PHYSICAL AND PSYCHOLOGICAL FACTORS IN THE USE OF MOUTHPIECE FORCE IN TRUMPET PLAYING

Thesis Submitted in Partial Fulfilment of the
Requirements for the Degree of Doctor of Philosophy,

by

Patrick Kenny

Psychology Department
University of Strathclyde
Glasgow

1985

ABSTRACT

1) A strain gauge device was developed to measure the forces exerted in two dimensions on the trumpet mouthpiece during performance.

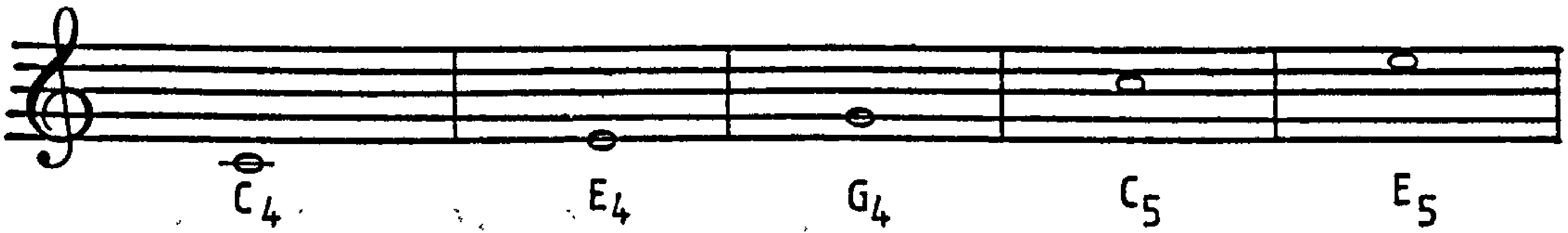
2) A group of top professional performers (n = 30) and a group of intermediate players (n = 30) were tested on a wide variety of musical material with mouthpiece force as the dependent variable. Data from the long note exercise in the normal register revealed significant effects for pitch and intensity on the dependent variable. No proficiency effects were observed. The scale and arpeggio exercises revealed a remarkable consistency in the application of mouthpiece force by all the subjects, such that force levels at each selected pitch/intensity combination appeared to be independent of context. The lip flexibility exercise however revealed significant contextual influences, and it was suggested that musical complexity might be responsible for this. Some proficiency related effects were observed: the professional players had greater levels of maximum tolerable forces, displayed more endurance and were more consistent in the application of mouthpiece force than the intermediate players across the range of tests.

3) An investigation was carried out on the assessments by trumpet players of the mouthpiece force usage of other players. The judgments involved rating photographs of players performing notes at known force levels. Groups of subjects differing in proficiency showed a high consensual agreement in judgments of mouthpiece force. However, equivalent proficiency groups rating the same photographs according to the criterion of apparent effort yielded almost identical rankings, suggesting that mouthpiece force judgments are made on the basis of this erroneous cue. The extent of consensual agreement and judgmental accuracy was found to be independent of proficiency level.

4) A psychophysical investigation of mouthpiece force using ratio production, magnitude production and magnitude estimation methods produced logarithmic sensory functions for both individual and group data. Judgmental accuracy was found to be unrelated to proficiency level.

PREFACE

This thesis uses the system of staveless pitch identification. The basis of the system is that middle C is identified as C_4, the reason being that there are approximately four octaves of audible sound below this pitch. The C's below C_4 are C_3, C_2 and C_1, while the tones above are C_5, C_6 and C_7.



Also, all the written music in this study is the notation for the standard Bb trumpet. Thus, a written C_4 is realy Bb_3 in terms of absolute pitch.

The study refers throughout to mouthpiece "force". In fact, the usual way musicians refer to this phenomenon is as mouthpiece "pressure". Strictly speaking, pressure is measured as force per unit area. To use this measure would unnecessarily complicate the experiment. By measuring mouthpiece force, questions as to the effective area of the mouthpiece rim become irrelevant, and simplicity is maintained. However, the forces are described throughout as kilograms rather than Newtons. Newtons are the standard scientific force measurement. It was felt to be permissible to use kilograms because this is a unit more readily accessible to nonphysicists. In any case the conversion is simply accomplished; 1kg is equivalent to 1N.

