (between subjects) by speed or resolution (within subjects))was conducted upon mean duration and urgency judgements.

The results of the Anovas are presented in Tables 7.4.(urgency judgements) and 7.5 (duration judgements). As is shown in Table 7.4, the effect of speed upon urgency was significant ($F(2,76)=20.1$, $p=0.00$), there was no effect of musicality and no interaction. There was also a significant effect of resolution upon urgency ($F(2,76)=6.87$, $p=0.002$) but again no effect of musicality and no interaction. There was however a non significant trend towards the predicted interaction between musicality and resolution ($F(2,76) = 2.11$, $p=0.127$).

FiG. 7.13 : Mean Urgency Judgements To Speed Stimuli.

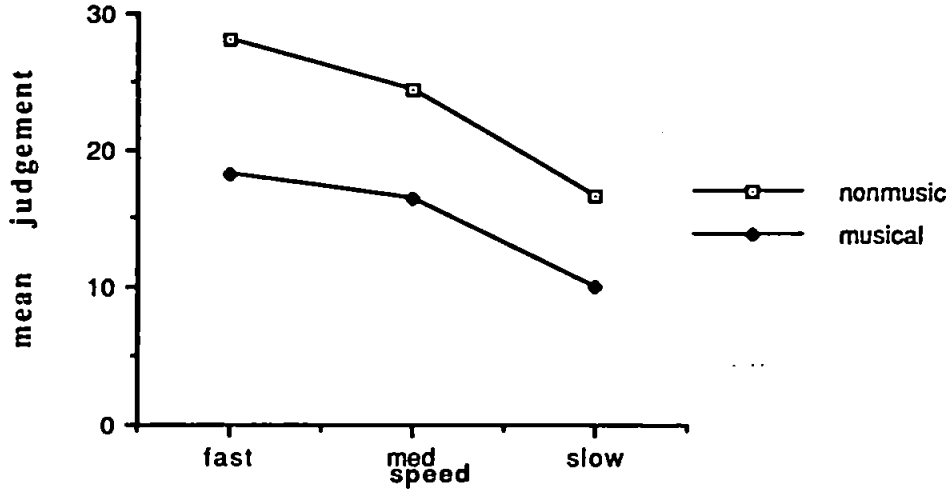


Fig.7.14 : Mean Urgency Judgements to Resolution Stimuli.

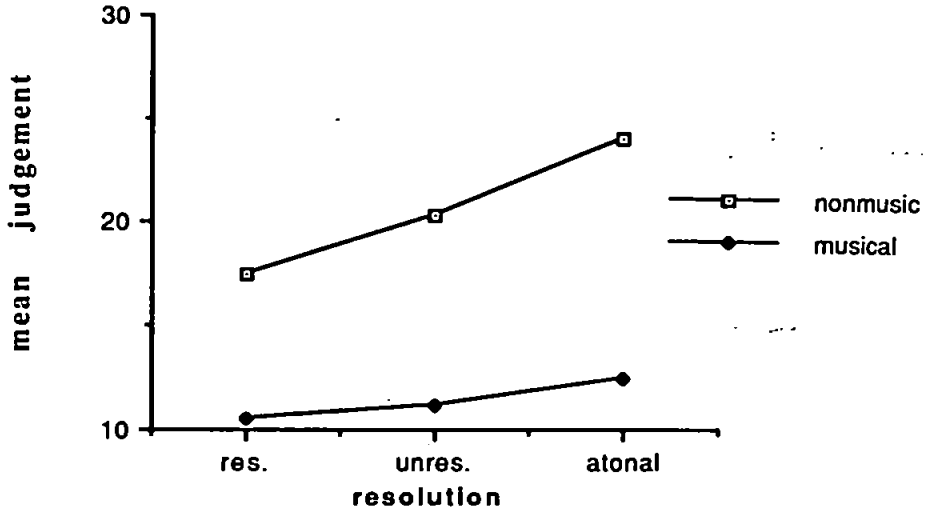


Fig.7.15 : Mean Duration Judgement to Speed stimuli

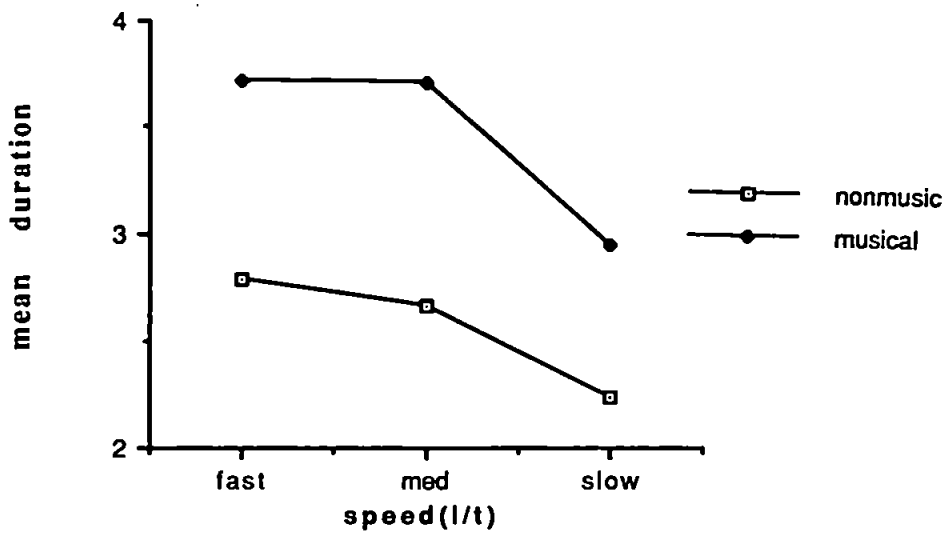


Fig. 7.16 : Mean Duration Judgements to Resolution Stimuli

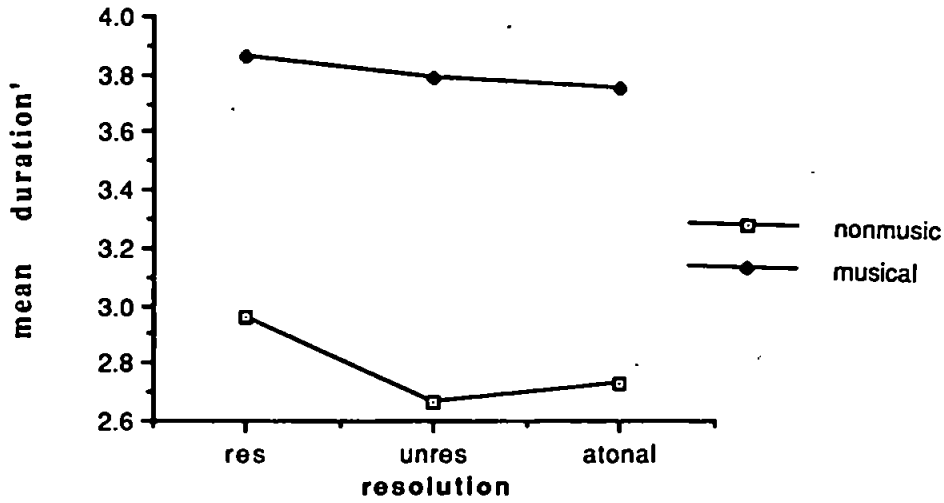


Fig. 7.17 : Median Urgency Judgement to Speed Stimuli

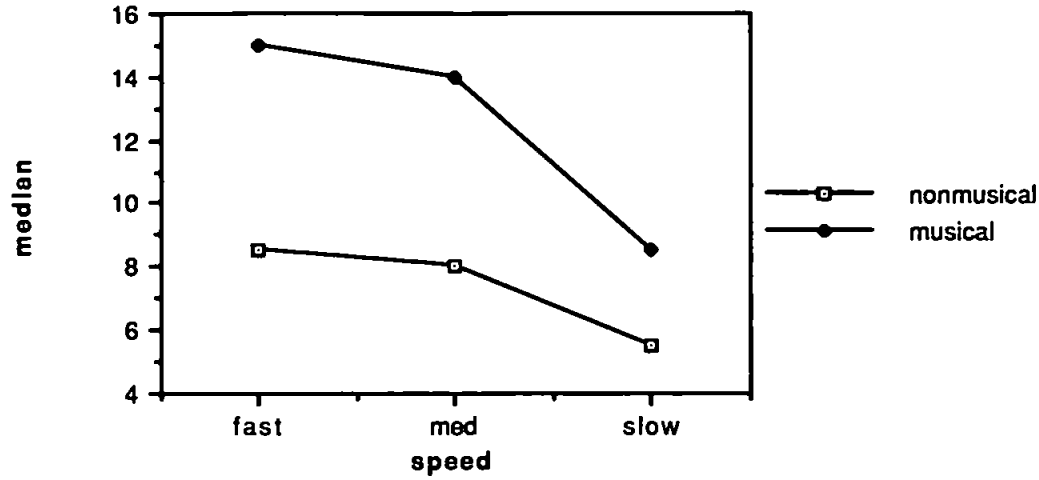


Fig. 7.18 : Median Urgency Judgement to Resolution Stimuli

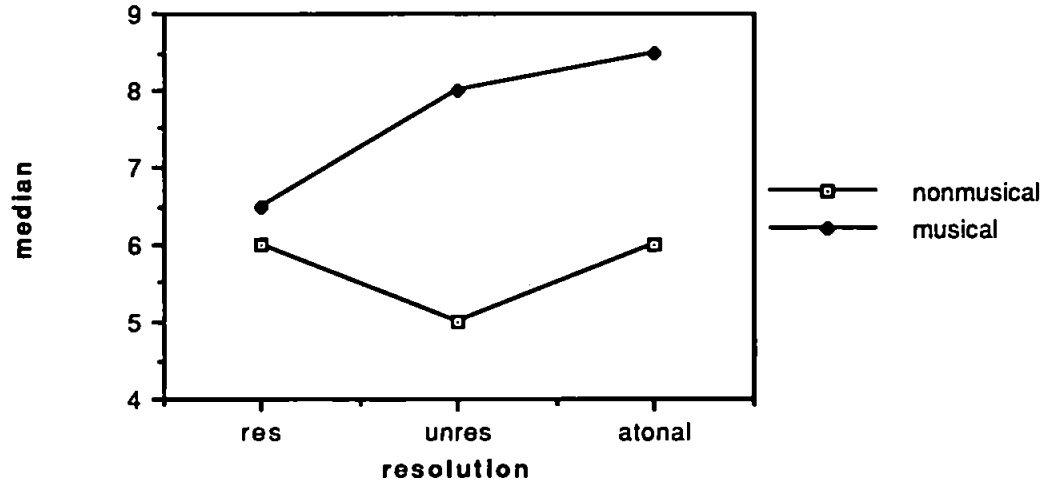


Fig. 7.19 : Median of Duration Judgements to Speed Stimuli

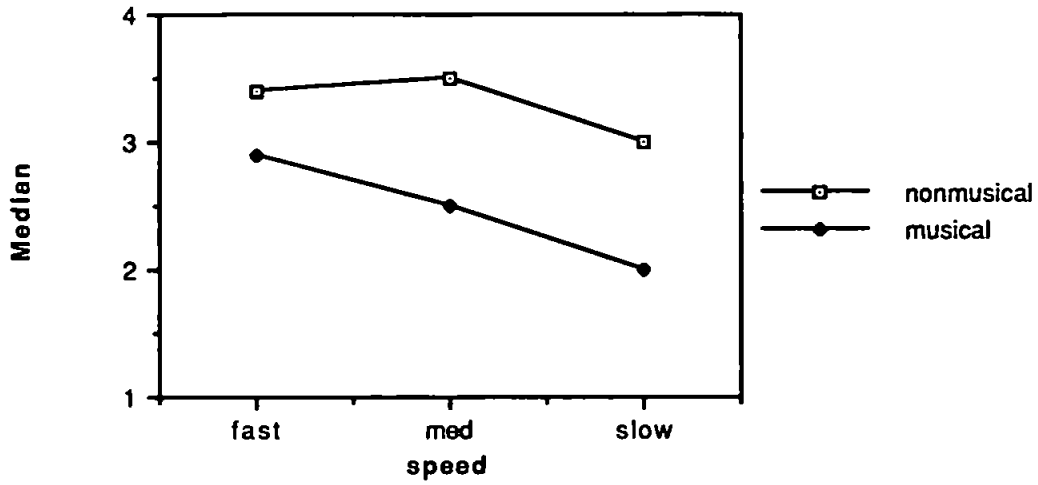
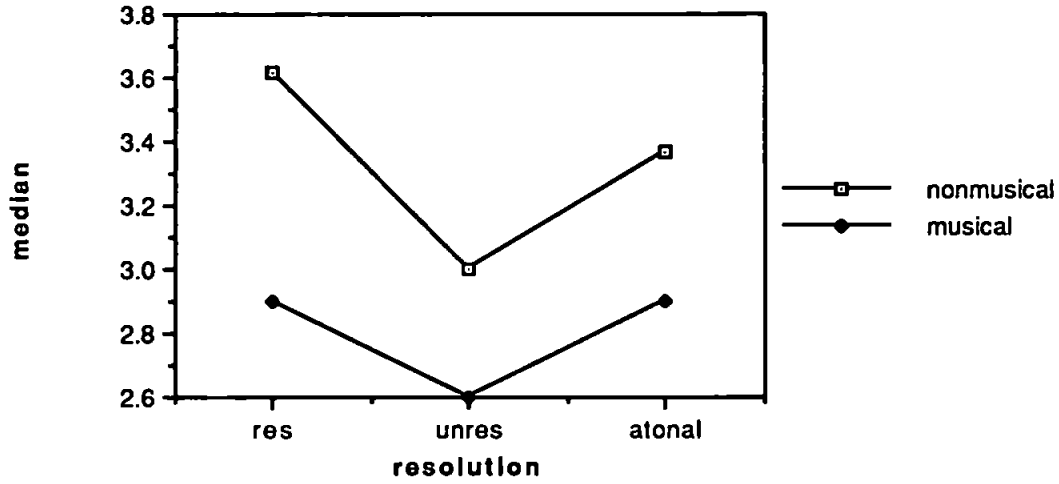


Fig. 7.20 : Median of Duration Judgements to Resolution Stimuli



SOURCE	SUM OF SQ	df	MEAN SQ	F	P
SPEED					
Within Cells	44106.3	38	1160.1		
Musicality	1876.3	1	1876.3	1.61	0.211
Speed	1926.5	2	963.2	20.11	0.00
Error(Speed)	3639.1	76	47.88		
Mus*Speed	49.19	2	24.59	0.513	0.60
RESOLUTION					
Within Cells	33022.7	38	1160.1		
Musicality	2368.8	1	2368.8	2.72	0.10
Resolution	328.0	2	164.0	6.87	0.002
Error(Res.)	1813.9	76	23.86		
Mus*Res.	101.1	2	50.55	2.11	0.127

TABLE 7.4 : TWO WAY ANOVA, (MUSICALITY * PARAMETER LEVEL) FOR URGENCY JUDGEMENTS.

SOURCE	SUM OF SQ	df	MEAN SQ	F	P
SPEED					
Within Cells	115.2	38	3.033		
Musicality	22.28	1	22.28	7.34	0.01
Speed	9.65	2	4.82	20.7	0.00
Error(Speed)	17.67	76	0.23		
Mus*Speed	0.543	2	0.271	1.16	0.316
RESOLUTION					
Within Cells	198.9	38	5.23		
Musicality	28.52	1	28.52	5.44	0.025
Resolution	0.776	2	0.388	1.69	0.190
Error(Res.)	17.38	76	0.228		
Mus* Res	0.257	2	0.128	0.562	0.572

TABLE 7.5 : TWO WAY ANOVA, (MUSICALITY * PARAMETER LEVEL) FOR DURATION JUDGEMENTS.

Table 7.5 shows that there was a significant effect of musicality upon duration judgements to speed stimuli ($F(1,38)=7.34$, $p=0.01$) and a significant effect of speed upon duration judgements ($F(2,76)=20.76$, $p=0.00$), but no interaction. There was also a significant effect of musicality upon duration judgements to resolution stimuli ($F(1,38)=5.44$, $p=0.025$). There was however no significant effect of resolution upon duration judgements and no interaction.

The significant findings of the study were that for both musicians and non musicians, speed increased urgency judgements, and resolution decreased them. Musicians judged the speed and resolution stimuli as being shorter than the non musicians (they were more accurate). For both groups increases in speed increased perceived duration.

7.2.4. Discussion.

The trends exhibited by subjects mean judgements were as predicted, with increases in speed resulting in increases in perceived urgency and perceived duration, and increases in resolution resulting in increases in perceived duration and decreases in perceived urgency, for all subjects. As in Experiment 10, increases in perceived duration corresponded to increases in perceived urgency for speed and to decreases in perceived urgency for resolution.

Where means and medians differed from the expected trends it was with the unresolved and atonal stimuli. Musicians judged the atonal stimulus longer, not as expected shorter, than the unresolved one. This was reflected in the median judgement by the non musicians of the atonal stimulus as more, not as expected less, urgent than the unresolved stimulus. The violation of expected trends by responses to the atonal stimulus does not have serious implications. The theoretical predictions relied on the fact that resolved stimuli communicate an expected ending and unresolved stimuli violate a communicated ending. Atonal stimuli do not communicate an ending at all, it is therefore possible that they are not on a continuum of resolution. Although in the previous experiment they did follow the expected trend, it is not important that they did not in this case because if they are not on a continuum of resolution then responses to them might be ambiguous because they contain no expectancy information. This might mean that

atonal stimuli do not encourage future orientated attending as much as resolved and unresolved stimuli.

In terms of the theoretical predictions, as expected there were no differences in musical and non musical subjects' responses to the speed stimuli. The demonstrated increase in urgency with increases in stimulus speed replicates previous findings and implies that it is a strong effect that holds regardless of musical experience. Although the information processing accounts of time perception are supported by the finding that increases in stimulus speed increase perceived duration, they are not an adequate account. As was demonstrated in Experiment 10, such theories cannot account for increases in perceived duration associated with resolution. The findings support the idea that speed is a fundamental parameter because no difference were exhibited between the responses of musicians and non musicians. The idea that fundamental parameters are at the low structural coherence end of the Reiss Jones continuum is supported because such stimuli are attended to analytically and thus mnemonic strategies would have lead to increases in perceived duration with the information content of the stimulus (more speed), as was demonstrated. The effect of increasing perceived duration would contribute to the effect of increased urgency because longer stimuli are perceived as being more urgent.

The finding that resolved stimuli were perceived as being less urgent also supported the predictions and replicated previous findings (Experiment 10, Edworthy et al 1991). What was unexpected was that there was no significant interaction between musicality and resolution, only a trend towards the interaction when subjects made urgency judgements. Therefore both non musicians and musicians judged this effect to be of the same size. Our theoretical predictions suggested that musicians, being more attuned to cultural music cues, would perceived the violation of the expected ending of the unresolved stimuli more acutely and would judge those stimuli more urgent than the nonmusicians.

The other unexpected finding was that there was no significant effect of resolution upon perceived duration. Since this finding has been demonstrated by Reiss Jones and Experiment 10 we can only conclude that it is weak and did not reach significance in this case. Examination of the graphs indicates that responses to the atonal stimuli may have obscured the effect. In the graphs, the trend remains for unresolved stimuli to be perceived as shorter than the resolved stimuli by all

subjects (means and medians). Fig. 7.16 also indicates that this effect was larger for musicians, as predicted.

The significant effects of musicality upon duration judgements resulted from the superior accuracy of musicians at judging duration. Non musicians consistently gave higher and less accurate duration judgements. This is not unexpected when we remember that musicians are probably more used to judging a set duration regardless of what fills it.

In sum, the findings were as expected if speed is a fundamental parameter, increases in speed increased duration and urgency for musicians and non musicians. To conclude that resolution was a cultural parameter we expected increases in resolution to result in larger increases in perceived duration and decreases in perceived urgency for musicians than for nonmusicians, ie an interaction. Although only the decrease in urgency with increasing resolution reached significance, the trends in the means supported the predictions. It is possible that including atonal stimuli in the resolution stimulus set may have made the effect harder to detect.

In order to investigate more fully the effects of resolution, it is recommended that stimulus sets do not include atonal stimuli. Furthermore, it is possible that in this study the two sets of subjects were too similar in terms of their musical culture to demonstrate conclusively that responses to resolution are culturally determined. Even non musicians are attuned, through daily life to the musical norms in our culture. It is possible that comparing our subjects with subjects from a different musical culture might reveal differences in responding to the resolution, but not the speed stimuli, thus supporting the hypothesis.

This study has provided partial support for the hypothesis that speed and resolution are two different types of parameter. The findings support the idea the speed is a fundamental parameter similar in nature to those at the low structural coherence end of the Reiss Jones continuum. There is limited support for the idea that resolution is a cultural parameter similar to those at the coherent end of the Reiss Jones continuum. Further research needs to be conducted upon the proposed 'cultural' parameters.

CHAPTER EIGHT

CONCLUSION.

This thesis has investigated the perceived urgency of sound in order to improve auditory warning construction. Techniques to measure perceived urgency were selected and used to measure and quantify the effects of different sound parameters upon perceived urgency. The concept of urgency itself was considered, and perceived duration was identified as a contributor to urgency when it was communicated through some sound parameters. It was proposed that there are at least two different types of parameter that can communicate perceived urgency, one innate or fundamental and one cultural. In the following chapter, the conclusions that can be drawn from this program of research are discussed.

8.1. Empirical Findings.

Perceived urgency is a difficult continuum to measure because urgency judgements are subjective in nature. In Chapters Two and Three it was demonstrated that psychophysical techniques, which relate objective quantifiable changes to subjective judgements, provide a viable means of measuring perceived urgency. Such techniques allow objective changes in sound parameters to be related to changes in perceived urgency.

The biases and problems inherent in different psychophysical techniques were investigated and four techniques were selected for further investigation - free modulus magnitude estimation, fixed modulus magnitude estimation, category estimation and cross modality matching. In Chapter Three, the perceived urgency of stimuli that varied in speed was measured by each of the selected techniques. On the basis of the cross modality validation procedure, free modulus magnitude estimation or cross modality matching itself were recommended as the most reliable techniques to use for scaling perceived urgency. All of the techniques demonstrated that increases in speed resulted in increases in perceived urgency.

Having found reliable measurement techniques, the effects of variations in other sound parameters upon perceived urgency were investigated. In Chapter Four cross modality matching was used to measure the effects of variations in pitch, repetition

units and inharmonicity upon perceived urgency. Increases in all of the parameters resulted in increases in perceived urgency. The exponents of the matching functions were used to quantify the amount of change in each parameter that was required to communicate a unit change in perceived urgency. Speed was the most economical parameter through which to communicate perceived urgency because it took smaller changes in speed than in any other parameter to produce a unit change in urgency, inharmonicity was the least economical.

Inharmonicity was harder than the other sound parameters to quantify objectively. Two different ways of quantifying it were tested and it was shown that a linear relationship between perceived urgency and inharmonicity was only demonstrated if inharmonicity was quantified by counting the number of inharmonic components in the stimulus. It was recommended that inharmonicity was described in this way so that the effect of variations in inharmonicity upon perceived urgency could be quantified.

In Chapter Five the different sound parameters were combined in the same stimuli, and an attempt was made to see if any one parameter had more influence on urgency judgements, that is, was more salient or discriminable. The experiments thus far resulted in four matching functions that quantified the changes in each of four parameters (speed, pitch, repetition units and inharmonicity) that had to be made to produce a unit change in perceived urgency. It was therefore possible to use the matching functions to calculate how to communicate an equal level of urgency in each parameter. Stimuli were constructed that communicated theoretically equal levels of urgency through the different parameters. It was shown that pitch was contributing more than the other parameters to urgency judgements even though all of the parameters were at theoretically equal levels of urgency.

In Chapter Six the possible determinants of urgency were examined. An attempt was made to see what factors might determine whether or not acoustic parameter manipulations resulted in changes in perceived urgency. In particular, the perceived duration literature was examined to see if the concepts of time and urgency were related. The information processing and Reiss Jones accounts of time perception were identified as being likely to offer explanations for changes in perceived urgency with reference to perceived duration.

Chapter Seven assessed the information processing and Reiss Jones accounts of time perception and investigated the relationship between perceived urgency and

perceived duration. Sound parameters were varied in ways known to increase perceived urgency and the effects of these variations upon perceived duration was noted. It was shown that increases in speed resulted in increases in perceived duration and increases in resolution resulted in increases in perceived duration. There was a non-significant trend for increases in inharmonicity to decrease perceived duration. Variations in pitch had no effect upon perceived duration.

Of the parameter changes shown to affect perceived duration, one (increases in speed) was associated with increases in perceived urgency, and one (increases in resolution) was associated with decreases in perceived urgency. An attempt was made to account for these findings in terms of differences in the nature of the two parameters. It was suggested the speed was an innate parameter and resolution was cultural. This suggestion was investigated in the final experiment by testing musical and non-musical subjects, and assuming that the former group would be more attuned to cues in the cultural parameter than the latter group. Increases in speed resulted in increases in perceived duration and perceived urgency for both groups. Increases in resolution resulted in decreases in perceived urgency but had no significant effect upon perceived duration for both groups. There were no differences in response between the musical and non musical subjects to either the speed of the resolution stimuli.

The empirical work in this research program has demonstrated some potentially important findings. It has been shown that psychophysical techniques, in particular free modulus magnitude estimation and cross modality matching, provide a useful means of measuring perceived urgency. It has also been demonstrated that increases in speed, pitch, repetition units and inharmonicity result in increases in the perceived urgency of a sound; while increases in resolution result in decreases in perceived urgency. The effects of variations in some of the parameters upon urgency have been quantified, so that it is possible to say how much a parameter has to be varied to communicate a unit change in urgency. When the parameters are set at equal levels of urgency, pitch influences urgency judgements the most. It was suggested that pitch was more salient or discriminable than the other acoustic parameters. It has been shown that increases in speed and resolution result in increases in perceived duration. The findings for speed and resolution parameters have been shown to apply to both musical and non musical subjects.

The theoretical and practical implications of these empirical findings are discussed below.

8.2. Theoretical Implications

The research presented in this thesis has contributed theoretically to the areas of psychophysics, time perception and to the concept of urgency itself. These contributions are discussed below.

matching function.

In Chapters 2-5 it was demonstrated that psychophysical techniques could be used to quantify the effects of different sound parameters upon perceived urgency, thus extending to scope of such techniques to cover a previously un-scaled parameter. The exponent of Steven's Power Law was used to describe the relationship between changes in the different sound parameters and changes in perceived urgency. The power law was used to construct stimuli that conveyed the same level of urgency through different sound parameters. In Chapter 5 it was shown that there was a high correlation between the predicted and obtained urgency values of stimuli constructed in this way. The Power Law is thus a suitable model for predicting the perceived urgency of different sounds.

Scaling the perceived urgency of sounds highlighted some issues surrounding Stevens' proposed division of continua into Prothetic and Metathetic. In usual scaling tasks an objective parameter is manipulated and subjects are required to judge the subjective value of that parameter, for example length is manipulated and subjects judge how long the stimulus appears. In urgency scaling however an acoustic parameter, for example speed, is manipulated and subjects are required to judge not the speed , but the urgency of the stimulus. This situation is termed 'second order' scaling.

In first order scaling of the usual kind, placing a stimulus into one of Stevens Prothetic or Metathetic categories is a simple matter of deciding whether the continuum is qualitative or quantitative. For example, length is quantitative and thus probably a Prothetic continuum. In second order scaling it is not clear whether it is the parameter that is manipulated (speed) or the parameter that is judged (urgency), that determines the continua type of the stimulus. Urgency itself is probably a Metathetic continua, whereas speed is Prothetic. Stevens claims that ratio scaling techniques cannot be successfully used to scale Metathetic continua. The successful application of ratio scaling techniques to the stimuli employed in Chapters 3-5 suggests that it is the manipulated, not the judged parameter that determines the continua of the stimulus. Most of the manipulated parameters employed, speed,

repetition units and the number of inharmonic components are Prothetic and thus the successful application of ratio scaling techniques would be predicted if it is the manipulated parameter that determines continua type. It has thus been demonstrated that second order scaling can be employed. Although it is not entirely clear whether it is the manipulated or judged parameter that determines the continua type of the stimulus in Stevens terms, the sucessful application of ratio scaling scaling techniques suggests that it is the manipulated parameter.

The research undertaken in this thesis has provided findings that allow urgency itself to be considered in theoretical terms. It has been demonstrated that various acoustic parameters can be manipulated to affect urgency, and that the pitch changes have especially salient effects upon urgency judgements. As discussed in Chapter 6, there are various possible causes of these effects, such as learned associations, predictability, arousal, evolution and perceived time. The possibility that perceived time might contribute to the effect of increases in perceived urgency was investigated in detail.

In Chapter 7 it was demonstrated that some parameter changes that increase perceived urgency increase perceived duration (for example, speed) whilst other parameter changes that decrease perceived urgency increase perceived duration (for example, resolution). An attempt was made to reconcile these findings with the current theories of perceived duration to reveal the mechanism for a relationship between perceived urgency and perceived duration. The information processing models of perceived duration could not account for the finding that increases in stimulus resolution resulted in increases in perceived duration. The Reiss Jones model of perceived duration was supported as the superior description. According to that model, stimuli varying in stimulus speed are non coherent and therefore encourage analytic attending and thus mnemonic strategies are employed to make time judgements. This results in increases in information in the stimulus, increases in speed, being perceived as longer. Stimuli that vary in resolution are more meaningful and thus encourage future- orientated attending. This results in unresolved stimuli, that violate the expected stimulus ending, being judged shorter relative to the resolved stimuli. It was therefore possible to reconcile the findings that some parameters increasd perceived duration while other parameters decreased perceived duration within the framework of the Reiss Jones account of time perception.

increased

The finding that some of the acoustic changes that cause changes in perceived urgency are associated with changes in perceived duration was supported by the Reiss Jones account of perceived duration. Thus perceived duration may be part of what determines whether changes in acoustic parameters affect perceived urgency, perceived duration may be part of what makes a stimulus urgent or non urgent. It was suggested that the mechanism for this effect depends on the nature of the parameter that is manipulated to convey urgency. Increases in parameters such as speed increase perceived duration in the manner suggested by Reiss Jones for non coherent stimuli that encourage analytic attending. The increases in perceived duration contribute to the effect of increasing urgency by making the stimulus appear longer (and it is known that longer stimuli are perceived as being more urgent). Increases in parameters such as resolution increase perceived duration in the manner described by Reiss Jones for coherent stimuli that encourage future-oriented attending. In this case the increases in perceived duration correspond with decreases in perceived urgency because resolved stimuli appear longer that unresolved stimuli and yet the information content in both is the same. Because the resolved stimulus appears longer without altering the information content, it becomes apparently slower because the information within it occurrs over an apparently longer period of time, and slower stimuli are preceived as being less urgent.

Thus it appears that perceived duration contributes to the effects of perceived urgency and that the nature of the effect depends upon the type of stimulus.
There was some limited support for the idea that the acoustic parameters used to communicate urgency were either cultural or innate, corresponding approximately to Reiss Jones coherent and non coherent parameters.

To summarise, several theoretical developments have arisen from this thesis. The scope of psychophysical techniques has been broadened to include a new continua, and a variation of the traditional scaling procedure, second order scaling, has been introduced. Stevens Power Law has been demonstrated as a predictive device for assessing the perceived urgency of different sounds. Changes in perceived duration has been identified as a factor that contributes to changes in urgency, with the mechanism for the effect depending on the acoustic parameter employed. The information processing accounts of time perception have been shown to be inadequate and the Reiss Jones account has been supported.

8.3. Practical Implications.

The work reported in this thesis provides information on perceived urgency to enable the more ergonomic construction of auditory warnings. This has far-reaching practical implications for auditory warnings designers and users, these are discussed below.

It has been shown that different acoustic parameters can be used to communicate perceived urgency and these include speed, pitch, repetition, inharmonicity and resolution. Objective methods of quantifying changes in these parameters were revealed so that the changes could be related to changes in perceived urgency by psychophysical techniques.

The use of psychophysical techniques in this research has demonstrated that it is possible to measure and quantify the effect of different sound parameters upon perceived urgency. The exponents for each parameter show at a glance the strength of the effect of changes in one parameter relative to another upon perceived urgency. Thus it was revealed that speed was the most economical parameter to use to communicate perceived urgency, and that inharmonicity had little practical usefulness because huge changes in inharmonicity are required to produce a unit change in urgency.

This information will enable existing warnings to be modified to make them more ergonomic. The temporal and spectral qualities of different warnings can be analysed and this research used to adjust the different parameters so that they communicate the required levels of urgency. Designers of new warnings will also be able to use this research to see the relative strengths of different parameters for communicating perceived urgency. When warnings are constructed by varying only one parameter, the exponents will help the designer to choose which parameter to manipulate, usually the most economic. This is because, in most cases, it will be possible to communicate more levels of urgency through the most economic parameter because smaller changes are required in that parameter to communicate set increases or decreases in urgency while keeping the parameter values within an ergonomic range.

By using Stevens Power Law as a predictive model for perceived urgency, it will be possible to know which levels of a parameter are more urgent than which other

levels and by how much. Thus it will be possible to prioritise warnings and also to employ 'urgency mapping' (Montahan and Tansley 1989). These advantages also apply between parameters, it will be possible to communicate an equal amount of urgency through warnings that vary in different parameters. This means that warnings signalling different conditions can be kept distinct by varying them along different parameters, but that they can signal the same range of urgency levels.

In Chapter 5 it was demonstrated that when all the parameter values were at equal urgency levels, the pitch parameter had relatively more of an effect upon urgency judgements. Although this may imply that pitch is a particulary salient parameter to use for communicating urgency, it is recommended that it is employed with caution. Pitch is a Metathetic continua and thus it is not possible to be entirely confident about the accuracy of its scaling until more is known about the importance of Stevens continua divisions. Moreover, the salience of pitch changes means that it would be a useful parameter for distinguishing one group of warnings from another. It is recommended that the parameters that have been scaled with more confidence should be used to convey urgency.

In sum, the use of psychophysical techniques to measure perceived urgency has meant that the effects of different acoustic parameters upon perceived urgency can be quantified. Warnings designers can use this information to evaluate the relative and individual contribution of different parameters to perceived urgency when new warnings are designed or existing ones modified. This will help to implement many of the recommended improvements to warning design, such as urgency mapping and warning prioritisation. Such information is already being incorporated into the draft BSI standard for hospital warnings and will be used by the BSI committee for warnings in noisy environments. This in turn may effect a more ergonomic and efficient relationship between the warning systems and the operator.

8.4. Future Areas of Research

There are several avenues of research that could be explored so that the conclusions of the present thesis could be extended and clarified. These are discussed below.

In order to extend the practical applications of the present research it is important to know what the effect is of communicating urgency through several parameters simultaneously in the same stimulus. The effects upon perceived urgency of covarying several parameters simultaneously are not yet known. Psychophysical

techniques could be used to show whether the effects of the different parameters upon urgency are additive or whether some parameters dominate judgement.

For practical purposes it is also important to recognise that urgency is only one of several possible messages that each parameter can be used to communicate. As auditory signals gain increasing acceptance in the workplace there is an increasing demand not just for alarms, but also for 'trendsons'. Trendsons are trend monitoring sounds that are used to convey auditory feedback to the operator about the state of a system of sytems. Trendsons are currently being designed for helicopters that monitor for example, rotor over-speed and rotor under-speed (Loxley 1991). In trendson design auditory parameters are used not only to communicate levels of urgency but also to monitor the state of various systems. Thus in a trendson several parameters are employed simultaneously, as an example, one may communicate urgency, another may communicate that a system is slowing down or speeding up and another may communicate that a system level is dropping or rising. It is important to employ the parameter that best conveys each message.

Experiments could be conducted in which subjects are required to rate the effectiveness of a set of descriptors for each parameter. This might show, for example, that speed is an effective communicator of urgency and pitch is an effective communicator of something dropping. Each parameter could thus be employed to its best advantage when the trendson was constructed.

Potential problems arise with trendsons when the information that the parameters are conveying is contradictory. For example, a drop in oil pressure might be communicated by decreasing pitch to show that the level was decreasing. Dropping oil pressure however is an increasingly urgent condition and decreasing pitch gives the impression of decreasing urgency. Such problems would be avoided if the effectiveness of each parameter for communicating a particular message were known. The dropping could be communicated by a parameter that was very poor at communicating 'urgency' and good at communicating 'dropping'. Thus the probability of the warning being interpreted as communicating a drop in urgency would be minimised and the probability of the dropping being interpreted maximised.

In order to extend the theoretical implications of the present research it is important firstly to test the proposed mechanism by which an increase in perceived duration is associated with a decrease in perceived urgency for resolution stimuli, whereas an increase in perceived duration was associated with an increase in

perceived urgency for speed stimuli. In the latter case, the increase in perceived duration was thought to result in an increase in perceived urgency by virtue of the fact that longer stimuli are perceived as being more urgent. It was suggested that in the former instance the increase in perceived duration for the resolved stimuli meant that they were perceived as being slower because the information within them appeared to be presented in a longer period of time. The decrease in the perceived speed of the stimuli was thought to result in the decrease in perceived duration. This proposal can be tested by asking subjects directly whether resolved stimuli are perceived as being any slower than unresolved stimuli.

The idea that there are fundamental or innate and cultural parameters that can communicate urgency received only limited support in the present research. Before this idea is rejected it is suggested that an experiment is run comparing the responses of subjects from two different cultures to the resolved and speed stimuli. It is thought that the musical and non musical subjects employed in Experiment 11 may not have been different enough culturally to demonstrate a difference in their responding.

8.5. Summary

This thesis has investigated the perceived urgency of auditory warnings. Psychophysical techniques have been used to quantify the effects of different sound parameters upon perceived urgency. It has been shown that increases in speed, pitch, repetition and inharmonicity result in increases in perceived urgency, and that increases in resolution result in decreases in perceived urgency. The possibility the perceived duration is part of what makes a stimulus urgent or non urgent has been investigated. It was shown that for some parameters increases in perceived duration corresponded to increases in perceived urgency but of others increases in perceived duration corresponded to decreases in perceived urgency. The Reiss Jones account of perceived duration could account for these findings. Further research is required on the proposed mechanism by which changes in perceived duration effect perceived urgency and on the idea that there are two types of acoustic parameter, fundamental or innate and cultural.