CONTENTS

Abstract ... ... (ii)

Preface ... ... (iii)

CHAPTER 1 INTRODUCTION 1

The Behavioural Studies 2

Studies from the Physical Sciences 11

The role of the Present Study 43

The Thesis Plan... 46

CHAPTER 2 INSTRUMENTATION 48

Transducer No. 1: Feasibility Study 54

Transducer No. 2 56

Transducer No. 3 66

CHAPTER 3 FORCES ON THE EMBOUCHURE DURING TRUMPET PERFORMANCE 78

Introduction 78

Previous Studies 79

Independent Variables 84

Dependent Variables 90

Sequencing Effects 90

Stability of Mouthpiece Force 91

Nature of the General Design 91

Hypotheses 93

General Method 93

I <u>Long Note Study</u> 102

II Diverse Musical Exercises 120

III Tests to Determine the Limits of the use of Mouthpiece Force. 138

IV Endurance Test 150
V Judgmental Tasks 154
Stability of Mouthpiece Force 160
Concluding Discussion 165

CHAPTER 4 JUDGMENTS OF MOUTHPIECE FORCE ON THE BASIS OF VISUAL CUES 178
Introduction 178
Method 187
Results 191
Discussion 204

CHAPTER 5 PSYCHOPHYSICAL INVESTIGATION OF MOUTHPIECE FORCE 208
Introduction 208
Method 220
Results 226
Discussion 237

CHAPTER 6 A CONCLUDING NOTE 244

References 247

Appendix 259

CHAPTER ONE

INTRODUCTION

This thesis is the report of an empirical investigation into certain specific aspects of musical performance on the trumpet. There is a large and ever growing body of knowledge gained from the use of scientific methods in music. However, this study does not readily fit into any of the broad categories that have dominated such research up until now. This departure from the familiar modes of experimentation has been motivated largely by the questionable utility of the traditional research paradigms in this area. This introduction therefore proposes to present a review of the major studies in this field and attempt a critical examination of what substantive achievements (if any) have been made. It is only against this background that the purpose of the present study can be fully appreciated.

The most striking division which is evident in the literature is that between the physical and the behavioural sciences. Each approach has been fruitful in its own way, but there is a high degree of mutual exclusivity which has sometimes not been beneficial to either line of investigation. We will examine each in turn.

THE BEHAVIOURAL STUDIES

Almost all of the techniques available in general psychology for the study of human behaviour have at one time or another been applied in a musical context. This has generated an enormous body of data across a very wide range of phenomena. However, as this thesis concentrates on the performance of a specific musical instrument, a large part of the more puristic work covered by the psychology of music is not particularly relevant. There is indeed something of a dearth of psychological experimentation on the playing of individual musical instruments.

Potentially relevant research seems to be divided between studies in the general area of the acquisition of skills, and the studies with an evaluative or educational bias. The reason for this is probably that it seemed natural for psychologists to follow the well worn conceptual paths which behavioural researchers had beaten out in the past in the study of more traditional phenomena. The results of the application of these methods to musical <u>performance</u> have not been rewarding.

Human Skills Research

On the face of it one might have expected that human skills research would be a rich source of data about performance aspects of music. This however is not the case. The layman who consults a psychology textbook or journal on human skills in the hope of learning about those activities he understands as involving skill will be exceedingly disappointed. Psychologists use the word skill in a very different sense (Welford, 1958), preferring to look at skills in a much broader way. There is a tendency to study principally central processes and to develop very general theories purporting to explain skilled behaviour at a basic level. Psychologists have been more concerned with how a violinist takes his instrument out of its case rather than how he plays the Bruch Violin Concerto. Largely responsible for this state of affairs is the advent of information theory. Human skills research has for some time now concentrated on a view of man as a medium for the processing of information. The experimentation has been modelled on engineering and cybernetic theory (Shannon and Weaver, 1949). A huge body of work has been accomplished, with repeated use of the familiar types of task; vigilance, complex sensory detection, reaction time, continuous tracking, two channel listening and so on.

Several authorities (see for example Stammers and Patrick, 1975) view this type of emphasis in the study of perceptual motor skills as misplaced. Psychologists, in their attempts to investigate skills have frequently been too willing to sacrifice the utility of their research in order to maximise its empirical validity. The result has been the development of a discipline which is highly esoteric and academic in nature. It is an extreme outcome of the tension which permeates behavioural thinking between the scientific and the preconceived, the normative and the factual. Unfortunately, proceeding in this manner has left a very large gap between experimentation and real life phenomena. This thesis takes the view however that no stigma should attach to a more concrete study of musical performance which has the ultimate aim of helping to improve the standards of musical performance whether this aim be ultimately achieved by discovering ability or by improving achievement.