REFERENCES.

Achamamba, B. (1988). Perceptions of time and cognitive efficiency. Journal of Psychological Researches, 32(1-2), 63-67.

Aiba, T. & Stevens, S. (1964). Relation of brightness to duration and luminence under light and dark adaption. Vision Research, 4, 391-401. Cited by Marks, L. (1974).

Allen, L. (1976). Is there a constant minimum under light and dark adaption? Quarterly Journal of Experimental Psychology, 28, 71-76.

Allen, L. (1977). The time order error in judgements of duration, Canadian Journal of Psychology, 31, 24-31.

Allen, L. (1979). The perception of time. Perceptiohn and Psychophysics. 26(5), 340-354.

Allen, L. (1983). Magnitude estimation of Temporal intervals. Perception and Psychophysics, 33(1), 29-42.

Allen, L. & Kristofferson, A. (1974). Psychophysical theories of Duration discrimination. Perception and Psychophysics, 16(1), 26-34.

Allen, L. & Kristofferson, A. & Wiess, E. (1971). Duration discrimination of breif light flashes. Perception and Psychophysics, 22, 686-670. Cited by Fraisse, P. (1978).

Anderson, A. (1971). Test of adaption level theory as an explaination of recency effects in psychophysical integration. Journal of Experimental Psychology, 87(1), 57-63.

Atkinson, W. (1982). A general equation for sensory magnitude. Perception and Psychophysics, 31(1), 26-40.

Atteneave, F. (1962). Perception and related areas. In S. Koch (Ed.), Psychology. A Study of a Science. McGraw : Hill.

Avant, L. Lyman, P. & Antes, J. (1975). Effects of stimulus familiarity. Perception and Psychophysics, 17(3), 253-262.

Banks, W. (1974). A new psychophysical ratio scaling technique, random production. Bulletin of the British Psychometric Society, 1(4), 273-275.

Berg, M. (1979). Temporal duration as a function of information processing. Perceptual and Motor Skills, 49, 988-990.

Bergland, B., Ekman, U. & frankenhauser, M. (1969). Influence of auditory stimulus intensity upon apparent duration. Scandinavian Journal of Psychology, 10, 21-26.

Birnbaum, M. (1974). Using contextural effects to derive psychophysical scales. Perception and Psychophysics, 15(1), 89-96.

Bjokman, M. & Holmkvist, O. (1960). Time order errors in the construction of a subjective time scale. Scandinavian Journal of Psychology, 1, 7-13.

Block, R. (1990). Models of Psychologiacl Time. In R. Block (Ed), Cognitive Models of Psychological Time, Lawrence Erlbaum Assoc., New Jersey.

Block, R. (1985). Contextual coding in memory. In A. Michon & J, Jackson (Eds), Time Mind and Behavior, Heidelberg, Springer-Verlag. Cited By Blotz, M. (1989).

Block, R. (1978). Remembered Duration, Effects of event and sequency complexity. Memory and Cognition, 6(3), 320-326.

Block, R. (1978). Memory and the experience of duration in retrospect. Memory and Cognition, 2, 153-160. Cited by Block, R. (1990).

Block, R. & Reed, M. (1978). Remembered Duration. Journal of Experimental Psychology : Human Learning and Memory, 4(6), 656-665.

Block, R., Reed, M. & George, E. (1980). A watched pot sometimes boils. Acta Psychologia, 46, 81-94.

Bobko, D., Thompson, J. & Schifman, H. (1977). The Perception of breif temporal intervals. Perception, 6, 703-709.

Bobko, D., Schifman, H. & Castino, R. (1977a). Contextual effects in duration experience. American Journal of Psychology, 90(4), 577-586.

Bock, M., Lazarus, H. & Hoge, H. (1983). Effects of noise on the efficiency of danger signals. In G. Rossi. (Ed), Noise as A Public Health Problem - Proceedings of Forth International Congress, 1, 517-521.

Boltz, M. (1989). Time judgements of musical endings. Perception and Psychophysics, 46(5), 409-418.

Brown, S. (1985). Time perception and attention. Perception and Psychophysics, 38(2), 115-124.

Brown, S. & Stubbs, D. (1988). The psychophysics of prospective and retrospective timing. Perception, 17, 297-310.

Brown, D. & Hichcock, L. (1965). Time estimation. Perceptual and Motor Skills, 21, 727-734. Cited by Fraisse, P.(1978).

Bringer, W. (1988). Perceived duration as a function of pitch. Perceptual and Motor Skills, 67, 301-302.

Buffardi, L. (1971). Factors effecting the filled duration illusion. Perception and Psychophysics, 10(4), 292-294.

Butler, D. (1979). Tonal structure versus function. Music Perception, 2, 6-24.

Cahoon, R. (1969). Physiological arousal and time estimation. Perceptual and Motor Skills, 28, 259-268.

Cantor, N. & Thomas, E. (1977). Control of attention in the processing of temporal and spatial information. Journal of Experimental Psychology : Human Perception and Performance, 3(2), 243-250.

Church, R. (1984). Properties of the internal clock. In J. Gibbon & L. Allen (Eds), Timing and Time Perception, Annals of New York Academy of Science, 556.

Clynes, M. & Walker, J. (1983). Neurobiologic functions of rhythm time and pulse in music. In M. Clynes (Ed), Music, Mind and Brain, Pleunem Press, New York and London.

Clynes, M. & Walker, J. (1986). Music as times measure. Music Perception, 4(1), 85-120.

Cooper, J. & Couvillion, L. (1983). Accidental breathing system disconnections. Interim Report to the Food and Drug Administration. Cited by Kerr, J. (1985).

Cross, D. (1973). Sequential dependancies and regression in psychophysical judgement. Perception and Psychophysics, 14(3), 547-552.

Curtis, D. (1970). Magnitude estimates and category judgements of brightness and brightness intervals. Journal of Experimental Psychology, 83(2), 201-208.

Curtis, D., Atteneave, F. & Harrington, T. (1968). A test of a two stage model of magnitude estimation. Perception and Psychophysics, 3, 25-31.

Curton, E. & Lordahl, D. (1974). Effects of attentional focus and arousal on time estimation. Journal of Experimental Psychology, 103(5), 861-867.

Dawson, W. & Brinker, R. (1971). Validation of ratio scales of oppinion by multi-modality matching. Perception and Psychophysics, 9(5), 413-417.

Delay, E. & Richardson, M. (1981). Time estimation in humans. Perceptual and Motor Skills, 53, 747-750.

Deutsch, D. (1980). The processing of structured and unstructured musical sequences. Perception and Psychophysics, 28, 381-389.

Dilollo, V. (1964). Contrast effects in judgements of lifted weights. Journal of Experimental Psychology, 68(4), 383-387.

Divenyi, P. & Sacks, R. (1987). Discrimination of time intervals bounded by tone bursts. Perception and Psychophysics, 24(5), 429-436.

Dohering, D. (1961). Accuracy and consistency of time estimation. American Journal of Psychology, 74, 27-35. Cited by Schiffman, H, et al (1977).

Duda, P. (1975). Tests on the psychophysical meaning of the power law. Journal of Experimental Psychology : Human perception and Performance, 104(2), 188-194.

Edworthy, J., Loxley, S., Geelhoed, E. & Dennis, I. (1988). An experimental investigation into the effects of spectral, temporal and musical parameters on the perceived urgency of auditory warnings. Report on MOD Project No. SLS42B/205. RAE Farnborough.

Edworthy, J., Loxley, S. & Hellier, E. (1989). A preliminary investigation into the use of sound parameters to convey helicopter trend information. Report on MOD Project No. SLS42B/568. RAE Farnborough.

Edworthy, J., Loxley, S. & Dennis, I. (1991). Improving auditory warning design : relationship between warning sound parameters and perceived urgency. Human Factors, 33(2), 205-231.

Efron, R. (1964). Temporal perception aphasia. Brain, 86, 403-424. Cited by Fraisse, P. (1984).

Eisler, H. (1962). Empirical test of a model relating magnitude and category scales. Scandinavian Journal of Psychology, 3, 88-96.

Eisler, H. (1965). Psychophysics in general and the general psychophysical differential equation in particular. Scandinavian Journal of Psychology, 6, 85-102.

Eisler, H. (1974). The derivation of Stevens Power Law. In H. Moskowitz, B. Scharf, and J. Stevens (Eds.), Sensation and Measurement. Dordnecht : Holland. Cited by Allan, L. (1983).

Eisler, H. (1975). Subjective duration and Psychophysics, Psychological Review, 82, 429-450. Cited by Fraisse, P. (1984).

Eisler, H. (1976). Subjective duration and psychophysics. Psychological Review, 82(6), 429-450.

Eisler, H. (1981). Applicability of the parallel clock model to duration discrimination. Perception and Psychophysics, 29(3), 225-233.

Eisler, H. (1984). Subjective duration in rats. In J. Gibbon & L. Allen (Eds.), Timing and Time Perception, p.43, Annals of the new York Academy of Science.

Eisler, H. & Montgomery, H. (1974). On theoretical and realisable ideal conditions in psychophysics. Perception and Psychophysics, 16(1), 157-166.

Ekman, G. (1958). Two generalized ratio scaling methods. Journal of Psychology, 45(1), 287-295.

Ekman, G. (1961). Some aspects of psychophysical research. In W. Rosenblith, Sensory Communication, M.I.T. Press.

Ekman, G., Berglund, B. & Berglund, U. (1966). Loudness as a function of duration of auditory stimulation. Scandinavian Journal of Psychology, 7, 201-210. Cited by Berglund, B. et al (1969).

Ekman, G., Hosman,B., Lindman, R., Jundberg, L. & Akesson, C. (1968). Interindividual differences in scaling performance. Perceptual and Motor Skills, 26, 815-823.

Ekman, G., Frankenhauser, H & Berglund, U. (1969). Apparent duration as a function of stimulation. Perceptual and Motor Skills, 34, 421-422. Cited by Berglund, B. et al (1969).

Engen, T. (1971). Scaling methods. In J. Kling, & L. Riggs (Eds.), Experimental Psychology. Methuen and Co. Ltd : London.

Erickson, C. & Hake, H. (1957). Anchor effects in absolute judgements. Journal of Experimental Psychology, 53, 132-138. Cited by Seigel, W. (1972).

Fagot, R. (1963). On the psychophysical law and estimation procedures in psychophysical scaling. Psychometrika, 28(2), 145-160.

Fagot, R. & Porkorny, R. (1989). Bias effects on magnitude and ratio estimation power function exponents. Perception and Psychophysics, 45(3), 221-230.

Fechner, G. (1860). Elemente der Psychophysik. Vols. 1&2. Breikopf & Hartel. Cited by Stevens, S. (1966).

Federal Aviation Administration. (1977). Aircraft Alerting Systems. US Department of Transportation. FAA Research and Development Service. Washington DC. Cited by Kantowitz, B. et al (1988).

Fidell, S. & Teffeteller, S. (1981). Scaling the annoyance of intrusive sounds. Sound and Vibration Research, 78(2), 291-298.

Fischer, R. & Fraser, J. (1962). Biological time. In J. Fraser, (Ed), The Voices of Time, New York, Brasillier. Cited by Reiss-Jones, M. (1976).

Foley, H., Cross, D., Foley, M. & Reeder, R. (1983). Stimulus range, number of categories and the virtual exponent. Perception and Psychophysics, 34(6), 505-512.

Foley, H., Cross, D. & O'Reilly, J. (1990). Prevasiveness and magnitude of context effects ; Evidence for relativity of absolute magnitude judgements. Perception and Psychophysics, 48(6), 551-558.

Fraisse, P. (1963). The Psychology of Time. New York, Harper and Row.

Fraisse, P. (1975). Psychology of Time. Greenwood Press.

Fraisse, P. (1978). Time and Rhythm perception. In C. Carterette & M. Freidman (Eds.), Handbook of Perception 8, p.203. Academic Press.

Fraisse, P. (1981). Cognition of time in human activity. In G. Ydewalle &W. Lens (Eds.), Cognition in Human Learning and Motivation, Lauwrence Erlbaum Associates.

Fraisse, P. (1984). Perception and estimation of time. Annual Review of Psychology, 35(1), 1-36.

Fraisse, P. & Orsini, F. (1962). Etude experimentale de conduites temporales. Anee Psychol. 58, 1-6. Cited by Fraisse, P. (1984).

Frankenhauser, M. (1959). Estimation of Time. Almquist & Wiksells. Cited by Reiss-Jones, M. (1976).

Freed, D. & Martens, D. (1986). Deriving Psychophysical Relations fro Timbre. Proceedings of the International Computer-Music Association, 393-405.

Fucci, D., Harris, D., Petrinso, L. & MacMath, E. (1987). Effects of psychophysical scaling method, body test site and skin contractor surface area on vibrotactile magnitude functions. Perceptual and Motor Skills, 43(3), 1127-1138.

Fucci, D., Harris, D., Petrosino, L. & Randolf-Tyler, E. (1988). Auditory Psychophysical Scaling Exposure effects. Perceptual and Motor Skills, 66, 643-648.

Gabrielsson, A. (1974). Performance of rhythm patterns. Scandinavian Journal of Psychology, 15, 63-72.

Gage, F. (1934). An experimental investigation into the measurability of auditory stimuli. Proceedings of the Royal Society (London), 116B, 103.

Galanter, E. & Messick, S. (1961). The relation between category and magnitude scales of loudness. Psychological Review, 68(6), 363-372.

Garner, W. (1968). Perception and learning of temporal patterns. Quarterly Journal of Experimental Psychology, 20, 97-109

Garner, W. & Hake, J. (1951). The ammount of information in absolute judgements. Journal of Experimental Psychology, 46, 446-459.

Gavin, H. (1959). Contribution a l'etude de la perception des burees breves. Journal de la Psychologie Normale et Pathologique, 56, 455-468. Cited by Fraisse, P. (1978).

Gescheider, G. (1990). Stimulus context and absolute magnitude estimation. Proceedings of 6th Annual Meeting of International Society for Psychophysics, 43-48.

Getty, D. (1975). discrimination of short temporal intervals. Perception and Psychophysics, 18, 1-8. Cited by Reiss-Jones et al (1989).

Goldner, J., Reuder, M., Riba, B. & Jarmon, D. (1971). Neutral v's ego orienting instructions. Perception and Psychophysics, 9, 84-88.

Goldstein, J. (1973). An optimal processor theory for the central formation of the pitch of complex tones. Journal of Acoustical Society of America, 54, 1494-1516.

Goldstone, S., Lhamon, W. & Sechzer, J. (1978). Light intensity and judged duration. Bulletin of the Psychonomic Society, 12, 83-84. Cited by Fraisse, P. (1984).

Gomez, L. & Robertson, L. (1979). The filled duration illusion. Perception and Psychophysics, 25(5), 432-438.

Goude, G. (1962). On Fundamental Measurement in Psychology. Stockholm : Almquist and Wicksell. Cited by Eisler, H. (1965).

Graham, C. & Ratoosh, P. (1962). Notes on some interactions of sensory psychological perceptions and behavior. In S. Koch (Ed.), 1962.

Green, D. & Luce, R. (1974). Variability in magnitude estimates. Perception and Psychophysics, 15(2), 291-300.

Gravetter, F. & Lockhead, G. (1973). Criterial range as a frame of reference for stimulus judgement. Psychological Review, 80(3), 203-216.

Guilford, J. & Dingman, H. (1954). A validation study of ratio judgement methods. American Journal of Psychology, 67, 395-410. Cited by Marks, L. (1974)

Halpern, L., Blake, R. & Hillenbrand, J. (1986). The psychoacoustics of a chilling sound. Perception and Psychophysics, 39(3), 77-80.

Harton, J. (1939). The influence of difficulty on the estimation of time. Journal of Experimental Psychology, 23, 428-432. Cited by Brown, S. (1985).

Haverland, E. (1979). Magnitude estimation, a new way of measuring subjective test variables. US Army Tropic Test Center Report, No.790601, 1 - 58.

Hawkes, G., Joy, R. & Evans, W. (1962). Autonomic effects on estimates of time. Journal of Psychology, 53, 183-191.

Hawkins, M. & Telford, J. (1979). Effects of interest and relatedness on perceived duration. Bulletin of Psychonomic Society, 8(4), 301-302

Hellier, E. (1988). An investigation into the effects of repetition rate, speed and length on the perceived urgency of auditory warnings. Unpublished undergraduate dissertation.

Hellier, E. & Edworthy, J. (1989). Quantifying the Perceived Urgency of Auditory Warnings. Canadian Acoustics, 17(4), 3-11.

Hellman, R. & Zwislocki, J. (1961). Some factors effecting the estimation of loudness. Journal of the Acoustical Society of America, 33, 687-694. Cited by Warren, R. (1981).

Hellman, R. & Zwislocki, J. (1968). Loudness determination at low frequencies. Journal of the Acoustical Society of America, 43, 60-64. Cited by Marks, L. (1974).

Hellman, R. & Zwicker, E. (1990). Magnitude scaling : a meaningful method for measuring loudness and annoyance? Proceedings of 6th Annual Meeting of International Society for Psychophysics, 123-128.

Helson, H. (1964). Adaption Level Theory. New York, Harper Row.

Hicks, R. & Brundige, R. (1974). Judgement of Temporal Duration. Acta Psychologia, 38, 447-453.

Hicks, R., Miller, G. & Kinsbourne, M. (1976). Prospective and retrospective judgements of time. American Journal of Psychology, 89, 719-730. Cited by Block, R. (1990).

Hicks, R., Miller, G. Gaes, G. & Bierman, K. (1977). Concurrent Processing Demands and experience of time in Passing. American Journal of Psychology, 90(3), 431-446.

Higashiyama, A. & Tashiro, T. (1989). Magnitude Estimates for electrical pulses, Evidence for two neural mechanisms. Perception and Psychophysics, 45(6), 537-549.

Hoagland, H. (1934). Physiologic control of judgements of duration. Journal of General Psychology, 9, 267-287. Cited by Block, R. (1990).

Hogan, W. (1975). Time perception and stimulus preference. Journal of Personality and Social Psychology, 31, 32-35. Cited by Reiss-Jones et al (1989).

Hogan, W. (1978). A theoretical reconciliation of competing views of time perception. American Journal of Psychology, 91(3), 417-428.

Hoge, H. & Shick, A. (1988). Are there invariants of sound interpretation? In B. Berglund et al (Eds), Noise as a Public Health Problem, 2, 253-258.

Holland, M. & Lockhead, G. (1968). Sequential effects in absolute judgements of loudness. Perception and Psychophysics, 3, 409-414.

Hollingworth, L. (1910). The central tendency of judgement. Journal of Philosophical Psychology, 7, 461-469. Cited by Bobko, D. et al (1977).

Horing, A. (1864). Verschue uber das ... Tubinge. Cited by Fraisse, P. (1978).

Houtgast, T. (1976). Subharmonic pitches of a pure tone at low S/N ratio. Journal of Acoustic Society of America, 60, 405-409.