There are many reasons why a more direct approach is necessary here. Firstly, on a general note, Gagné (1962) has pointed out several examples of principles of skill which were established in the laboratory which have not held up when applied to real training problems. Secondly, differences in musical performance involve very fine differences at an already extremely high level of skill, and individual differences will assume great importance. Skills research, in looking for general laws of human behaviour does not seem very appropriate here. Thirdly and most specifically, skilled performance of the trumpet is achieved almost entirely without the aid of visual information. The role of visual information during the learning process is probably minimal also. However, much of the work on human skills has involved vision. While the findings based on this type of data are not without relevance at least conceptually, the perceptual area of skills research with the most applicability to brass (and indeed most wind instrument) performance is that of proprioception. The difficulties preventing extensive exploration of this sensory modality are obvious, and hence much of the research remains at a highly theoretical level (Bahrick 1957).

The philosophy of the present experiment is inspired by the approach which Seymour (1966) advocates, and is exemplified in much of the work in the field of industrial training psychology. Seymour observes that any particular task will have its own unique characteristics. These characteristics will decide the most effective training methods. It is vital to determine a system of principles and rules which tailor the training methods to the task characteristics. The attempt to improve performance begins with task analysis, (called job analysis in industry). This is where the information on the task behaviour is collected and the component activities are identified. Of course this type of approach makes it necessary to categorise behaviour. The important thing to keep in mind however, as Stammers and Patrick (supra) note, is that the utility of any taxonomy is completely dependent on the relevance and usefulness of the categorisation for the purpose of training, rather than the academic elegance of the taxonomic scheme itself.

As regards the present knowledge of the various functions and components of the processes involved in trumpet performance, we are still at a very early stage, and this thesis does not attempt to develop detailed training programs of any sort. It was decided instead to isolate a single phenomenon, namely mouthpiece force and examine this in some detail. (The reasons for selecting this particular variable for analysis will be discussed more fully later).

It is obvious that the present approach is fraught with many of the same conceptual and methodological dangers facing the traditional skills research methods, and some of the most pertinent will now receive attention. To begin with, skill acquisition research must ultimately consider how the discrete components of behaviour are organised into an integrated hierarchical structure. Isolated examination of these components can often only give a very limited picture of the overall skilled behaviour. Probably, the most effective suggestions for facilitating learning will emerge only with this wider view. However as Fitts and Posner (1973) point out, the breaking up into parts is often the only available method of approach, and that the most important thing is for the "dissection" to be accomplished in the right way. Secondly, it is probably naive to think that research into trumpet performance (or any other instrument for that matter) will readily yield many "universally valid optimums" (Wagner et al 1973).

Paillard (1960) observes that it has been shown that there can be a variety of possible muscular activity underlying superficially identical movements. This should not deter experimentation in this area. On the contrary, even more data should be collected in an attempt to develop a model of appropriate complexity for the behaviour in question.

Although much criticism has been levelled at the classical skills research methods, there are a number of paradigms which have been developed which could very usefully be applied to the study of trumpet performance. For example, devices could be designed which allow performance of the musical task with periodic or continuous presentation of physiological information to the subject. Such feedback studies could be expected to provide illuminating results. Another area for potential exploration is that of methods of practice; for instance the question of massed versus distributed training. Apart from some work on the clarinet, (see Sloboda 1985), there are scarcely any reported investigations into the question of the effects of practice methods on musical performance.

Evaluative and Educational Studies

As might be expected, behavioural research into the performance of musical instruments is better represented in the psychological literature with an educational slant. There is quite a variety of material. To begin with there have been a number of attempts to evaluate the effectiveness of different training programs. Studies on the trumpet included Testa's (1972) examination of the effect of jaw thrust instruction on the trumpet performance and overjet of young players, Schlaefer's (1975) comparative study of experimental methods for increasing range and endurance on the cornet and the trumpet, and Lacey (1969) and Nichols et al (1971) have tested the efficacy of more general training programs. The advantage of these types of study is that they have very high face validity and are hence more attractive to performers and educators. The fact that they have this seductive quality about them is very unfortunate because the experimentation itself has been largely flawed through lack of adequate controls. In Testa's study for instance, the evaluation of the training method was not accomplished by the single blind procedure (which should be mandatory in these types of experiment). Despite Testa's assertion that the testing instrument (the TEA or Trumpet Embouchure Assessment) called for little "subjective" assessment, there is clearly too much scope here for the intrusion of unwanted artifacts. Moreover, another feature of Testa's study which it is important to note is that it is an attempt to assess a method of correcting a fairly obvious aberration from normal trumpet performance which is caused by a relatively speaking rare physical abnormality, namely the pronounced protrusion of the upper mandible. Clearly, the study is really concerned with a form of remedial training. It may admittedly be of immense benefit to a limited number of performers, but does not have much relevance for the majority of trumpet students. Despite this, Testa's work is the best attempt to investigate empirically the specific effects of trumpet instruction and lights the way for future research to follow.