Howell, D. (1982). Statistical methods for use in Psychology. PWS Publishing.

James, S. & James, M. (1989). Effects of warning format and keyboard layout on reaction times to auditory warnings. Proceedings of Institute of Acoustics, 11(5), 25-29.

Jesteadt, W., Luce, R. & Green, D. (1977). Sequential effects in loudness judgements. Journal of Experimental Psychology : Human Perception and Performance, 3, 92-104.

Johnson, D. & Mulally, C. (1969). Correlation and regression model for category judgements. Psychological Review, 76(2), 209-215.

Jones, F. & Marcus, M. (1961). The subject effect in judgements of subjective magnitude. Journal of Experimental Psychology, 61(1), 40-44.

Jones, F. & Woskow, M. (1966). Some effects of context on the slope in magnitude estimation. Journal of Experimental Psychology, 71(2), 177-180.

Jones, B. & Lehur-Huang, Y. (1982). Space time dependencies in psychophysical judgement of extent and duration. Psychological Bulletin, 91, 128-142.

Kantowitz, B., Kantowitz, S. & Sorkin, B. (1988). Liklihod alarm displays. Human Factors, 30(4), 445-459.

Kerr, J. (1985). Warning Devices. British Journal of Anaesthetics, 57, 696-676.

Kerr, J. & Hayes, B. (1983). An alarming situation in the intensive therapy unit. Intensive Care Medicine, 9, 103-109.

King, M. & Lockhead, G. (1980). Response scales and sequential effects in judgement. Perception and Psychophysics, 30(6), 599-603.

Kowal, K. (1987). Apparent duration and numerosity as a function of melodic familiarity. Perception and Psychophysics, 42(2), 122-131.

Kreuger, L. (1984). Perceived numerosity. Perception and Psychophysics, 35(6), 536-542.

Kristofferson, A. (1984). Quantal and deterministic timing in human duration discrimination. in Gibbon, J. & Allen, L. (Eds.). Timing and Time Perception, p.3. Annals of the New York Academy of Science.

Kristofferson, A. & Allen, L. (1973). Sucessiveness and Duration discrimination. In S. Koch (Ed.), Attention and Performance, 738-50, New York, Academic. Cited By Fraisse, P. (1984).

Kruup, K. (1971). Influence of method on time judgements. Australian Journal of Psychology, 5, 34-40. cited by Hicks, R. et al (1977).

Kuwano, S. & Namba, S. (1990). Continuous judgement of loudness and annoyance. Proceedings of 6th Annual Meeting of International Society for Psychophysics, 129-139.

Landis, C. (1925). Studies of emotional reactions. American Journal of Physiology, 74, 188-206. cited by Cahoon, R. (1969).

Lane, H., Catania, A. & Stevens, S. (1961). Voice level : Autophonic scale, perceived lounness and the effects of the side tone noise. Journal of the Acoustic Society of America, 33(1), 160-167.

Latour, P. (1967). Evidence for internal clocks in the human operator. Acta Psychologia, 27, 341-348.

Lazarus, H. & Hoge, H. (1986). Industrial safety - acoustic signals for danger in factories. Applied Ergonomics, 17(1), 41-46.

Lockhead, G. & King, M. (1983). A memory model of sequential effects in scaling tasks. Journal of Experimental Psychology : Human Perception and performance, 9(3), 461-473.

Longuet-Higgins, C. & Lee, C. (1982). Perception of musical rhythms. Perception, 11, 115-128.

Louge, A. (1976). Individual differences in the magnitude estimation of loudness. Perception and Psychophysics, 19(3), 279-280.

Lower, M., Weeler, M., Patterson, R., Edworthy, J., Shailer, M., Milroy, R., Rood, G. & Chillery, J. (1986). Design and production of auditory warnings for helicopters 1 : the sea king. Institute of Sound and Vibration Research, Report No. AC598.

Loxley, S. (1991). The Design and Evaluation of Trend Monitoring Sounds. Submitted as M.Phil Thesis, Polytechnic South West, Plymouth, England.

Luce, D. (1985). Time Perception. In J. Gibbon & L. Allen (Eds.).

Luce, R. & Green, D. (1974). Response Ratio Hypothesis for magnitude estimation. Journal of Mathematical Psychology, 11, 1-14.

Macmillan, N., Maschetto, C., Bailostolky, F. & Engel, L. (1974). Size judgement the presence of the standard increases the exponent of the power law. Perception and Psychophysics, 16(2), 340-346.

Marks, L. (1968). Stimulus range, number of categories and the form of the category scale. American Journal of Psychology, 81(4), 467-479.

Marks, L. (1974). Sensory Processes : The New Psychophysics. New York : Academic Press.

Marks, L. (1988). Magnitude estimation and sensory matching. Perception and Psychophysics, 43(6), 511-525.

Marks, L., Szczesiul, R. & Ohlott, P. (1986). On the cross modal perception of intensity. Journal of Experimental Psychology : Human Perception and Performance, 12, 517-534. Cited by Marks, L. (1988).

Martin, J. (1972). Rhythmic versus serial structure in speech and other behaviors. Psychological Review, 79, 487-509. Cited by Reiss-Jones, M. (1976).

Mashour, M. (1965). Note on the validity of the Power law. Scandinavian Journal of Psychology, 6, 220-224.

Mashour, M. & Hosman, J. (1968). On the new psychophysical law : a validation study. Perception and Psychophysics, 3, 367-375.

Masin, S. (1983). Experimental check on Stevens explaination of partition paradox. Perception and Psychophysics, 34(3), 294-296.

Massaro, D. (1960). Temporal course of perceived auditory duration. Perception and Psychophysics, 14, 233-235. Cited by Fraisse, P. (1984).

Matsuda, F. (1965). Development of time estimation. Japaneese Journal of Psychology, 36, 285-294.

Mcbride, R. (1983a). Taste intensity and the case of exponents greater than one. Australian Journal of Psychology, 35, 175-184.

McBride, R. (1983b). Jnd/category scale convergence in taste. Perception and Psychophysics, 34(1), 77.

Mcbride, R. (1983c). Psychophysics, could fechners'assumption be correct? Australian Journal of Psychology, 35, 85-88.

Mcbride, R. (1986). Cautionary note on log-log plots. Australian Journal of Psychology, 38(2), 177-178.

Mclain, L. (1983). Interval estimation. Perception and Psychophysics, 34, 185-189. Cited by Brown, S. (1985).

McIntyre, J. (1986). Ergonomics : Anaesthetists use of auditory alarms in the operating room. Journal of Clinical Monitoring and Computing, 2, 47-55.

Mclonchie, R. & Ruchiman, J. (1971). Human time estimation. Perceptual and Motor Skills, 45, 854-857. Cited by Brown, S. (1985).

McRobert, H., Bryan, M . & Tempest, W. (1965). Magnitude estimation of loudness. Journal of Sound Vibration, 2, 391. Cited by Stevens, S. (1971).

Meisselman, H. (1980). Variables Affecting the Psychophysical Function for Taste. In H. Van der Starre (Ed.), Olfaction and Taste 8, London IRL Press.

Mellers, B. (1983). Evidence against absolute scaling. Perception and Psychophysics, 33(6), 523-526.

Mellers, B. & Birnbaum, M. (1982). The loci of contextural effects in absolute judgement. Journal of Experimental Psychology : Human Perception and Performance, 8(4), 582-601.

Meredith, L. & Wilsoncroft, W. (1989). Time Perception. Perceptual and Motor Skills, 68, 373-374.

Metfessel, M. (1947). A proposal for quantitative reporting of comparative judgements. Journal of Psychology, 24, 229-235. Cited by Marks, L. (1974).

Michon, J. (1967). Studies in subjective duration 2. Acta Psychologia, 24, 205-219.

Miller, G. & Heise, G. (1950). The trill threshold. Journal of the Acoustical Society of America, 22, 654-665. Cited by Reiss-Jones, M. (1976).

Milwiski, A. & Iaccino, J. (1982). Strategies in cross modality matching. Perception and Psychophysics, 31(3), 273-275.

Miller, G., Hicks, R. & Willette, M. (1978). Effects of concurrent verbal rehersal and temporal set on temporal judgements. Acta Psychologia, 34(2), 40-47.

Mo, S. (1971). Judgement of temporal duration as a function of numerosity. Psychonomic Science, 24, 71-72.

Momtahan, C. (1990). Mapping of psychoacoustic parameters to the perceived urgency of auditory warning signals. MA Thesis, submitted to Carleton University, Ottowa, Ontario.

Momtahan, C. & Tansley, B. (1989). An ergonomic analysis of alarm signals in the operating and recovery rooms. Paper presented at the

Annual Conference of the Canadian Acoustical Association. Halifax, Canada.

Montgomery, H. (1975). Direct estimation, the effect of methodological factors on scale type. Scandinavian Journal of Psychology, 16, 19-29.

Montgomery, H. & Eisler, H. (1974). Is an equal interval scale an equal discriminability scale? Perception and Psychophysics, 15(3), 441-448.

Moore, B., Glasberg, B. & Peters, R. (1982). Relative dominance of individual partials in determining the pitch of complex tones. Journal of the Acoustical Society of America, 77(5), 1853-1860.

Mulligan, R. & Schiffman, H. (1979). Temporal experience as a function of memory organisation. Bulletin of the Psychonomic Society, 14, 417-420.
Cited by Brown, S. et al (1988).

Mundy, C. & Castle, A. (1953). Electrical responses in the brain in relation to behavior. Journal of Psychology, 44, 318-329. Cited by Cahoon, R. (1969).

Nakajima, Y., Nishimura, S. & Teranish, R. (1988). Ratio judgements of empty durations with numeric scales. Perception, 17, 93-118.

Neter, J. & Wasserman, W. (1985). Applied Linear Statistical Models. (2nd ed.). Richard Irwin Inc., Illinois.

Nisly, S. & Wasserman, G. (1989). Intensity dependance of perceived duration. Psychological Bulletin, 106(3), 483-496.

O'Carroll, T. (1986). Survey of alarms in the intensive therapy unit. Anaesthesia, 41, 742-744.

Ornstein, R. (1969). On the Experience of Time. Penguin Books.

Parducci, A. (1974). Contextural effects, a range frequency analysis. In E. Carette, & M. Freidman (Eds.), Handbook of Perception (Vol2), New York : Academic Press. Cited by Mellers, B. & Birnbaum, M. (1982).

Parker, M. & Schubert, M. (1984). Experience of a critically ill patient experienceing theraputic paralysis in the ICU. Intensive Care Medicine, 12, 69-74. Cited by Kerr, J. (1985).

Patterson, R. (1982). Guidelines for auditory warning systems in civil aircraft. CAA Paper 82017.

Patterson, R. (1985). Design of auditory warnings for aircraft, industry and hospitals. In Brown, I. et al (Eds), Ergonomics International -85, Proceedings of the Ninth Congress of the International Ergonomics Association. 163-165.

Patterson, R., Edworthy, J., Shailer, M.,Lower, M. & Weeler, M.(1986). Alarm sounds for medical eqiupment in intensive care units and operating theatres. Institute of Sound and Vibration Research, Report No. AC527.

Peleg, M . & Campanella, O. (1988). On the mathematical form of psychophysical relationships, with special focus on mechanical properties of soild objects. Perception and Psychophysics, 44(5), 451-455.

Petrosino, L., Fucci, D., Harris, D. & Randolf-Tyler, E. (1988). Lingual vibrotactile-auditory magnitude estimations and cross modality matching, comparison of supra threshold responses in men and women. Perceptual and motor skills, 67, 291-300.

Plateau, J. (1872). Sur la mesure des sensations physiques et sur la loi qui lie l'intensitie de ces sensations a l'intensitiede la cause excitante. Bull. Acad. Roy. Belg., 13, 376-388. Cited by Warren, R. & Poulton, E. (1962).

Plomp, R. & Steencken, H. (1969). Effect of phase on the timbre of complex tones. Journal of the Acoustical Societyof America, 46(2), 409-421.

Pollach, I. (1964). Neutralization of stimulus bias in auditory rating scales. Journal of the Acoustic Society of America, 36, 1272-1276.

Poulton, E. (1968). New psychophysics, six models for magnitude estimation. Psychological Bulletin, 69(1), 1-19.

Poulton, E. (1979). Models for judging biases in sensory magnitude. Psychological Bulletin, 86, 777-803.

Poulton, E. (1982). Biases in quantitative judgements. Applied Ergonomics, 13, 31-42.

Poulton, E. & Simmonds, D. (1963). Value of standard and very first variable in judgements of reflectance of grays. Journal of Experimental Psychology, 65(3), 297-304.

Poulton, E. & Freedman, P. (1966). Unwanted asymmetrical transfer effects with balanced experimental designs. Psychological Bulletin, 66(1), 1-8.

Poulton, E., Edwards, R. & Fowler, T. (1980). Eliminating subjective bias in judging loudness. Perception and Psychophysics, 27(2), 93-103.

Poulton, E. & Simmonds, D. (1985). Subjective zeros, subjectively equal stimulus spacing and contraction biases in the very first judgements of lightness. Perception and Psychophysics, 37, 420-428.

Povel, D. (1979). Temporal structure of performed music. Acta Psychologia, 49, 309-320. Cited by Reiss-Jones, M. (1990).

Povel, D. (1981). Internal representation of simple temporal paterns. Journal of Experimental Psychology : Human Perception and Performance, 7(1), 3-18.

Povel, D. & Essens, P. (1985). Perception of temporal patterns. Music Perception, 2(4), 411-440.

Poynter, W. (1983). Duration judgement and the segmentation of experience. Memory and Cognition, 11, 77-82.

Poynter, W. (1979). Human time perception and memory processes. Masters thesis , Arizona State University. Cited by Poynter, W. et al (1983).

Poynter, W. & Homa, D. (1983). Duration judgement and the experience of change. Perception and Psychophysics, 33, 548-560.

Pradham, P. & Hoffman, P. (1963). Effect of spacing and range of stimuli on magnitude estimation judgements. Journal of Experimental Psychology, 66(6), 533-541. Cited by Warren, R. (1981).

Price-Williams, D. (1954). The kappa effect. Nature, 363-364. Cited by Matsuda, M. (1979).

Rai, S. (1973). A comparison of time estimation of music, noise and light filled intervals. Indian Journal of Psychology, 48(4), 37-43.

Raab, D. & Osman, E. (1962). Effect of the temporal overlap. Journal of Optical Society of America, 52, 1174. Cited by Marks, L. (1974).

Refinetti, R. (1989). Magnitude estimation of warmth, inter and intra subject variability. Perception and Psychophysics, 46(1), 81-84.

Reiss-Jones, M. (1976). Time our lost dimension. Psychological Review, 83, 323-335.

Reiss-Jones, M., Kidd, G. & Wetzel. (1981). Evidence for rhythmic attention. Journal of Experimental Psychology : Human Perception and Performance, 7, 1059-1073.

Reiss-Jones, M., Boltz, M. & Kidd, G. (1982). Controlled attending as a function of melodoc and temporal context. Perception and Psychophysics, 32, 211-218.

Reiss-Jones, M. & Boltz, M. (1989). Dynamic attending and responses to time. Psychological Review, 96, 459-491.

Richardson, S. & Ross, J. (1930). Loudness and Telephone Current. Journal of General Psychology, 3, 288-306. Cited by Stevens, S. (1956).

Robinson, G. (1976). Biasing Power Law exponents by magnitude estimation instructions. Perception and Psychophysics, 19(1), 80-84.

Rood, G. (1989). Auditory warnings for fixed and rotary winged aircraft. Proceedings of Institute of Acoustics, 11(5), 59-71.

Ross, J. & Dilollo, V. (1970). A consistent failure of the PowerLaw for lifted weight. Perception and Psychophysics, 8(5), 289-290.

Ross, J. and Dilollo, V. (1971). Judgement and response in magnitude estimation. Psychological Review,78(6), 515-527.

Schab, F. & Crowder, R. (1989). Accuracy of temporal coding. Memory and Cognition, 17(4), 384-397.

Schaffer, L. (1984). Rhythm and timing in skill. Psychological Review, 89, 109-122. Cited by Reiss-Jones, M. (1990).

Schiffman, H. Bobko, D. & Thompson, J. (1977). Role of stimulus context on apparent duration. Bulletin of the Psychonomic Society, 10(6), 484-486.

Schiffman, H. & Bobko, D. (1974). Effects of stimulus complexity. Journal of Experimental Psychology, 103(1), 156-159.

Schmuckler, M. (1989). Expectation in Music. Music Perception, 7(2), 109-150.

Schneider, B. (1981). Is the sensory code truly inaccessible? Behavior and Brain Sciences, 4, 175.

Schneider, B. & Bisset, R. (1988). Ratio and difference judgements of area, length and volume. Are there really two classes of sensory continua? Journal of Experimental Psychology : Human Perception and Performance, 14(3), 503-512.

Schreiber, J. & Schreiber, J. (1989). Structured alarm systems for the operating room. Journal of Clinical Monitoring, 5(3), 201-207.

Sebel, A. & Wilsoncroft, W. (1989). Auditory and visual differences in time estimation. Perceptual and Motor Skills, 57, 295-300.

Siegel, S. (1972). Memory effects in absolute judgement. Journal of Experimental Psychology, 94(2), 121-131.

Stanford, L., McIntyre, J. & Hogan, T. (1985). Audible alarm signals for anaestesia monitoring equipment. International Journal of Clinical Monitoring and Computing, 1, 251-256.

Stanford, L., McIntyre, J., Nelson, H. & Hogan, T. (1988). Affective resopnses to commercial and exerimental auditory alarm signals for anaeathesia delivery.. nternational Journal of Clinical Monitoring and Computing, 5, 111-118.

Stevens, J. (1975). Psychophysics. New York, Wiley.

Stevens, J. & Tulving, E. (1957). Estimation of loudness by a group of untrained observers. American Journal of Psychology, 70, 600-605. Cited by Marks, L. (1974).

Stevens, J. & Mack, T. (1959). Scales of apparent force. Journal of Experimental Psychology, 58, 405-413.

Stevens, J. & Marks, L. (1980). Cross modality matching functions generated by magnitude estimation. Perception and Psychophysics, 27(5), 379-389.

Stevens, S. (1955). The Measurement of loudness. Journal of the Acoustic Society of America, 27, 815-829.

Stevens, S. (1956). The direct estimation of sensory magnitudes : loudness. American Journal of Psychology, 69,1-25.

Stevens, S. (1957). On the Psychophysical Law. Psychological Review, 64, 153-181. Cited by Stevens, S. (1971).

Stevens, S. (1958). Problems and Methods in psychophysics. Psychological Bulletin, 55(4), 177-196.

Stevens, S. (1959). On the validity of the loudness scale. Journal of the Acoustic Society of America, 31(1), 995-1003.

Stevens, S. (1960). The psychophysics of sensory function. American Scientist, 48, 226-253.

Stevens, S. (1961). Towards a resolution of the Fechner-Thurstone legacy.Psychometrika, 26(1), 35-47.

Stevens, S. (1966). On the operation known as judgement. American Scientist, 54(1), 385-401.