Much less satisfactory is Schlaefer's attempt to compare the "Anatowind" method of trumpet teaching with traditional methods for increasing range and endurance on the trumpet. Suspicion is immediately aroused by the almost evangelical style in which the experiment is reported, and by the extravagance of the claims made for the universal superiority of Joseph Simmons' patented method. A closer inspection of the study reveals methodological flaws which undermine the validity of the conclusions. Chief amongst these is the lack of blind scoring. Equally problemmatic is the extremely small size of the sample, with only five subjects assigned to each treatment group. Schlaefer also relies heavily on a non-statistical review and interpretation of the data. The experiment is in fact a classic example of the confusion of empirical fact and mere opinion in an area where rigorous scientific standards ought to be most stringently observed.

The studies of Lacey and Nichols et al appear to have made little impact in the literature; the latter is in any case unobtainable.

Attempts to evaluate teaching methods are obviously of great interest. As to what type of training methods are actually in use, irrespective of their efficacy, there have occasionally been surveys made to throw light on this question. Examples of such surveys are those by Bellamah (1960), Neff (1974) and Deye (1947) on the trumpet, and Tanner (1976) on brass instruments in general. These surveys however were largely carried out by means of loosely structured and open ended interview techniques, and are in any event a little dated now.

Musical performance is ideally suited to studies looking at the question of ability and aptitude. Surprisingly, very little of the work has concentrated on the playing of musical instruments. Lamp and Keys (1935) made an early study investigating the feasibility of predicting aptitude for specific musical instruments. Years later Pflieger (1971) looked at the relationship between certain psychomotor abilities and performance success on particular musical instruments. The author is not aware of any other work of this kind. Clearly, such research is of the greatest potential importance. More often than not the pupil's choice, (if he indeed has a choice) whether to take up a musical instrument, or of what specific instrument to select is made quite arbitrarily. It is sad to think that such an important decision, with far-reaching consequences in terms of possible wastage of time and resources, should be left to chance. More will be said on this subject later.

This short section just about covers the contribution that behavioural science has made to the improvement of trumpet performance. There are other studies involving the trumpet which were uncovered in an extensive search of the literature, but which only have tangential relevance. Fiske (1975 and 1977) has done some work on performance adjudication reliability, and there has been some interesting research exploring personality characteristics and their relationship with for instance, achievement in instrumental music (Kaplan 1961, and Kemp 1981a), and also with the particular instrument played by the musician (Kemp 1981b).

The major advances in the scientific analysis of musical performance have been by the physical sciences, and it is to these experiments that we will now turn.

STUDIES FROM THE PHYSICAL SCIENCES

There is a surprisingly large amount of research into physical factors in trumpet performance, much of which makes extensive use of highly advanced and sophisticated technology. The reason why this has occurred is perhaps as follows. Of all the musical instruments, it is probably the trumpet which offers the least external evidence of how expert performance is achieved. With manual manipulation confined to the relatively simple operation of the three valves of the modern Bb instrument, it is generally agreed that the most critical determinants of performance variables are located in the embouchure and within the body of the player (indeed on the baroque or natural trumpet which consists of a single length of tubing, there is no manual manipulation at all). One of the most astounding features of the best trumpet players is that they are able to perform a wide range of techniques both without any apparent effort and without much observable change in external physical appearance between these different techniques. Performers and teachers therefore, in their attempts to understand and communicate the physical operations underlying optimal trumpet performance have had to rely on the limited information from three sources. Firstly, the marginal external changes which are observable; secondly, the players subjective judgments of the physiological correlates of their own performances and thirdly, the information from the audible aspects of performance. Such sources have been shown to be notoriously unreliable and misleading guides to the real underlying mechanisms, and the inadequacy of such information has led to a wide variety of theories and opinions as to how expert trumpet performance is achieved, many of which are in direct contradiction to each other. The widespread lack of agreement among leading brass authorities has been documented by several opinion surveys (see for example Richtermeyer 1966, and Trosper 1962). In fact Weast (1961) went so far as to describe brass teaching as being in a "state of chaos". Many of the physical studies are a reaction against the conjecture and lack of objectivity inherent in the teaching literature, and are attempts to move towards a more scientifically based understanding of the phenomenon of brass performance. We will now consider in detail some of the different approaches which have been used.

X-Ray Studies

Radiological techniques initially appeared to offer the most effective way of studying the controversial intra-oral, intra-pharyngeal, intra-thoracic and intra-abdominal variables, and these methods form the basis of the majority of the physical investigations into trumpet performance.