Stevens, S. (1966a). Matching functions between loudness and ten other continua. Perception and Psychophysics, 1, 5-9.

Stevens, S. (1969). On predicting exponents for cross modality matches. Perception and Psychophysics, 6(4), 251-257.

Stevens, S. (1970). Neural events and the psychophysical law. Science, 70, 1043-1050.

Stevens, S. (1971). Issues in Psychophysical measurement. Psychological Review, 78(5), 426-450.

Stevens, S. & Poulton, E. (1956). The estimation of loudness on unpracticed observers. Journal of Experimental Psychology, 51, 71-78.

Stevens, S. & Galanter, E. (1957). Ratio scales and category scales for a dozen perceptual continua. Journal of Experimental Psychology, 54, 377-411.

Stevens, S. & Guiaro, M. (1962). Loudness, reciprocality and partition scales. Journal of the Acoustic Society of America, 34, 1406-1471.

Stevens, S. & Greenbaum, H. (1966). Regression effect in psychophysical judgement. Perception and Psychophysics, 1, 439-446.

Szeto, A., Valerio, N. & Novak, R. (1991). Audible pedestrian traffic signals. Journal of Rehabilitation Research and Development, 28(2), 57-64.

Teghtsoonian, R. (1973). Range effects in psychophysical scaling and Stevens law. American Journal of Psychology, 86(1), 3-27.

Teghtsoonian, R. & Teghtsoonian, M. (1970). Scaling apparent distance. Psychonomic Science, 21, 215-216. Cited by Warren, R. (1981).

Teghtsoonian ,R. & Teghtsoonian, M. (1971). The apparent length of ratios and diameters. American Journal of Psychology, 84, 437-438. Cited by Teghtsoonian, R. (1973).

Teghtsoonian, R. & Teghtsoonian, M. (1978). Range and regression effects in magnitude scaling. Perception and Psychophysics, 24(4), 305-314.

Teghtsoonian, R. & Teghtsoonian, M. (1986). Scaling loudness over short ranges- a reply to Poulton. Perception and Psychophysics, 39(1),73-75.

Thomas, E. & Brown, I. (1974). Time perception and the filled duration illusion, Perception and Psychophysics, 16(3), 449-458.

Thomas, E. & Weaver, W. (1975). Cognitive processing and time perception. Perception and Psychophysics, 17(4), 363-367.

Thomas, E. & Cantor, N. (1978). Interdependance in the processing of temporal and non temporal information. In J. Requin (Ed), Attention and Performance, Hillsdale, NewYork : Erlbaum. Cited by Reiss-Jones, M. (1989).

Thorning, A. & Ablett, R. (1985). Auditory warning systems in commercial transport aircraft. In Brown, I. et al (Eds), Ergonomics International -85, Proceedings of the Ninth Congress of the International Ergonomics Association. 166-168.

Todd, N. (1985). A model of expressive timing in tonal music. Music Perception, 3(1), 33-57

Tomlinson, M. (1987). Alarm standard aims to strike the right note. New Scientist, Jan, 40.

Torgerson, W. (1961). Distances and Ratios in Psychophysical Scaling. Acta Psychologia, 19(1), 201-205.

Treisman, M. (1963). Temporal discrimination and the indifference interval. Psychological Monographs, 77(13), No.576.

Underwood, G. (1975). Attention and the perception of duration. Perception, 4, 291-296.

Underwood, G. & Swain, G. (1973). Selectivity of attention and the perception of duration. Perception, 2, 101-105.

VanOrden, K., Sturr, T. & Taub, H. (1987). Context effects in brightness estimation. Perception and Psychophysics, 41(5), 416-418.

Von Sturmer, G. (1966). Stimulus variation and sequential judgements of duration. Quarterly Journal of Psychology, 354-357.

Vroon, P. (1970). Effects of presented and processed information on experince of duration. Acta Psychologia, 34, 115-121.

Wagner, M. & Baird, J. (1981). Quantitative analysis of sequential effects. Perception and Psychophysics, 29(4), 359-364.

Walker, J. & Scott, K. (1981). Auditory visual conflicts in perceived duration. Journal of Experimental Psychology : Human Perception and Performance, 7(6), 1327-1339.

Ward, L. (1972). Category judgements of loudness in the absesce of an experimenter induced identification function. Journal of experimental Psychology, 94(2), 179-184.

Ward, L. (1973). Repeated magnitude estimates with a variable standard.Perception and Psychophysics, 13(2), 193-200.

Ward, L. (1985). Mixed modality psychophysical scaling. Perception and Psychophysics, 38(6), 517-522.

Ward, L. (1986). Mixed modality psychophysical scaling. Perception and Psychophysics, 39(6), 407-417.

Ward, L. (1987). Remembrance of sounds past, memory and psychophysical scaling. Journal of Experimental Psychology, Human Perception and Performance, 13(2), 216-227.

Ward, L. & Lockhead, G. (1971). Response system processes in absolute judgement. Perception and Psychophysics, 9(1), 73-78.

Warren, R. (1970). Elimination of biases in Loudness judgements. Journal of the Acoustic Society of America, 48, 1397-1403.

Warren, R. (1981). The measurement of sensory intensity. Behavoir and Brain Sciences, 4, 175-223.

Warren, R. & Poulton, E. (1962). Ratio and Partition Judgements. American Journal of Psychology, 75(1), 109.

Weiss, D. (1981). The impossible dream of Fechner and Stevens. Perception, 10, 407-417.

Werboff, J. (1957). Relationship between eeg activity and behavior. Dissertation Abstracts, 17, 2325. Cited by Cahoon, R. (1969).

White, C. (1963). Temporal numeroisity and the psychological unit of duration. Psychological Monographs, 77, 12, No.575.

Wilson, K. & Stelmak, R. (1982). Magnitude estimation and auditory brainstem evoked responses. Perception and Psychophysics, 31(6), 561-565.

Woodrow, H. (1934). The effect of practice upon time order error. Psychological Review, 42, 127-152. cited by Bobko, D. et al (1977)

Woodrow, H. (1966). Time Perception. In S. Stevens (Ed), Handbook of Experimental Psychology p.1224. J.Wiley &Sons.

Zakay, D., Nittzan, D. & Glickshon, J. (1983). Influence of task difficulty and external tempo on subjective time estimation. Perception and Psychophysics, 34(5), 451-456.

Zelkind, I. (1973). Factors in Time Estimation. Journal of General Psychology, 88, 295-301.

Zener, K. & Graffon, N. (1962). Perceptual experience. In S. Koch (Ed.), Psychology. A Study of a Science. McGraw : Hill.

Zwislocki, J. & Goodman, D. (1980). Absolute scaling of sensory magnitudes. Perception and Psychophysics, 28(1), 28-38.

Appendices.

APPENDIX 1A :

SPECIFICATION OF CED SOFTWARE FOR THE CONSTRUCTION OF AUDITORY WARNINGS.

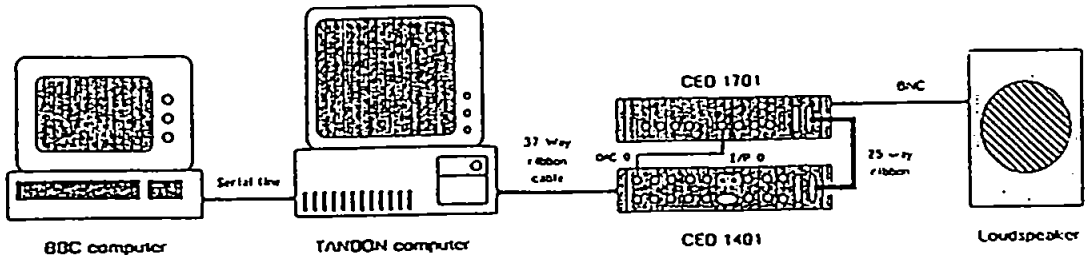
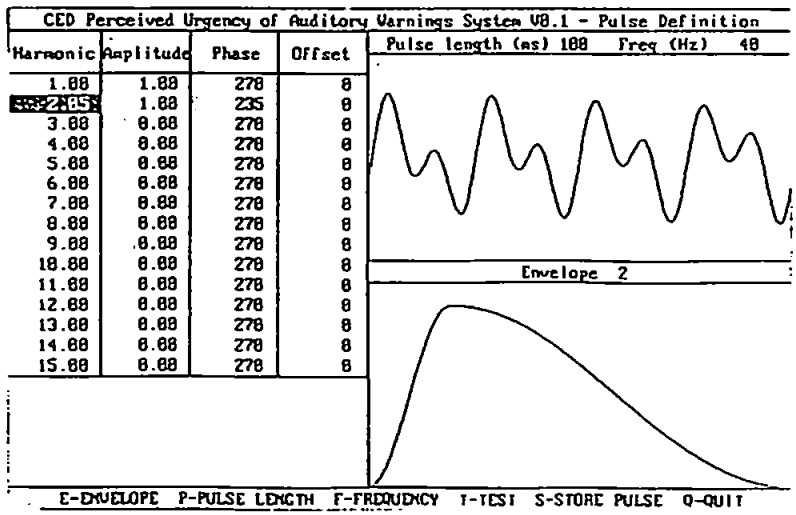


Fig. 1.1 : Set-up of the Tandon Hardware

Pulse definition

The first section is used to define a pulse type in terms of harmonics, length and amplitudes. The pulse type will be used as a template to create a burst in the second part. The display shows the harmonic information in tabular form on the left hand side of the screen and graphs of the signal and its envelope on the right.



Harmonic	Amplitude	Phase	Offset
1.00	1.00	270	0
2.00	1.00	235	0
3.00	0.00	270	0
4.00	0.00	270	0
5.00	0.00	270	0
6.00	0.00	270	0
7.00	0.00	270	0
8.00	0.00	270	0
9.00	0.00	270	0
10.00	0.00	270	0
11.00	0.00	270	0
12.00	0.00	270	0
13.00	0.00	270	0
14.00	0.00	270	0
15.00	0.00	270	0

Fig. 1.2 : Pulse Definition Screen Display

Harmonic content

The pulse can contain up to 15 harmonics from the 0th to the 99th. Note that the frequency of the pulse is is not set at this stage in the system and so if a high frequency is set later, high harmonics may be over the limit of 5000 Hz. and so will not be heard. It is up to the operator to check for this.

To alter a value, highlight the position of the value in the table with the cursor keys and press RETURN. Next type in the new value and
press RETURN. Four parameters may be defined for each harmonic. These are:

1. The number- of the harmonics.
Defines the frequency of the harmonic relative to the frequency of the pulse set later. E.g. 5.0 will set the value of the harmonic to five times the fundamental frequency. Valid values are from O.O to 99.

2. The amplitude (weighting) of the harmonics.
This can be in the range O to 100 but is defined relative to the total amplitude. That is, the fractional amplitude of the harmonic equals the harmonic amplitlude divided by the sum of the harmonic amplitudes. The overall amplitude is defined in the second part of the system.

3. The phase of the harmonic.
The phase has units of degrees, from O to 360. A phase of 270 degrees gives a sine wave which is useful when the harmonic is offset into the pulse since the amplitude still starts at O and so avoids a clicking sound when the harmonic sounds.

4. The offset into the pulse.
This is the time (in milliseconds) from the start of the pulse at which the harmonic will start. Normally this will be 0, starting when the pulse starts but can range between -pulselength and +pulse length. A positive value will wait the prescribed number of milliseconds before the harmonic is added to the pulse, a negative value means that the harmonic will end this Length of time before the pulse ends.

Pulse length

The P key allows the user to enter a value for the total pulse length milliseconds. This can range from the maximum offset so far defined to 409 ms.

Envelopes

The E key will switch between the three envelopes allowed, the one selected is shown on the graph in the bottom right hand corner. The envelope is multiplied by the signal values before the pulse is played and so confines the wave to the limits of the envelope. The details of the three envelopes are:

a) Standard envelope-
(i) Cosine gate from O to 1 in 20 ms
(ii) 1 until 20 ms from end of pulse
(iii) Cosine gate from 1 to O in final 20 ms

b) Slow offset envelope-
(i) Cosine gate from O to 1 in first 20 ms
(ii) Cosine gate dropping to O at end of pulse

c) Slow onset envelope:
(i) Cosine gate from O to 1 20 ms from end of pulse
(ii) Cosine gate from 1 to O in last 20 ms

Burst definition

This section defines a burst in terms of a pulse type already defined in the first part. A burst is a series of up to 12 pulses of one pulse type with varying gaps between them. The frequency and attenuation of the pulses can also vary. The frequency at whicn the burst will be played out 10 kHz. and the maximum available memory for a burst is 25000 integers which makes the maximum length of a burst 2.5 seconds. The information on the pulses in the burst is again in tabular form on the screen. Below this is space for a graph of the burst.

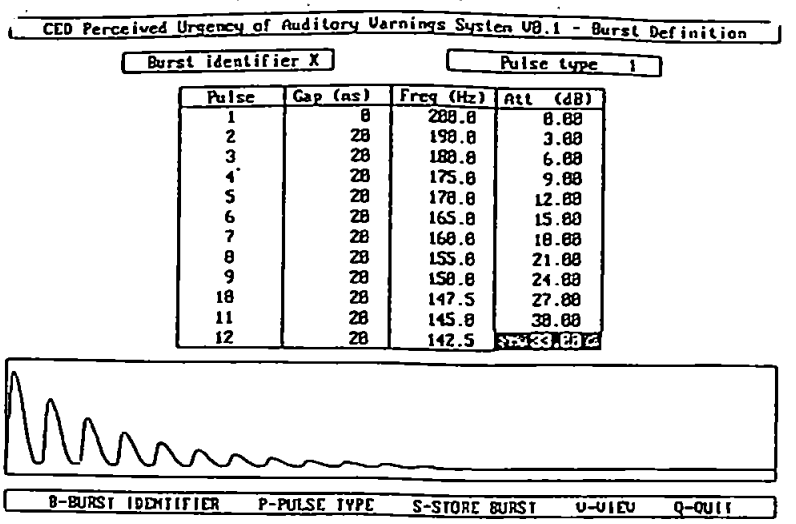


Fig. 1.3. Screen Display for Burst Definition.

The pulse type

When the program is first entered, and whenever P is pressed, you may type in the number of the pulse type to use for the burst. The computer will try to find the file for this pulse as it needs the information enclosed. If the file is not on the hard disk the computer will not allow most of the options in this part to be used (e.g. to view the burst). It is therefore important to determine the pulse type before running this section of the system.

The burst identifier

This is the letter or digit which will be used to produce the burst in the playback section.

The pulses

The cursor can be moved round thc table with the arrow keys as before except that you cannot move down the table until the amplitude of thc current pulse is specified (for example you can't define the second pulse before you have defincd the first). To alter a value, position the cursor on the valec and press RETURN, then enter the new value. The three parameters to specify are:

1. The time gap (in milliseconds) between the end of the previous pulse (or the start of the burst in thc case of the first pulse) and the start of this pulse. A negative gap will merge the two pulses but beware that this may cause the total amplitude to overflow, producing a very strange sound.

2. The fundamental frequency of the pulse to the nearest 2.5Hz. between 0 and 5000. High harmonics may be cut off if the frequency is too high.

3. The attenuation of the pulse in decibels. This is a value between O dB abd 72dB.

APPENDIX 3A

COMPONENTS OF SPEED STIMULI.

For an explanation of pulse and burst terminology see Appendix 1A.

Pulse Used in Burst Creation.

Pulse 1

200 ms. long

Fundamental frequency 300Hz

Regular harmonics

Phase 270

Offset 0

Amplitude 1

Regular amplitude envelope.

Practice Stimuli.

Burst A.

Pulse 1

Attenuation : 5,0,0..

Six pulses

5 ms. gap between pulses

Pitch = 300, 200 Hz etc.

Burst B.

Pulse 1

Attenuation : 5,0,0..

Eight pulses

5 ms. gap between pulses

Pitch = 300, 200 Hz etc.

Burst C.

Pulse 1

Attenuation : 5,0,0..

Twelve pulses

5 ms. gap between pulses

Pitch = 300, 200 Hz etc.

Experimental Stimuli.

Burst 1 (predicted least urgent).
Pulse 1
Attenuation : 5,0,0..
Four pulses
Pitch = 300, 200 Hz.
475 ms. gap between pulses
Total length = 2225 ms.

Burst 7.
Pulse 1
Attenuation : 5,0,0..
Five pulses
Pitch 300, 200 Hz.
300 ms. gap between pulses
Total length = 220ms.
2

Burst 6.
Pulse 1
Attenuation : 5,0,0..
Eight pulses
Pitch = 300, 200Hz.
118 ms. gap between pulses
Total length = 2426 ms.

Burst 5.
Pulse 1
Attenuation : 5,0,0..
Ten pulses
Pitch = 300,200Hz.
50 ms. gap between pulses
Total length = 2450ms.

Burst 2.
Pulse 1
Attenuation : 5,0,0..
Six pulses
Pitch = 300, 200 Hz.
237 ms. gap between pulses
Total length = 2385 ms.

Burst 3.
Pulse 1
Attenuation : 5,0,0..
Nine pulses
Pitch = 300, 200 Hz.
59 ms. gap between pulses
Total length = 2272 ms.

Burst 4 (predicted most urgent).
Pulse 1
Attenuation : 5,0,0..
Twelve pulses
Pitch = 300, 200 Hz.
9 ms. gap between pulses
Total length = 2499 ms.

APPENDIX 3B

t TEST COMPARING THE A AND B COMPONENTS OF EACH SUBJECTS REGRESSION EQUATIONS AGAINST MEANS OF 0 AND 1(EX.1).

Subject.	A components (intercept) of regression equation when each subjects first two judgements were regressed against against the last six judgements.	B components (slope) of regression equation when each subjects first two judgements were regressed against the last six judgements.
1	-5.3	0.968
2	-0.468	1.08
3	-0.990	1.07
4	-0.969	1.07
5	4.91	1.03
6	-40.6	1.71
7	1.56	0.983
8	-7.64	1.05
9	0.480	0.705
10	-5.27	1.49
11	-1.07	1.03
12	-0.62	0.923

t test of Mu=0 vs Mu not equal to 0.
t=-1.33, Df=11, p=0.21

t Test of Mu=1vs Mu not equal to1.
t=1.22, Df = 11, p=0.25

APPENDIX 3C

SUBJECTIVE AND OBJECTIVE STIMULUS VALUES FITTED TO STEVENS POWER FUNCTION BY THE METHOD OF LEAST SQUARES.

X	Y		
LOG STIM. MAG	LOG MEAN JUDGEMENT	XSQUARED	XY
0.568	0.750	0.322	0.426
0.698	0.946	0.487	0.660
0.757	0.995	0.573	0.753
0.895	1.168	0.801	1.045
0.984	1.340	0.968	1.318
1	1.358	1	1.358
1.077	1.488	1.159	1.602
SUM	SUM	SUM	SUM
5.979	8.045	5.31	7.164

$$A = \frac{N(\text{Sum } XY)-(\text{Sum } X)(\text{Sum } Y)}{N(\text{Sum } X^2)-(\text{Sum } X)} = \frac{7(7.164)-(5.979*8.045)}{7*5.312-5.979}$$

A (exponent) = 1.430

$$B = \frac{(\text{Sum } X^2)(\text{Sum } Y)-(\text{Sum } X)(\text{Sum } XY)}{N(\text{Sum } X^2)-(\text{Sum } X)}$$

$$= \frac{5.312*8.045-5.979*7.164}{7*5.312-5.979} = 0.682, \quad (\text{Antilog}=1.170)$$

B (constant) = 1.170

The Stevens Power Function for perceived urgency is:

Perceived urgency = 1.170*pulse rate^1.43

APPENDIX 3D

t TEST COMPARING THE A AND B COMPONENTS OF EACH SUBJECTS REGRESSION EQUATIONS AGAINST MEANS OF 0 AND 1(EX.2).

Subject.	A components (intercept) of regression equation when each subjects first two two judgements were regressed against against last six	B components (slope) of regression equation when each subjects first two judgements were regressed against last six
1	-4.81	1.16
2	-0.79	0.990
3	-1.01	0.950
4	9.39	0.797
5	5.29	0.979
6	-20.6	1.39
7	-2.94	1.04
8	6.67	0.799
9	-1.49	0.975
10	-1.72	1.07
11	-0.03	1.04
12	-7.39	1.16

t test of Mu=0 vs Mu not equal to 0.
t=-0.73, Df=11, p=0.48

t Test of Mu=1vs Mu not equal to1.
t=0.63, Df = 11, p=0.54

APPENDIX 3E

SUBJECTIVE AND OBJECTIVE STIMULUS VALUES FITTED TO STEVENS POWER FUNCTION BY THE METHOD OF LEAST SQUARES.

X	Y		
LOG STIM. MAG	LOG MEAN JUDGEMENT	XSQUARED	XY
0.568	1.248	0.322	0.708
0.698	1.326	0.487	0.925
0.757	1.365	0.573	1.033
0.895	1.647	0.801	1.474
0.984	1.754	0.968	1.752
1	1.774	1	1.774
1.077	1.875	1.159	2.019
SUM	SUM	SUM	SUM
5.979	10.989	5.31	9.661

$$A = \frac{N(\text{Sum } XY)-(\text{Sum } X)(\text{Sum } Y)}{N(\text{Sum } X^2)-(\text{Sum } X)} = \frac{7(9.661)-(5.979*10.989)}{7*5.312-5.979}$$

A (exponent) = 1.34

$$B = \frac{(\text{Sum } X^2)(\text{Sum } Y)-(\text{Sum } X)(\text{Sum } XY)}{N(\text{Sum } X^2)-(\text{Sum } X)}$$

$$= \frac{5.312*10.98-5.979*9.661}{7*5.312-5.979} = 0.475, \quad (\text{Antilog}=2.985)$$

B (constant) = 2.985

The Stevens Power Function for perceived urgency is:

Perceived urgency = 2.985*pulse rate^1.34

APPENDIX 3F

t TEST COMPARING THE A AND B COMPONENTS OF EACH SUBJECTS REGRESSION EQUATIONS AGAINST MEANS OF 0 AND 1(EX.3).

Subject.	A components (intercept) of regression equation when each subjects first two judgements were regressed against against last six	B components (slope) of regression equation when each subjects first two judgements were regressed against last six
1	-12.9	1.30
2	-2.99	0.973
3	-9.29	1.36
4	- 9	1.28
5	8.88	0.927
6	4.25	0.984
7	-0.36	0.731
8	-0.08	1.33
9	- 8	1.13
10	-6.91	1.01
11	1.75	1.09
12	-97.86	2.25

t test of Mu=0 vs Mu not equal to 0.
t=-1.36, Df=11, p=0.20

t Test of Mu=1vs Mu not equal to1.
t=1.79, Df = 11, p=0.10

APPENDIX 3G

SUBJECTIVE AND OBJECTIVE STIMULUS VALUES FITTED TO STEVENS POWER FUNCTION BY THE METHOD OF LEAST SQUARES.

X LOG STIM. MAG	Y LOG MEAN JUDGEMENT	XSQUARED	XY
0.568	0.787	0.322	0.447
0.698	1.211	0.487	0.845
0.757	1.334	0.573	1.009
0.895	1.729	0.801	1.547
0.984	1.792	0.968	1.763
1	1.807	1	1.807
1.077	1.941	1.159	2.09
SUM	SUM	SUM	SUM
5.979	10.601	5.31	9.510

$$A = \frac{N(\text{Sum } XY)-(\text{Sum } X)(\text{Sum } Y)}{N(\text{Sum } X^2)-(\text{Sum } X)} = \frac{7(9.510)-(5.979*10.601)}{7*5.312-5.979}$$

A (exponent) = 2.22

$$B = \frac{(\text{Sum } X^2)(\text{Sum } Y)-(\text{Sum } X)(\text{Sum } XY)}{N(\text{Sum } X^2)-(\text{Sum } X)}$$

$$= \frac{5.312*10.60-5.979*9.510}{7*5.312-5.979} = -0.328, \quad (\text{Antilog}=0.469)$$

B (constant) = 0.469

The Stevens Power Function for perceived urgency is:

Perceived urgency =

0.469*pulse rate^2.22

APPENDIX 3H

SUBJECTIVE AND OBJECTIVE STIMULUS VALUES FITTED TO THE MATCHING FUNCTION BY THE METHOD OF LEAST SQUARES.

X	Y		
LOG STIM. MAG	LOG MEAN JUDGEMENT	XSQUARED	XY
0.568	1.675	0.322	0.951
0.698	1.798	0.487	1.255
0.757	1.836	0.573	1.389
0.895	2.076	0.801	1.858
0.984	2.188	0.968	2.152
1	2.211	1	2.211
1.077	2.350	1.159	2.530
SUM	SUM	SUM	SUM
5.979	14.13	5.31	12.34

$$A = \frac{N(\text{Sum } XY)-(\text{Sum } X)(\text{Sum } Y)}{N(\text{Sum } X^2)-(\text{Sum } X)} = \frac{7(912.34)-(5.979*14.13)}{7*5.312-5.979}$$

A (exponent) = 1.35

$$B = \frac{(\text{Sum } X^2)(\text{Sum } Y)-(\text{Sum } X)(\text{Sum } XY)}{N(\text{Sum } X^2)-(\text{Sum } X)}$$

$$= \frac{5.31*14.13-5.979*12.34}{7*5.312-5.979} = 0.867, \quad (\text{Antilog}=7.367)$$

B (constant) = 0.867

The matching function between line length and perceived urgency has a slope of:

1.35

APPENDIX 4A

STIMULI VARYING IN PITCH, USED IN EXPERIMENT 5.

Harmonics over 4000Hz were given an amplitude of 0 so all harmonics were played at the same level, regardless of their fundamental.

Pulse Used in Practice Burst Creation.

PULSE 1
200 ms long
Fundamental frequency 300Hz, regular harmonics.
Phase 270
Offset 0
Amplitude , regular envelope.

Practice Stimuli.

BURST A.
Pulse 1
Attenuation:5,0,0...
Six pulses
5 ms gap between pulses
Pitch = 300,200Hz etc

BURST B.
Pulse 1
Attenuation:5,0,0...
Twelve pulses
5 ms gap between pulses
Pitch = 300,200Hz etc

BURST C.
PULSE 1
Attenuation : 5,0,0...
Eight pulses
5 ms gap between pulses
Pitch = 300,200 Hz etc

Pulses Used In Experimental Burst Creation.

PULSE 2.
200 ms long
Fundamental frequency 210Hz
Regular harmonics
Phase 270
Offset
Amplitude 1
Regular amplitude envelope

PULSE 3.
200 ms long
Fundamental frequency 250Hz
Regular harmonics
Phase 270
Offset 0
Amplitude 1
Regular amplitude envelope

PULSE 4.
200 ms long
Fundamental frequency 260Hz
Regular harmonics
Phase 270
Offset 0
Amplitude 1
Regular amplitude envelope

PULSE 5.
200 ms long
Fundamental frequency 320Hz
Regular harmonics
Phase 270
Offset 0
Amplitude 1 (except last three harmonics, amplitude 0)
Regular amplitude envelope

PULSE 6.
200 ms long
Fundamental frequency 440Hz
Regular harmonics
Phase 270
Offset 0
Amplitude 1 (last 5 = 0)
Regular amplitude envelope

PULSE 7.
200 ms long
Fundamental frequency 500Hz
Regular harmonics
Phase 270
Offset 0
Amplitude 1 (last 7 = 0)
Regular amplitude envelope

PULSE 8.
200 ms long
Fundamental frequency 680Hz
Regular harmonics
Phase 270
Offset 0
Amplitude 1 (last 10 = 0)
Regular amplitude envelope

Experimental Stimuli.

BURST 14.
Pulse 2
Attenuation:5,0,0...
Six pulse
0 ms gap between pulses
Pitch = 210Hz

BURST 11
Pulse 3
Attenuation:5,0,0...
Six pulses
0 ms gap between pulses
Pitch = 250Hz

BURST 8.
Pulse 4
Attenuation:5,0,0...
Six pulses
0 ms gap between pulses
Pitch = 260Hz

BURST 9.
Pulse 5
Attenuation:5,0,0...
Six pulses
0 ms gap between pulses
Pitch = 320Hz

BURST 10.
Pulse 6
Attenuation:5,0,0...
Six pulses
0 ms gap between pulses
Pitch = 440Hz

BURST 13.
Pulse 7
Attenuation:5,0,0...
Six pulses
0 ms gap between pulses
Pitch = 500Hz

BURST 12.
Pulse 8
Attenuation:5,0,0...
Six pulses
0 ms gap between pulses
Pitch = 680Hz

APPENDIX 4B

SUBJECTIVE AND OBJECTIVE STIMULUS VALUES FITTED TO THE MATCHING FUNCTION BY THE METHOD OF LEAST SQUARES.

X	Y		
LOG STIM. MAG	LOG MEAN JUDGEMENT	XSQUARED	XY
2.322	2.171	5.391	5.041
2.397	2.166	5.745	5.191
2.414	2.229	5.827	5.380
2.505	2.245	6.275	5.623
2.643	2.289	6.997	6.049
2.698	2.304	7.279	6.216
2.832	2.368	8.020	6.706
SUM	SUM	SUM	SUM
17.81	15.77	45.53	40.20

$$A = \frac{N(\text{Sum } XY)-(\text{Sum } X)(\text{Sum } Y)}{N(\text{Sum } X^2)-(\text{Sum } X)} = \frac{7(40.20)-(17.81*15.77)}{7*45.53-17.81}$$

A (exponent) = 0.384

$$B = \frac{(\text{Sum } X^2)(\text{Sum } Y)-(\text{Sum } X)(\text{Sum } XY)}{N(\text{Sum } X^2)-(\text{Sum } X)}$$

$$= \frac{45.53*15.77-17.81*40.20}{7*45.53-17.81} = 1.28, \quad (\text{Antilog}=18.89)$$

B (constant) = 18.89

The matching function between line length and perceived urgency, as communicated by pitch, has a slope of:

0.384

APPENDIX 4C

COMPONENTS OF REPETITION STIMULI.

Pulse Used In Burst Creation.

PULSE 1

200 ms long

Fundamental frequency 300Hz

Regular harmonics

Phase 270

Offset 0

Amplitude 1

Regular amplitude envelope

Stimuli Used In Practice Trials.

BURST A.

Pulse 1

Attenuation:5,0,0...

Six pulses

5 ms gap between pulses

Pitch = 300,200Hz etc

BURST B.

Pulse 1

Attenuation:5,0,0...

Twelve pulses

5 ms gap between pulses

Pitch = 300,200Hz etc

BURST C.

PULSE 1

Attenuation : 5,0,0...

Eight pulses

5 ms gap between pulses

Pitch = 300,200 Hz etc

Experimental Stimuli.

BURST 15.
Pulse 1
Attenuation:5,0,0...
2 units of repetition (4 pulses)
0 ms gap between pulses
Pitch = 300,200 Hz. etc.

BURST 21.
Pulse 1
Attenuation:5,0,0...
2.5 units of repetition (5 pulses)
0 ms gap between pulses
Pitch = 300,200 Hz. etc.

BURST 16.
Pulse 1
Attenuation:5,0,0...
3 units of repetition (6 pulses)
0 ms gap between pulses
Pitch = 300,200 Hz.etc.

BURST 20.
Pulse 1
Attenuation:5,0,0...
3.5 units of repetition (7 pulses)
0 ms gap between pulses
Pitch = 300,200 Hz. etc.

BURST 17.
Pulse 1
Attenuation:5,0,0...
4 units of repetiton (8 pulses)
0 ms gap between pulses
Pitch = 300,200 Hz. etc.

BURST 18.
Pulse 1
Attenuation:5,0,0...
5 units of repetition (10 pulses)
0 ms gap between pulses
Pitch = 300,200 Hz. etc.

BURST 19.
Pulse 1
Attenuation:5,0,0...
6 units of repetition (12 pulses)
0 ms gap between pulses
Pitch = 300,200 Hz. etc.

APPENDIX 4D

SUBJECTIVE AND OBJECTIVE STIMULUS VALUES FITTED TO THE MATCHING FUNCTION BY THE METHOD OF LEAST SQUARES.

X	Y		
LOG STIM. MAG	LOG MEAN JUDGEMENT	XSQUARED	XY
0.301	2.044	0.090	0.615
0.397	2.118	0.157	0.840
0.477	2.165	0.227	1.032
0.544	2.191	0.295	1.191
0.602	2.222	0.362	1.337
0.698	2.262	0.487	1.578
0.778	2.289	0.605	1.780
SUM	SUM	SUM	SUM
3.797	15.29	2.223	8.373

$$A = \frac{N(\text{Sum } XY)-(\text{Sum } X)(\text{Sum } Y)}{N(\text{Sum } X^2)-(\text{Sum } X)} = \frac{7(8.373)-(3.797*15.29)}{7*2.223-3.797}$$

A (exponent) = 0.502

$$B = \frac{(\text{Sum } X^2)(\text{Sum } Y)-(\text{Sum } X)(\text{Sum } XY)}{N(\text{Sum } X^2)-(\text{Sum } X)}$$

$$= \frac{2.223*15.29-3.797*8.373}{7*2.223-3.797} = 1.91, \quad (\text{Antilog}=81.28)$$

B (constant) = 81.28

The matching function between line length and perceived urgency, as communicated by units of repetition, has a slope of:

0.502

APPENDIX 4E

COMPONENTS OF INHARMONICITY STIMULI

All harmonics are regular with amplitudes of 1 unless otherwise stated.

Pulses used In Practice Burst Creation.

Pulse 9.
200ms long
Fundamental frequency 300Hz
Irregular harmonic=3.2, amplitude=1.5
Phase 270
Offset 0
Regular amplitude emvelope.

Pulse 10.
200ms long
Fundamental frequency 300Hz
Irregular harmonic=4.4, amplitude=1.5
Phase 270
Offest = 0
Regular amplitude envelope.

Pulse 11.
200ms long
Fundamental frequency 300Hz
Irregular harmonic=5.6, amplitude=1.5
Phase 270
Offset 0
Regular amplitude envelope.

Stimuli Used In Practice Trials.

Burst D.
Pulse 9
Attenuation 5,0,0...
Six pulses
0 ms gap between pulses
Pitch = 300Hz

Burst E.
Pulse 10
Attenuation 5,0,0...
Six pulses
0 ms gap between pulses
Pitch = 300Hz

Burst F.

Pulse 11

Attenuation 5,0,0...

Six pulses

0 ms gap between pulses

Pitch = 300Hz

Pulses Used In Experimental Bursts.

Pulse 12.

200ms long

Fundamental frequency 300Hz

Regular Harmonics

Phase 270

Offset 0

Amplitude 1

Regular amplitude envelope.

Pulse 13.

200ms long

Fundamental frequency 300Hz

Irregular harmonic=8.5

Phase 270

Offset 0

Amplitude 1

Regular amplitude envelope.

Pulse 14.

200ms long

Fundamental frequency 300Hz

Irregular harmonics=4.5, 8.5, 12.5

Phase 270

Offset 0

Amplitude 1

Regular amplitude envelope.

Pulse 15.

200ms long

Fundamental frequency 300Hz

Irregular harmonics=3.1, 4.5, 8.5, 11.1,12.5

Phase 270

Offset 0

Amplitude 1

Regular amplitude envelope

Pulse 16.

200ms long

Fundamental frequency 300Hz

Irregular harmonics = 3.1, 4.5, 5.9, 8.5, 11.1, 12.5, 13.9

Phase 270

Offset 0

Amplitude 1

Regular amplitude envelope.

Pulse 17.

200ms long

Fundamental frequency 300Hz

Irregular harmonics = 3.1, 4.5, 5.9,7.1, 8.5, 9.9, 11.1, 12.5, 13.9.

Phase 270

Offset 0

Amplitude 1

Regular amplitude envelope.

Pulse 18.
200ms long
Fundamental frequency 300Hz
Irregular harmonics = 2.2, 3.1, 4.5, 5.9, 6.4, 7.1, 8.5, 9.9, 10.7, 11.1, 12.5, 13.9, 14.8, 15.3
Phase 270
Offset 0
Amplitude 1
Regular amplitude envelope.

Experimental Stimuli.

Burst 22.
Pulse 12
Attenuation 5,0,0..
Six pulses
0ms gap between pulses
Pitch = 300Hz.

Burst 23.
Pulse 13
Attenuation 5,0,0...
Six pulses
0ms gap between pulses
Pitch = 300Hz.

Burst 24.
Pulse 14
Attenuation 5,0,0...
Six pulses
0ms gap between pulses
Pitch = 300Hz.

Burst 25.
Pulse 15
Attenuation 5,0,0...
Six pulses
0ms gap between pulses
Pitch = 300Hz.

Burst 26.
Pulse 16
Attenuation 5,0,0..
Six pulses
0ms gap between pulses
Pitch = 300Hz.

Burst 27.
Pulse 17
Attenuation 5,0,0...
Six pulses
0ms gap between pulses
Pitch = 300Hz.

Burst 28
Pulse 18

Attenuation 5,0,0...

Six pulses

0ms gap between pulses

Pitch = 300Hz.

Values of Harmonics In Experimental Stimuli.

Burst	Harmonics
22,	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15., ,
23,	1,2,3,4,5,6,7,8.5,9,10,11,12,13,14,15., ,
24,	1,2,3,4.5,5,6,7,8.5,9,10,11,12.5,13,14,15., ,
25,	1,2,3.1,4.5,5,6,7,8.5,9,10,11.1,12.5,13,14,15., ,
26,	1,2,3.1,4.5,5.9,6,7,8.5,9,10,11.1,12.5,13.9,14,15., ,
27,	1,2,3.1,4.5,5.9,6,7.1,8.5,9.9,10,11.1,12.5,13.9,14,15., ,
28,	1,2.2,3.1,4.5,5.9,6.4,7.1,8.5,9.9,10.7,11.1,12.5,13.9, 14.8,15.3.

APPENDIX 4F

SUBJECTIVE AND OBJECTIVE STIMULUS VALUES FITTED TO THE MATCHING FUNCTION BY THE METHOD OF LEAST SQUARES.