Developed as a diagnostic aid in medicine, radiology became to be used in other problem areas, notable among which is the science of speech production (Holbrook and Carmody 1941). The speech researcher Perkell (1969) suggested that there are divisions of physiological articulatory activity which correspond to the resultant acoustic distinctions. When eventually Hall (1954) used X-ray photography to study trumpet performance, the link with speech research became important, because one of the most heated arguments in trumpet teaching concerned the positioning of the jaw and tongue, and that argument tended to be conducted with reference to articulatory positions during the production of <u>speech</u> sounds. This disagreement between the prominent brass authorities is revealed by the opinion surveys of Deye (1947) and Aurand (n.d.), and is reflected in the disparity between the major recommendations and instructions found in trumpet teaching manuals. There can be found four categories of theory as to the unvoiced vowel formations used in trumpet playing, each with its own adherents. The single syllable theorists believe either that the "oo" formation* is used at all times (see for example Arban 1936; Eby 1933; St Jacome 1894), or that "ee" or perhaps some other formant is employed. The third group of authorities hold that vowel formations change with different registers of the instrument, usually "ah" for the lower tones and "ee" for the higher tones (see for example Bach 1925; Clarke 1934; Colin 1948). Finally, the fourth group believe that players use personal vowel or syllable sounds which change with the tone quality desired. For this view see Deye (1947), Sweeney (1948).

In order to look into this question, Hall conducted a photographic study using single exposure, lateral head radiography. The players performed long notes, and different pitches were tried to see if register affected physiological formation. There were strict controls for sound intensity. Data were taken from the individual X-ray photographs. A matrix (overlay) was used to ensure standardisation of measurement. On the basis of data taken from nine subjects, all professional performers, Hall's

---

* The vowel sounds discussed here correspond to the phonetician Heffner's (1950) beat (ee), boot (oo) and father (ah).

principal conclusion was that trumpet performers tend to use their own physiological formations, most commonly approximating to "ah", the intermediate phonetic vowel sound. Furthermore, Hall observed that the trumpet players tended to use the same basic formation for each register. Variations in formation between different registers were not large, and in nearly all the cases these changes were not as great as the changes between the extreme vowels. However, Hall did note some consistency in the minor observable changes in those few subjects who did appear to manipulate intra-oral formation with change in pitch, namely in the direction of the "ee" position for the upper register, and in the direction of the "oo" position for the lower register. There are several problems with Hall's study which need to be examined. Firstly, the use of only nine subjects, albeit professional players must raise doubts as to the generality of the conclusions, especially as individual differences were found to be of primary importance. Secondly, there are inherent deficiencies in single exposure radiographic methods. Despite the fact that the procedures do yield pictures from which satisfactory measurements can be taken, there is the limitation that these pictures have to be taken during the performance of isolated and sustained single tones, whereby the physiological structures are held immobile for the duration of the exposure. It is unlikely that single samples of intra-oral positions for long tones will yield enough information on performance variables. This is because it is probable that contextual factors will be a major influence on physiological formation during normal musical performance. In speech research it has been shown (Stetson et al 1940) that the physiological characteristics of speech sounds vary dramatically with phonetic context, and there is good reason to think that musical context will have a similar effect on the physiological correlates of trumpet performance.

However, Hall's study was challenging in several respects. It suggested that there were no major salient physiological formations of a universal kind which can be identified in the playing of advanced performers, and it highlighted the importance of individual differences and personal adaptations to the task. Clearly, any models of intra-oral formation would need to be of a more complex character, and require the application of far more refined techniques.

More rapid progress seemed to be promised by the advent of cinefluorography. This method involves the taking of motion pictures from a fluoroscopic screen (Ramsey et al 1960). The technique does not require the static conditions necessitated by the use of single exposure procedures. It can examine continuous musical passages rather than isolated sounds, and provides more than one cross sectional time sample during musical performance, allowing movements as well as stationary positions to be studied.

Haynie (1967) pioneered the application of X-ray cineradiography to trumpet playing, and eventually adapted the procedure to incorporate the use of instant replay videotape. However, despite the use of the X-ray image intensifier (fluoroscope) on more than seventy subjects, Haynie did not attempt to express structural positions and movements in quantitative form. However, his study shows the diversity of factors which appear to interact to produce the observed intra-oral formations. Firstly, there is a wide range of anatomical structures among humans. People's intra-oral structures exist in as wide a variety as do their facial features. This presents a problem to research into structural and movement patterns because there is a difficulty in finding universal reference points for measurement. Secondly, Haynie finds a large amount of variation between the observed movements and formations for the performance of the same task between individual players. This is consistent with Hall's conclusion that players tend to use their own personal basic formations, but extends this finding to a much wider range of performance material, including actual musical passages. These individual adaptations might be dictated to some extent by the unique physiological anatomy of each player's intra-oral structures, (Frohrip 1972), but this need not necessarily be so, and even given similar physical makeup the same task can probably be performed in quite different ways. To add yet further to the confusion, contextual factors seem to operate. Haynie notes that tongue position and movement is markedly affected by the musical performance context itself. Effects are observed for the techniques of register and intensity changes as well as tonguing and slurring, all of which may interact with general levels of pitch and intensity. Given the overall complexity of the situation, it is not surprising to find that the subsequent cineradiographic investigations have been somewhat hard to interpret.