X	Y		
LOG STIM. MAG	LOG MEAN JUDGEMENT	XSQUARED	XY
0.	1.801	0	0
0.477	1.862	0.227	0.888
0.698	1.896	0.487	1.323
0.854	1.943	0.714	1.641
0.954	1.922	0.910	1.834
1.146	1.922	1.313	2.203
SUM	SUM	SUM	SUM
4.12	11.34	3.651	7.892

$$A = \frac{N(\text{Sum } XY)-(\text{Sum } X)(\text{Sum } Y)}{N(\text{Sum } X^2)-(\text{Sum } X)} = \frac{6(7.892)-(4.12*11.34)}{6*3.651-4.12}$$

A (exponent) = 0.121

$$B = \frac{(\text{Sum } X^2)(\text{Sum } Y)-(\text{Sum } X)(\text{Sum } XY)}{N(\text{Sum } X^2)-(\text{Sum } X)}$$

$$= \frac{3.651*11.34-4.12*7.892}{6*3.651-4.12} = 1.80, \quad (\text{Antilog}=64.62)$$

B (constant) = 64.62

The matching function between line length and perceived urgency, as communicated by number of inharmonic components, has a slope of:

0.121

APPENDIX 4G

COMPONENTS OF INHARMONICITY STIMULI (Ex.8).

All harmonics are regular with amplitudes of 1 unless otherwise stated.

Pulses Used In Practice Burst Creation.

Pulse 9.
200ms long
Fundamental frequency 300Hz
Irregular harmonic=3.2,
Amplitude = 1.5
Phase 270
Offset 0
Regular amplitude envelope.

Pulse 10.
200ms long
Fundamental frequency 300Hz
Irregular harmonic = 4.4,
Amplitude = 1.5
Phase 270
Offset 0
Regular amplitude envelope.

Pulse 11.
200ms long
Fundamental frequency 300Hz
Irregular harmonic = 5.6, amplitude = 1.5
Phase 270
Offset 0
Regular amplitude envelope.

Stimuli Used In Practice Trials.

Burst D.
Pulse 9
Attenuation 5,0,0...
Six pulses
0 ms gap between pulses
Pitch = 300Hz

Burst E.
Pulse 10
Attenuation 5,0,0...
Six pulses
0 ms gap between pulses
Pitch = 300Hz

Burst F.

Pulse 11

Attenuation 5,0,0...

Six pulses

0 ms gap between pulses

Pitch = 300Hz

Pulses Used In Experimental Bursts.

Pulse 19

200ms long

Fundamental frequency 300Hz

Regular harmonics

Phase 270

Offset 0

Regular amplitude envelope.

Pulse 20

200ms long

Fundamental frequency 300Hz

Irregular harmonic = 3.1,

Amplitude = 1.5

Phase 270

Offset 0

Regular amplitude envelope.

Pulse 21

200ms long

Fundamental frequency 300Hz

Irregular harmonic = 3.3,

Amplitude = 1.5

Phase 270

Offset 0

Regular amplitude envelope

Pulse 22

200ms long

Fundamental frequency 300Hz

Irregular harmonic = 3.5,

Amplitude = 1.5

Phase 270

Offset 0

Regular amplitude envelope.

Pulse 23

200ms long

Fundamental frequency 300Hz

Irregular harmonic = 3.7,

Amplitude = 1.5

Phase 270

Offset 0

Regular amplitude envelope.

Pulse 24

200ms long

Fundamental frequency 300Hz

Irregular harmonic = 3.8,

Amplitude = 1.5

Phase 270

Offset 0

Regular amplitude envelope.

Pulse 25
200ms long
Fundamental frequency 300Hz
Irregular harmonic = 3.9, amplitude = 1.5
Phase 270
Offset 0
Regular amplitude envelope.

Experimental Stimuli.

Burst 29.
Pulse 19
Attenuation 5,0,0...
Six pulses
0ms gap between pulses
Pitch = 300Hz

Burst 30
Pulse 20
Attenuation 5,0,0...
Six pulses
0ms gap between pulses
Pitch = 300Hz

Burst 31
Pulse 21
Attenuation 5,0,0...
Six pulses
0ms gap between pulses
Pitch = 300Hz

Burst 32
Pulse 22
Attenuation 5,0,0...
Six pulses
0ms gap between pulses
Pitch = 300Hz

Burst 33
Pulse 23
Attenuation 5,0,0..
Six pulses
0ms gap between pulses
Pitch = 300Hz

Burst 34
Pulse 24
Attenuation 5,0,0...
Six pulses
0ms gap between pulses
Pitch = 300Hz

Burst 35
Pulse 25
Attenuation 5,0,0...
Six pulses
0ms gap between pulses
Pitch = 300Hz

APPENDIX 4H

SUBJECTIVE AND OBJECTIVE STIMULUS VALUES FITTED TO THE MATCHING FUNCTION BY THE METHOD OF LEAST SQUARES.

X LOG STIM. MAG	Y LOG MEAN JUDGEMENT	XSQUARED	XY
0.477	1.581	0.227	0.754
0.491	1.674	0.241	0.821
0.518	1.734	0.268	0.898
0.544	1.753	0.295	0.953
0.568	1.731	0.322	0.983
0.579	1.701	0.335	0.984
0.591	1.717	0.349	1.014
SUM	SUM	SUM	SUM
3.768	11.891	2.039	6.41

$$A = \frac{N(\text{Sum } XY)-(\text{Sum } X)(\text{Sum } Y)}{N(\text{Sum } X^2)-(\text{Sum } X)} = \frac{7(6.41)-(3.768*11.89)}{7*2.039-3.768}$$

A (exponent) = 0.85

$$B = \frac{(\text{Sum } X^2)(\text{Sum } Y)-(\text{Sum } X)(\text{Sum } XY)}{N(\text{Sum } X^2)-(\text{Sum } X)}$$

$$= \frac{2.039*11.89-3.768*6.41}{7*2.039-3.768} = 1.24, \quad (\text{Antilog}=17.45)$$

B (constant) = 17.45

The matching function between line length and perceived urgency, as communicated by stimulus inharmonicity, has a slope of:

0.85

APPENDIX 5A

COMPONENTS OF STIMULI EMPLOYED IN EXPERIMENT 9.

PULSES USED IN BURST CREATION.

These pulses were used to define and construct the low, medium and high urgency levels of the pitch variations.

PULSE 26(NON URGENT).
200ms. long
Fundamental frequency = 70Hz.
15 regular harmonics.
Phase 270
Amplitude 1, regular envelope.

PULSE 27(MID URGENT).
200ms. long
Fundamental frequency = 170Hz.
15 regular harmonics
Phase 270
Amplitude 1, regular envelope.

PULSE 28(URGENT)
200ms. long
Fundamental frequency = 380Hz.
10 regular harmonics.
Phase 270
Amplitude 1, regular envelope.

STIMULI USED IN PRACTICE TRIALS.

BURST G.
Pulse 26
2 pulses (1 unit of repetition)
400ms gap between pulses
Pitch= 65, 43Hz
Total length = 800ms
Attenuation 5,0...

BURST H.
Pulse 28
12 pulses (6 unit of repetition)
7ms gap between pulses
Pitch = 400, 266Hz
Total length = 2484ms
Attenuation 5,0...

BURST I.
Pulse 27
4 pulses (2 unit of repetition)
50ms gap between pulses
Pitch = 300, 200Hz

Total length = 1000ms
Attenuation 5,0...

EXPERIMENTAL STIMULI.

S=SPEED. R=UNITS OF REPETITION. P=PITCH.
L=LOW URGENCY LEVEL. M=MID URGENCY LEVEL. H=HIGH URGENCY LEVEL.

BURST 36
SHRHPH
Pulse 28
7 pulses(3.5 unit repetition)
14ms between pulses
Pitch = 380,252.5 Hz
Total length = 1498ms
Attenuation = 5,0,0...

BURST 37
SHRHPM
Pulse 27
7 pulses(3.5 unit repetition)
14ms between pulses
Pitch = 170,112.5 Hz
Total length = 1498ms
Attenuation = 5,0,0...

BURST 38
SHRHPL
Pulse 26
7 pulses(3.5 unit repetition)
14ms between pulses
Pitch = 70,47.5 Hz
Total length = 1498ms
Attenuation = 5,0,0...

BURST 39
SHRMPH
Pulse 28
4 pulses(2 unit repetition)
14ms between pulses
Pitch = 380,252.5 Hz
Total length = 856ms
Attenuation = 5,0,0...

BURST 40
SHRMPM
Pulse 27
4 pulses(2 unit repetition)
14ms between pulses
Pitch = 170,112.5 Hz
Total length = 856ms.
Attenuation = 5,0,0...

BURST 41
SHRMPL
Pulse 26
4 pulses(2 unit repetition)
14ms between pulses
Pitch = 70,47.5 Hz
Total length = 856ms
Attenuation = 5,0,0...

BURST 42
SHRLPH
Pulse 28

BURST 43
SHRLPM
Pulse 27

2 pulses(1 unit repetition)
14ms between pulses
Pitch = 380,252.5 Hz
Total length = 428ms.
Attenuation = 5,0,0...

2 pulses(1 unit repetition)
14ms between pulses
Pitch = 170,112.5 Hz
Total length = 428ms
Attenuation = 5,0,0...

BURST 44
SHRLPL
Pulse 26
2 pulses(1 unit repetition)
14ms between pulses
Pitch = 70,47.5 Hz
Total length = 428ms.
Attenuation = 5,0,0...

BURST 45
SMRHPH
Pulse 28
7 pulses(3.5 unit repetition)
70ms between pulses
Pitch = 380,252.5 Hz
Total length = 1890ms
Attenuation = 5,0,0...

BURST 46
SMRHPM
Pulse 27
7 pulses(3.5 unit repetition)
70ms between pulses
Pitch = 170,112.5 Hz
Total length = 1890ms
Attenuation = 5,0,0...

BURST 47
SMRHPL
Pulse 26
7 pulses(3.5 unit repetition)
70ms between pulses
Pitch = 70,47.5 Hz
Total length = 1890ms
Attenuation = 5,0,0...

BURST 48
SMRMPH
Pulse 28
4 pulses(2 unit repetition)
70ms between pulses
Pitch = 380,252.5 Hz
Total length = 1080ms
Attenuation = 5,0,0...

BURST 49
SMRMPM
Pulse 27
4 pulses(2 unit repetition)
70ms between pulses

BURST 50
SMRMPL
Pulse 26
4 pulses(2 unit repetition)
70ms between pulses

Pitch = 170,112.5 Hz
Total length = 1080ms
Attenuation = 5,0,0...

Pitch = 70,47.5 Hz
Total length = 1080ms
Attenuation = 5,0,0...

BURST 51
SMRLPH
Pulse 28
2 pulses(1 unit repetition)
70ms between pulses
Pitch = 380,252.5 Hz
Total length = 540ms
Attenuation = 5,0,0...

BURST 52
SMRLPM
Pulse 27
2 pulses(1 unit repetition)
70ms between pulses
Pitch = 170,112.5 Hz
Total length = 540ms
Attenuation = 5,0,0...

BURST 53
SMRLPL
Pulse 26
2 pulses(1 unit repetition)
70ms between pulses
Pitch = 70,47.5 Hz
Total length = 540ms
Attenuation = 5,0,0...

BURST 54
SLRHPH
Pulse 28
7 pulses(3.5 unit repetition)
140ms between pulses
Pitch = 380,252.5 Hz
Total length = 2380ms
Attenuation = 5,0,0...

BURST 55
SLRHPM
Pulse 27
7 pulses(3.5 unit repetition)
140ms between pulses
Pitch = 170,112.5 Hz
Total length = 2380ms
Attenuation = 5,0,0...

BURST 56
SLRHPL
Pulse 26
7 pulses(3.5 unit repetition)
140ms between pulses
Pitch = 70,47.5 Hz
Total length = 2380ms

BURST 57
SLRMPH
Pulse 27
4 pulses(2 unit repetition)
140ms between pulses
Pitch = 380,252.5 Hz
Total length = 1360ms

Attenuation = 5,0,0...

BURST 58
SLRMPM
Pulse 27
4 pulses(2 unit repetition)
140ms between pulses
Pitch = 170,112.5 Hz
Total length = 1360ms
Attenuation = 5,0,0...

BURST 60
SLRLPH
Pulse 28
2 pulses(1 unit repetition)
140ms between pulses
Pitch = 380,252.5 Hz
Total length = 680ms
Attenuation = 5,0,0...

BURST 62
SLRLPL
Pulse 26
2 pulses(1 unit repetition)
140ms between pulses
Pitch = 70,47.5 Hz
Total length = 680ms
Attenuation = 5,0,0...

Attenuation = 5,0,0...

BURST 59
SLRMPL
Pulse 26
4 pulses(2 unit repetition)
140ms between pulses
Pitch = 70,47.5 Hz
Total length = 1360ms
Attenuation = 5,0,0...

BURST 61
SLRLPM
Pulse 27
2 pulses(1 unit repetition)
140ms between pulses
Pitch = 170,112.5 Hz
Total length = 680ms
Attenuation = 5,0,0...

APPENDIX 5B

POWER CALCULATIONS FOR MATCHED SAMPLE t TEST.

To calculate the effect size (Y) we used the equation, Y=u1-u2/o(x1-x2).

Estimate o(standard deviation of difference scores from parent population)
Difference scores from Experiment 1(x1-x2), (where the same magnitude estimation task was used as in the present experiment), were used in this calculation. Three pairs of consecutive stimuli were chosen to calculate the difference scores so that the scores would reflect the difference across the range of stimuli. The mean of subjects first two scores were used in the calculation because subjects will be making two judgements in Experiment 9. Stimulus notations are pulse rates.

Stimulus			Stimulus			Stimulus		
3.69	4.98	diff	7.78	9.63	diff	10	11.93	diff.
32	27.5	4.5	48.5	64.5	16	56.5	80.5	24
4.5	6	1.5	9	10	1	10.5	12.5	2
2	3	1	4.5	5	.5	5.5	7	1.5
1.5	2	.5	4.5	6.5	2	7.5	11	3.5
16.5	17.5	1	31.5	39	7.5	46	60	14
17.5	17.5	0	37.5	45	7.5	50	75	25
1.5	6	4.5	7	10.5	3.5	10	12.5	2.5
12.5	22.5	10	55	70	15	77.5	82.5	5
1.5	2.5	1	4	7.5	3.5	7.5	12.5	5
6.5	8	1.5	12.5	30	17.5	27.5	42.5	15
3	25	22	30	60	30	55	85	30
16.5	16.5	0	25	37.5	12.5	45	45	0
st.dev.=6.35			st.dev.=8.81			st.dev.=10.62		

Mean standard deviation of difference scores = 8.59(o)

Estimate u1-u2(expected mean of the differnce scores)
This estimate reflects the smallest difference in mean scores that would be practically or theoretically useful. To make this estimate the data from previous experiments that varied the parameters of interest were examined.

Expt 1(speed): Response range 9-34 = 34 data points/7 stimuli = average of 4.8 data points between consecutive stimuli.
Expt 5(pitch): Response range 164-249 = 85 data points/7stimuli = average of 12 data points between consecutive stimuli.
Expt 6(repetition units): Response range 143-219 = 76 data points/7 stimuli = average 10 data points between consecutive stimuli.

We are interested of differences of 5 data points or more between the same urgency levels of the different parameters eg PH,RH,SH. Because we have predicted that the same levels will be judged equally urgent across parameters, we are only interested in a difference that would mean that they were not judged at the same level. On the basis of previous data that difference in at least 4.8(5) data points, that is what has been found to separate urgency levels.

Calculate the effect size

Y = 5/8.59 = 0.582

Y = 0.582 (medium effect size)

In order to estimate the required sample size the power must be selected. Power tables in Howell(1982) show that when o (the power) = 3.70 there is a 96% change of correctly detecting an existing effect and a 4% chance of missing an existing effect, when the significance level is 0.05.

Now that the power has been calculated to is possible to work out the required sample size (N),

N = (o/Y)^2

N = (3.70/0.582)^2 = 40.4

The required sample size for the proposed experiment is 40 subjects.

APPENDIX 5C.

MEANS AND STANDARD DEVIATIONS OF EACH SUBJECTS RESPONSE TO EACH STIMULUS (UNTRANSFORMED).

UNTRANSFORMED MEANS OF SCORES TO EACH STIMULUS.

BURST	N	MEAN	MEDIAN	TRMEAN	STDEV	SEMEAN
36	80	42.49	15.50	35.47	53.21	5.95
39	80	32.80	14.00	27.65	40.51	4.53
45	80	38.21	15.00	30.62	51.17	5.72
54	80	35.03	14.00	28.45	47.65	5.33
37	80	20.21	8.00	15.75	28.82	3.22
48	80	23.98	12.50	20.21	28.84	3.22
57	80	32.92	12.50	26.34	46.01	5.14
51	80	24.03	9.50	19.79	29.81	3.33
46	80	24.20	10.00	20.27	30.27	3.38
40	80	31.57	12.00	25.52	44.99	5.03
42	80	17.55	7.00	13.76	24.80	2.77
60	80	17.02	7.25	13.77	23.19	2.59
38	80	17.64	6.50	13.53	25.16	2.81
49	80	26.01	10.00	21.34	35.10	3.92
55	80	19.25	9.00	14.80	26.97	3.02
47	80	26.46	10.00	22.32	32.98	3.69
41	80	19.99	8.00	16.40	25.34	2.83
43	80	10.66	4.50	7.65	16.79	1.88
52	80	21.95	10.00	17.78	30.71	3.43
58	80	13.99	5.00	10.51	21.14	2.36
50	80	14.06	6.00	10.63	21.23	2.37
56	80	16.57	5.50	12.45	26.34	2.95
61	80	16.52	8.00	12.76	22.86	2.56
59	80	11.00	5.00	8.03	17.43	1.95
53	80	11.10	5.00	8.12	17.86	2.00
44	80	12.41	6.00	9.73	17.60	1.97
62	80	9.78	4.00	6.73	17.49	1.96

APPENDIX 5D.

SPEARMANS RANK CORRELATION COEFFICIENT BETWEEN PREDICTED AND OBTAINED URGENCY RANKINGS.

PREDICTED	OBTAINED
1	1
2	2
2	3
2	4
5	5
5	8
5	15
5	6
5	9
5	11
11	10
11	14
11	16
11	18
11	19
11	12
18	13
18	22
18	25
18	20
18	21
24	23
24	24
24	26
27	27

Spearmans rank correlation coefficient between predicted and obtained order = 0.901. Significant at p=0.01.

APPENDIX 5E.

MULTIPLE REGRESSION OF MEAN LOG. MAGNITUDE ESTIMATES AGAINST SPEED, PITCH AND REPETITION. (FULL AND REDUCED MODELS)

	URGENCY	SP.	PCH.	REPS	S*P	S*R	P*R
1	1.339	3	3.00	2.93	9.00	8.79	8.7900
2	1.277	3	3.00	2.00	9.00	6.00	6.0000
3	1.246	3	2.01	2.93	6.03	8.79	5.8893
4	1.232	2	3.00	2.93	6.00	5.86	8.7900
5	1.186	1	3.00	2.93	3.00	2.93	8.7900
6	1.184	2	3.00	2.00	6.00	4.00	6.0000
7	1.138	1	3.00	2.00	3.00	2.00	6.0000
8	1.109	3	3.00	0.84	9.00	2.52	2.5200
9	1.099	3	2.01	2.00	6.03	6.00	4.0200
10	1.093	2	3.00	0.84	6.00	1.68	2.5200
11	1.088	2	2.01	2.93	4.02	5.86	5.8893
12	1.031	2	2.01	2.00	4.02	4.00	4.0200
13	1.018	1	3.00	0.84	3.00	0.84	2.5200
14	0.976	1	2.01	2.93	2.01	2.93	5.8893
15	0.954	3	0.93	2.93	2.79	8.79	2.7249
16	0.935	3	2.01	0.84	6.03	2.52	1.6884
17	0.905	1	2.01	2.00	2.01	2.00	4.0200
18	0.901	3	0.93	2.00	2.79	6.00	1.8600
19	0.892	2	0.93	2.93	1.86	5.86	2.7249
20	0.847	2	2.01	0.84	4.02	1.68	1.6884
21	0.818	2	0.93	2.00	1.86	4.00	1.8600
22	0.786	1	0.93	2.93	0.93	2.93	2.7249
23	0.774	1	2.01	0.84	2.01	0.84	1.6884
24	0.706	1	0.93	2.00	0.93	2.00	1.8600
25	0.686	3	0.93	0.84	2.79	2.52	0.7812
26	0.686	2	0.93	0.84	1.86	1.68	0.7812
27	0.623	1	0.93	0.84	0.93	0.84	0.7812

```
MTB > REGRESS C1 6 C2 C3 C4 C5 C6 C7 (FULL MODEL)

The regression equation is
URGENCY = 0.296 + 0.0425 SPEED + 0.209 PITCH + 0.0731 REPS - 0.00295
  S*P + 0.0224 S*R - 0.00730 P*R

Predictor      Coef       Stdev      t-ratio      p
Constant      0.29641    0.06060      4.89     0.000
SPEED         0.04250    0.02420      1.76     0.094
PITCH         0.20911    0.02389      8.75     0.000
REPS          0.07309    0.02403      3.04     0.006
S*P          -0.002946   0.008425    -0.35     0.730
S*R           0.022358   0.008330     2.68     0.014
P*R          -0.007305   0.008046    -0.91     0.375

s = 0.03022      R-sq = 98.2%      R-sq(adj) = 97.7%


MTB > LET C9=(C2+C3)
MTB > NAME C9 'SP/PI.'
MTB > PRINT C9
```

SP/PI.(SPEED AND PITCH)

6.00	6.00	5.01	5.00	4.00	5.00	4.00	6.00	5.01	5.00	4.01
4.01	4.00	3.01	3.93	5.01	3.01	3.93	2.93	4.01	2.93	1.93
3.01	1.93	3.93	2.93	1.93						

MTB > REGRESS C1 3 C9 C4 C6 (REDUCED MODEL 1)

The regression equation is
URGENCY = 0.112 + 0.169 SP/PI. + 0.173 REPS - 0.0350 S*R

Predictor	Coef	Stdev	t-ratio	p
Constant	0.11185	0.05669	1.97	0.061
SP/PI.	0.16883	0.01242	13.59	0.000
REPS	0.17339	0.02151	8.06	0.000
S*R	-0.035021	0.008493	-4.12	0.000

s = 0.05860 R-sq = 92.4% R-sq(adj) = 91.4%

MTB > LET C10 = (C2+C4)
MTB > NAME C10 'SP/REP'
MTB > PRINT C10

SP/REP (SPEED AND REPETITION)

5.93	5.00	5.93	4.93	3.93	4.00	3.00	3.84	5.00	2.84	4.93
4.00	1.84	3.93	5.93	3.84	3.00	5.00	4.93	2.84	4.00	3.93
1.84	3.00	3.84	2.84	1.84						

MTB > REGRESS C1 3 C10 C3 C6 (REDUCED MODEL 2)

The regression equation is
URGENCY = 0.338 + 0.0458 SP/REP + 0.189 PITCH + 0.0236 S*R

Predictor	Coef	Stdev	t-ratio	p
Constant	0.33770	0.04270	7.91	0.000
SP/REP	0.04580	0.01809	2.53	0.019
PITCH	0.189173	0.007300	25.92	0.000
S*R	0.023558	0.008821	2.67	0.014

s = 0.03206 R-sq = 97.7% R-sq(adj) = 97.4%

MTB > LET C11 = (C3+C4)
MTB > NAME C11 'PI/REP'
MTB > PRINT C11

PI/REP (PITCH AND REPETITION)

5.93	5.00	4.94	5.93	5.93	5.00	5.00	3.84	4.01	3.84	4.94
4.01	3.84	4.94	3.86	2.85	4.01	2.93	3.86	2.85	2.93	3.86
2.85	2.93	1.77	1.77	1.77						

MTB > REGRESS C1 3 C2 C11 C6 (REDUCED MODEL 3)

The regression equation is
URGENCY = 0.150 + 0.131 SPEED + 0.173 PI/REP - 0.0265 S*R

Predictor	Coef	Stdev	t-ratio	p
Constant	0.14978	0.05017	2.99	0.007

SPEED	0.13055	0.01832	7.13	0.000
PI/REP	0.17253	0.01096	15.75	0.000
S*R	-0.026456	0.007132	-3.71	0.001

s = 0.05152 R-sq = 94.1% R-sq(adj) = 93.4%

APPENDIX 5Ea :

MULTIPLE REGRESSION - FULL MODEL WITHOUT INSIGNIFICANT INTERACTIONS

regress c1(URGENCY) 4 c2(SPEED) c3(PITCH) c4(REPETITION) c5(S*R)

The regression equation is
C1 = 0.336 + 0.0367 C2 + 0.189 C3 + 0.0586 C4 + 0.0224 C5

Predictor	Coef	Stdev	t-ratio	p
Constant	0.33590	0.03927	8.55	0.000
C2	0.03666	0.01711	2.14	0.043
C3	0.189173	0.006712	28.18	0.000
C4	0.05863	0.01756	3.34	0.003
C5	0.022358	0.008128	2.75	0.012

s = 0.02948 R-sq = 98.2% R-sq(adj) = 97.8%

APPENDIX 5Eb :

F TEST OF SIMILARITY OF REGRESSION COEFFICIENTS

General F statistic =

error ss(reduced) - error ss(full) % error ss(full)
df(reduced) - df(full) df(full)

A. TO TEST Ho: B1=B2

To see if coefficients for speed and pitch are the same.

0.07899 - 0.01912 % 0.01912 =
0.05987%0.01912 23 - 22
22 1 22
= 68.89

68.89 is significant at p= 0.01, critical value = 7.95

B. TO TEST Ho: B1=B3

To see if coefficients for speed and repetition are the same.

0.02365 - 0.01912 % 0.01912 =
0.00543%0.01912
23 - 22 22 1 22
= 5.21

5.12 is not significant at p= 0.01, critical value = 7.95

5.12 is not significant at p= 0.025, critical value = 5.79

5.12 is significant at p= 0.05, critical value = 4.30

C. TO TEST Ho: B2=B3

To see if coefficients for pitch and repetition are the same.

0.06104 - 0.01912 % 0.01912 = 0.04192%0.01912
23 - 22 22 1 22
= 48.23

48.23 is significant at p= 0.01, critical value = 7.95

All f values were looked up with 1, n-4 df, ie 1,23.

APPENDIX 6A

GLOSSARY OF TIME PERCEPTION TERMS.

Estimation Task.

The subject is presented with a stimulus and must estimate its length in conventional time units eg. 30 seconds.

Production Task.

The subject has to produce a duration equal to a standard stimulus expressed in conventional time units eg. 'Produce 30 seconds.'

Reproduction Task.

The subject has to produce a duration equal to a previously presented standard eg. 'Reproduce this duration.'

Comparison Task.

The subject has to say whether a second duration is longer or shorter than a first.

Subjective Time Rate.

The subjective experience of time. When subjective time rate increases, time passes slowly. More time units are registered per minute, thus the interval appears to be longer than it is, perceived duration is thus increased and it is overestimated. When subjective time rate is decreased,time passes quickly. Less units of time are registered and thus an interval seems shorter than it is, perceived duration is decreased and the interval is underestimated.

Under Estimation.

Subjective time rate is slow. Occurs when a stimulus or interval is judged as being shorter than it is objectively. Perceived duration is decreased.

Over Estimation.

Subjective time rate is fast. Occurrs when a stimulus or interval is judged as being longer than it is objectively. Perceived duration is increased.

Decreased Perceived Duration.

Subjective time rate is slow. Results in shorter judgements being made relative to an objective value, a standard or another stimulus (under estimations). Time passes quickly.

Increased Perceived Duration.

Subjective time rate is fast. Results in longer judgements being made relative to an objective value, a standard or another stimulus (overestimation). Time passes slowly.

Prospective Judgements.

Judgements made when the subject knows in advance that he/she will be required to make a temporal judgement

Retrospective Judgements.

Judgements made when subjects are only told that they will be required to make a temporal judgement after the interval, stimulus or task has occurred.

Filled Interval Effect.

This refers to the fact that temporal judgements have been shown to be affected by whether the interval to be judged is empty or full, and by the nature of the filler. It is usually claimed that filled intervals increase perceived duration relative to empty ones.

Indifference Interval.

It is a common finding that short intervals are overestimated and long ones are underestimated. In the middle of these two extremes are 'indifference intervals' which are estimated accurately, neither under nor over estimated.

Watched Pot Phenomena.

An increase in perceived duration when a subject is attentively waiting for an event.

APPENDIX 7A

STIMULI EMPLOYED IN EXPERIMENT TEN.

Pulses Used in Burst Creation

All harmonics below 4000Hz were weighted equally.

PULSE 29
200 ms long
Fundamental frequency 200Hz
15 regular harmonics
Phase 270
Offset 0
Amplitude 1
Regular envelope

PULSE 30
200 ms long
Fundamental frequency 450Hz
8 regular harmonics
Phase 270
Offset 0
Amplitude 1
Regular envelope

PULSE 31
200 ms long
Fundamental frequency 700Hz
5 regular harmonics
Phase 270
Offset 0
Amplitude 1
Regular envelope

PULSE 32
200 ms long
Fundamental frequency 300Hz
13 regular harmonics
Phase 270
Offset 0
Amplitude 1
Regular envelope

PULSE 33
200 ms long
Fundamental frequency 300Hz
Irregular harmonics - 3.1,
4.5,8.5,11.1,
13.9.
Phase 270
Offset 0
Amplitude 1
Regular envelope

PULSE 34
200 ms long
Fundamental frequency 300Hz
Irregular harmonics - 2.2,3.1,4.5,
5.9,6.4,7.1,8.5,9.9,10.7,11.1,12.5
13.9.
Phase 270
Offset 0
Amplitude 1
Regular envelope

PULSE 35
200 ms long
Fundamental frequency 300Hz
Irregular harmonics - 3.2
Phase 270
Offset 0
Amplitude 1
Regular envelope

Practice Bursts

BURST J
Pulse 32
Attenuation 5,0..
10 pulses
0 ms gap
Pitch 300Hz

BURST K
Pulse 35
Attenuation 5,0..
6 pulses
0 ms gap
Pitch 300Hz

BURST L
Pulse 32
Attenuation 5,0..
4 pulses
300 ms gap
Pitch 300Hz

Experimental Bursts

a. Pitch Stimuli

BURST63
Pulse 29
Attenuation 5,0..
12 pulses
0 ms gap
Pitch 200Hz
Length 2400ms

BURST64
Pulse 29
Attenuation 5,0..
8 pulses
0 ms gap
Pitch 200Hz
Length 1600ms

BURST65
Pulse 29
Attenuation 5,0..
4 pulses
0 ms gap
Pitch 200Hz
Length 800ms

BURST 66
Pulse 30
Attenuation 5,0..
12 pulses
0 ms gap
Pitch 450Hz
Length 2400ms

BURST 67
Pulse 30
Attenuation 5,0..
8 pulses
0 ms gap
Pitch 450Hz
Length 1600ms

BURST 68
Pulse 30
Attenuation 5,0..
4 pulses
0 ms gap
Pitch 450Hz
Length 800ms

BURST 69
Pulse 31
Attenuation 5,0..
12 pulses
0 ms gap
Pitch 700Hz
Length 2400ms

BURST 70
Pulse 31
Attenuation 5,0..
8 pulses
0 ms gap
Pitch 700Hz
Length 1600ms

BURST 71
Pulse 31
Attenuation 5,0..
4 pulses
0 ms gap
Pitch 700Hz
Length 800ms

b. Inharmonicity Stimuli

BURST 72
Pulse 32
Attenuation 5,0..
12 pulses
0 ms gap
Pitch 300Hz
Length 2400ms

BURST 73
Pulse 32
Attenuation 5,0..
8 pulses
0 ms gap
Pitch 300Hz
Length 1600ms

BURST 74
Pulse 32
Attenuation 5,0..
4 pulses
0 ms gap
Pitch 300Hz
Length 800ms

BURST 75
Pulse 33
Attenuation 5,0..
12 pulses
0 ms gap
Pitch 300Hz
Length 2400ms

BURST 76
Pulse 33
Attenuation 5,0..
8 pulses
0 ms gap
Pitch 300Hz
Length 1600ms

BURST 77
Pulse 33
Attenuation 5,0..
4 pulses
0 ms gap
Pitch 300Hz
Lenght 800ms

BURST 78
Pulse 34
Attenuation 5,0..

BURST 79
Pulse 34
Attenuation 5,0..

BURST 80
Pulse 34
Attenuation 5,0..

12 pulses	8 pulses	4 pulses
0 ms gap	0 ms gap	0 ms gap
Pitch 300Hz	Pitch 300Hz	Pitch 300Hz
Length 2400ms	Length 1600ms	lenght 800ms

c.Repetition Stimuli

BURST 81	BURST 82	BURST 83
Pulse 32	Pulse 32	Pulse 32
Attenuation 5,0..	Attenuation 5,0..	Attenuation 5,0..
12 pulses	8 pulses	4 pulses
0 ms gap	0 ms gap	0 ms gap
Pitch 300,200Hz	Pitch 300,200Hz	Pitch 300,200Hz
Length 2400ms	Length 1600ms	Lenght 800ms
6 units	4 units	2 units

d. Speed Stimuli

BURST 84	BURST 85	BURST 86
Pulse 32	Pulse 32	Pulse 32
Attenuation 5,0..	Attenuation 5,0..	Attenuation 5,0..
11 pulses	7 pulses	4 pulses
20 ms gap	20 ms gap	20 ms gap
Pitch 300Hz	Pitch 300Hz	Pitch 300Hz
Length 2400ms	Length 1520ms	lenght 860ms
I/t = 4.5	I/t = 4.5	I/t = 4.5

BURST 87	BURST 88	BURST 89
Pulse 32	Pulse 32	Pulse 32
Attenuation 5,0..	Attenuation 5,0..	Attenuation 5,0..
8 pulses	5 pulses	3 pulses
114 ms gap	114 ms gap	114 ms gap
Pitch 300Hz	Pitch 300Hz	Pitch 300Hz
Length 2399ms	Length 1456ms	lenght 828ms
I/t = 3.1	I/t = 3.1	I/t = 3.1

BURST 90	BURST 91	BURST 92
Pulse 32	Pulse 32	Pulse 32

Attenuation 5,0..	Attenuation 5,0..	Attenuation 5,0..
4 pulses	3 pulses	2 pulses
533 ms gap	533 ms gap	533 ms gap
Pitch 300Hz	Pitch 300Hz	Pitch 300Hz
Length 2399ms	Length 1666ms	Lenght 933ms
I/t = 1.3	I/t = 1.3	I/t = 1.3

e. Resolution Stimuli

Pitch entries rounded to nearest 2.5Hz

BURST 93	BURST 94	BURST 95
Resolved	Resolved	Resolved
Pulse 32	Pulse 32	Pulse 32
Attenuation 5,0..	Attenuation 5,0..	Attenuation 5,0..
12 pulses	8 pulses	4 pulses
0 ms gap	0 ms gap	0 ms gap
Pitch 392,330,392,330	Pitch 392,330,392,	Pitch
392,330,294, 262,294,330,392,349,		330,262,294,330,
262Hz		
294,330,262Hz	262Hz	Lenght 800ms
Length 2400ms	Length 1600ms	

BURST 96	BURST 97	BURST 98
Unresolved	Unresolved	Unresolved
Pulse 32	Pulse 32	Pulse 32
Attenuation 5,0..	Attenuation 5,0..	Attenuation 5,0..
12 pulses	8 pulses	4 pulses
0 ms gap	0 ms gap	0 ms gap
Pitch 392,330,392.330,	Pitch 392,330,392,	Pitch
392,330,262, 262,294,330,392,349,		
330,262,330,262,		294Hz
330,262,294Hz	294Hz	Length 800ms
Length 2400ms	Length 1600ms	

BURST 99	BURST 100	BURST 101
Atonal	Atonal	Atonal
Pulse 32	Pulse 32	Pulse 32

Attenuation 5,0..
12 pulses
0 ms gap
Pitch 392,277,370,349
277,294,311,330,392,
370,277,349 Hz
Length 2400ms

Attenuation 5,0..
8 pulses
0 ms gap
Pitch 392,277,370,
349,277,294,311
330Hz
Length 1600ms

Attenuation 5,0..
4 pulses
0 ms gap
Pitch 392,277,370
349Hz
Length 800.sec

APPENDIX 7B

STIMULI EMPLOYED IN EXPERIMENT ELEVEN.

Pulse Used in Burst Construction

PULSE 36
200 ms long
Fundamental frequency 300Hz
13 regular harmonics
Phase 270
Offset 0
Amplitude 1
Regular envelope

Practice Trials.

BURST M	BURST N
Pulse 36	Pulse 31
Attenuation 5,0..	Attenuation 5,0..
10 pulses	6 pulses
0 ms gap	0 ms gap
Pitch 300Hz	Pitch 300Hz

BURST O
Pulse 36
Attenuation 5,0..
4 pulses
300 ms gap
Pitch 300Hz

Experimental Trials

a. Speed Stimuli

BURST 102	BURST 103	BURST 104
Pulse 36	Pulse 36	Pulse 36

Attenuation5,0..	Attenuation 5,0.	.Attenuation 5,0..
11 pulses	8 pulses	4 pulses
20 ms gap	114 ms gap	533 ms gap
Pitch 300Hz	Pitch 300Hz	Pitch 300Hz
Length 2400ms	Length 2399ms	length 2399ms

b. Resolution Stimuli.

Pitch entries rounded to nearest 2.5Hz

BURST 105
Resolved
Pulse 36
Attenuation 5,0..
12 pulses
0 ms gap
Pitch 392,330,392,330
262,294,330,392,349,
294,330,262Hz
Length 2400ms

BURST 106
Unresolved
Pulse 36
Attenuation 5,0..
12 pulses
Pitch 392,330,392,330
262,294,330,392,349
330,262,294Hz
0 ms gap
Length 2400ms

BURST 107
Atonal
Pulse 36
Attenuation 5,0..
12 pulses
0 ms gap
Pitch 392,277,370,349
277,294,311,330,392,
370,277,349 Hz
Length 2400ms