The research by Hiigel (1967), despite lacking any positive outcome as regards substantive teaching recommendations, is useful however in that it has dispelled some of the myths about the conceptualisation of tongue placement based on the enunciation of syllables. Hiigel took three series of cinefluorographic sound pictures as six brass players: (i) performed a prepared musical score, (ii) recited selected syllables, and (iii) as they performed the same musical score with "syllabic imagery": (the latter condition involves the performers playing the score while actively mimicking the related syllables during the performance). To analyse the movements of the intra-oral structures, sequences of individual cinefluorographic frames were traced and various measurements extracted from the tracings with the use of a superimposed standard grid template to provide reliable reference points for measurement. The experiment is a repeated measures design, allowing each subject to be his own control. Given the extent of the influence of individual differences and the small number of subjects, this is an extremely important consideration. Statistical analysis revealed that significant differences in tongue placement existed between the performance of pitches and the enunciation of syllables matched with these pitches. Furthermore, the use of syllabic imagery was found to be ineffective in promoting the associated tongue placements in actual performance. This finding really questions the utility of teaching systems which employ concepts of syllabic recitation as a means of describing or inducing particular types of tongue placement. Of course, such training methods may have beneficial effects, perhaps of an indirect kind, but until these can be conclusively demonstrated, there would appear to be no valid reasons, on the basis of current knowledge, for continuing to use them.

Moving away from a discussion of intra-oral factors in terms of syllabic enunciation, what can be said about the action of the tongue in trumpet performance? The literature is inconclusive on this matter. This lack of clarity is a result of the tendency of researchers in the area to leave uncontrolled several sources of variance. Firstly, of the seven major radiographic studies with the trumpet, only Amstutz (1970) uses more than 12 subjects. Secondly, some of the studies use a variety of brass instruments (including trumpets, horns and trombones) in an attempt to explore general factors in brass playing. With small numbers of subjects, this approach is misconceived.

Different brass instruments call for radically different techniques (there are scarcely any professional musicians who are expert at more than one type of brass instrument). Also, there is empirical support for the suggestion that the physical mechanisms underlying the performance of the various brass instruments are different, (see Meidt 1967). The third problem concerns the selection of the particular cinefluorographic frames for detailed analysis. The whole point of cinefluorography is that it enables the visualisation of the movement of hidden structures. To avoid complexity however, most of the studies proceed by selecting individual frames from which to extract information. But this information loses much of its value unless it is collated into a sequential frame by frame analysis. Unfortunately, the techniques currently used in the science of movement analysis (see McCutcheon et al 1977), were relatively undeveloped at the time of the cineradiographic investigations, and so these studies retain many of the disadvantages associated with single frame studies. Intimately linked with the last problem is the fourth source of variability to be considered, namely musical context. The various cinefluorographic studies to their credit used more convincing musical material than their long note predecessors. However, the diversity of test items introduced an amount of variability which makes it hard to integrate the findings of the different studies. For example, Haynie (supra) observed that in an ascending arpeggio exercise some players used a sort of "flicking" action of the tongue to accompany the change in harmonics, such that the tongue was rarely in a stationary position at all during the exercise. Haynie noted that tongue position also seemed to be influenced by attempts by the player to correct the intonation or change the style or sound. This shows that the choice of particular frames for analysis is a critical part of the procedure. Deficiencies in this respect have thus rendered many of the studies somewhat artificial. The final drawback in the literature is that there has been little attempt to systematically examine the relationship between <u>proficiency</u> and physiological correlates of performance. It is hard to see how useful teaching suggestions will emerge from such research unless this factor is properly investigated.

We shall now consider in more detail the individual studies to examine what they claim to have revealed as regards intra-oral phenomena in brass performance. All the cinefluorographic studies report that there is an overriding tendency for the tongue to arch as the register increases, thus narrowing the gap between the highest point of the posterior arch of the tongue and the palate. The extent and character of the movement is however disputed. Hiigel found a slight amount of upward movement as pitch ascended, but in no way matching the degree of movement found in syllabic recitation. Meidt however found greater degrees of movement in the predicted direction among some of his students, but again falling short of the extremes of syllabic structures. By far the most extensive study of trumpet players was that conducted by Amstutz on 25 performers. He found varying degrees of tongue arching with register change in all the subjects and always in the predicted direction. The uniform result of this study needs to be tempered with the findings of two similar pieces of research by Frohrip (1972) and De Young (1975). Neither of these researchers found such a monolithic relationship between register differences and tongue behaviour. Unfortunately, the latter studies were exclusively on the trombone. However, they do throw some doubt on the findings of Amstutz, and suggest that a full scale experiment should be undertaken using methods of selecting individual X-ray frames which are less arbitrary and open to hidden bias.

We have so far concentrated on the action of the tongue as revealed by X-ray. We will now take a broader view of the radiographic findings, paying special attention to the theoretical issues which are most relevant to the present experiment.

The studies investigating the size of the pharyngeal opening and its relationship with performance in different registers have produced highly contradictory results. Hall (supra) observed a tendency for the pharyngeal opening to become enlarged as pitch increased. This conclusion agreed with data collected in studies on clarinet performance (Anfinson 1965), and the human voice (Whitworth 1961). Meidt's study however had already introduced the suggestion that different instruments can involve distinctly different physical manipulations. He found that the french horn players in his study reversed the tendency observed in the trumpet players and constricted the pharyngeal opening with increasing register. The most startling discrepancy in the literature however involves the studies of Frohrip and De Young. Both studies examined the same brass instrument (the trombone) but reached directly opponent conclusions. Frohrip's sample of players tended to decrease the opening of the throat as register increased, whereas De Young's subjects were found to increase the size of the throat opening under similar conditions. Such disparities are eloquent of the need to perform experiments on a much larger group of performers under much more closely controlled conditions.

There seems to be much more agreement between the authorities concerning the question of teeth aperture and alignment. As regards alignment, the professional players in the seven major studies manipulate the jaw to more or less equalise the teeth. The players with a pronounced overjet which results from protruding upper teeth or receding lower jaw tend to compensate for this by projecting the lower jaw to the amount necessary to achieve equalisation, (for this see especially Haynie, 1967). This would seem to tie up with the teaching study of Testa (supra) which looked at the effect of jaw thrust instruction on trumpet performance and overjet.

The consensus among the X-ray studies suggests that the teeth aperture (the distance between the tips of the upper and lower teeth) varies systematically with register, with the size of the opening decreasing with ascending pitch and vica versa. Meidt's is the only study not to find this effect for trumpeters, although the predicted relationship was found among the horn players. Once more it must be noted that the sample sizes were very small.

Instrument pivot is another important parameter to consider. Studies into the degree of pivot are generally conducted by measuring the angle between the straight line along the face of the mouthpiece and the line joining the face of the lower teeth to the tip of the mandible. Amstutz and Haynie note a tendency for the angle of pivot to increase as pitch ascends. These studies reveal how deceptive the player's external physical appearance can be. The subject's head motions were not restricted in any way, and the X-ray pictures showed that the angle of pivot could be changed by tilting the trumpet (which is the most obvious movement for an onlooker to detect), or by tilting the head back (which is much less obvious), or by a combination of these movements. This is a good example of the problems facing teachers, whose judgments may be detrimentally affected by deceptive visual cues. This thesis will examine in some detail the judgmental problems associated with the assessment of the use of mouthpiece force.

Some interest has been shown in the intra-oral techniques involved in tonguing and attack, particularly as regards the extent of closure between the tongue and the palate necessary for clean attack (see especially Meidt), and also the vertical placement of the tip of the tongue. This type of research presents the severest problems. It is very hard to obtain readable photographs of the tip of the tongue in radiography as it does not show up very well on the image intensifier. It is therefore necessary to use a quantity of radiopaque paste, (e.g. microbarium sulphate) to delineate this area satisfactorily. Haynie (1967) notes that this coating acts as a definite handicap to normal performance, as the lips cannot be licked or moistened before starting the attack. Despite these reservations however, Haynie concludes that the extent of aperture of the teeth will be the determining factor in the positioning of the tip of the tongue in tongued attacks.

One of the most interesting areas of X-ray research attempts to uncover the relationship between different physiological operations and the resultant tonal qualities produced. Every trumpet player has his own unique, identifiable sound. This uniqueness must be reflected ultimately at a physiological level. It is crucial that this individual variation is thoroughly explored. Without this, it will be so much the harder to identify and extract the "universally valid optimums" which, as discussed earlier, is one of the main objectives of this type of study. Hall, in his long note study claims that the observed tongue placement of the subjects was indicative of discernible differences in sound quality. The electronic analysis of the sound was made by means of spectrograms made on a sonograph. An association was found between particular physiological formations and the formant pattern of the tone, in that there was a concentration of intensity at particular frequency areas. The extent of this relationship led Hall to suggest that the physiological formation of the phonatory channel is much more important in the determination of the characteristics of the sound than are the differences between individual instruments (assuming that they are adequately constructed). This also puts into perspective the obsessional frequency with which many trumpet performers change between different makes and models of mouthpiece and instrument in the hope of effecting a radical improvement in sound and performance.

There are some more general points which arise from the X-ray work. One observation is that the researchers have not given due consideration to the "post hoc ergo propter hoc" fallacy. When certain variables have been found to be significantly related, the researchers have too often (see especially Amstutz 1970), talked in terms of direct influences on performance rather than mere statistical associations. Although attributions of causality in the more physically orientated studies are less hazardous than in psychology generally, caution is advised in this particular area. As a demonstration of this, the author made some informal experimental attempts to examine the effects of deliberate manipulations of physical formation on performance with some of the subjects in this study. It was found that many subjects were able to play long notes without detectable variation in tone quality despite simultaneously and continuously effecting quite radical intra-oral manipulations, (for example by assuming extreme tongue positions). Researchers should therefore be put on their guard, and there would seem to be scope for experimentation which might systematically investigate this phenomenon.

Another point of interest which emerges is the relationship between the physiological and the psychological aspects of performance. With singers, Russell (1928) remarked on the pronounced disparity between the tongue positions the singers used as revealed by X-ray, and the tongue positions that they believed they were using. The phenomenological study of skilled behaviour has potentially enormous implications for teaching. Brass players often rely on subjective judgments about the way they themselves perform, and errors at this stage could be very detrimental to their methods of instruction as teachers. The present study recognises the importance of an empirical understanding of the structure of immediately lived experience, and the phenomenological aspects of one performance variable, namely mouthpiece force, will be examined in some detail in the experimental sections.

It is unfortunate that the cinefluorographic experiments were not more conclusive. This is a pity because it is unlikely that any large scale use can now be made of the technique by way of replication or elaboration of the studies which have received mention so far. The problem is that at the moment medical science cannot point to a level of X-radiation which could be described as a "safe dosage". The previous investigators were using radiation levels on subjects which would now be thought of as reckless, and this is despite the use of more and more powerful image intensifiers to reduce the overall exposure. Currently fears about the carcinogenic and destructive properties of X-rays have confined the use of radiography to when it is strictly medically necessary. Despite these developments, all is not lost because conceivably the much less harmful technique of Ultrasound may yet be used as a replacement to cinefluorography. In fact, several researchers have already used this method in the context of speech research, (see e.g. Kelsey et al 1969).

Electromyographic Studies

The next domain of research to be considered is that which has as its object the study of the function of the muscles in trumpet performance. Particular attention has been paid to the muscles of the face which form the embouchure or "set", but there has been research into non-facial muscles involved in trumpet playing.

Isley (1971) identifies the problems which have arisen from the inadequacy of the relevant anatomical literature. For a long time there has been general agreement about the location and physical description of the facial muscles. However, the anatomists have not been agreed as to the specific functions of these muscles, even as regards the kinesiology of common facial expressions, much less musical performance.

The problems facing the study of muscle function are similar to those inherent in the X-ray work in that phenomena are internal and little can be gained by outward observation. But in addition, the situation is complicated by the fact that there are a large number of different muscles involved, all working in a complex co-ordinated action. Thus, the study of the functional anatomy of the facial musculature had to await the advent of the modern technique of electromyography. Electromyography involves the recording of minute fluctuations of electrical potentials generated by bundles of muscle fibres when in contraction (see Basmajian 1967). The measurement model is based on the existence of a linear quantitative relationship between the amplitude of the integrated action potentials and the tension produced by the voluntary contraction of the human muscle (Lippold 1952).

The major part of the work which has been done on the electromyography of the muscles has concentrated on muscles away from the face, perhaps because of the inherent complexity of the intricately interlaced facial musculature. As with the X-ray work considered already, the impetus for study has come from the speech researchers, interested in this case in the neuromuscular basis for speech production, particularly as a means of studying organic and functional disturbances in articulation, such as with dysarthria (see Leanderson et al 1967). There has however been E.M.G. work on activities with more apparent relevance to trumpet performance, for example De Sousa and Vitti's (1965) study of the action of the buccinator muscle during smiling, blowing and whistling, and a similar study on the cheek muscle by Blanton et al (1970).

There have however been several studies specifically examining trumpet performance. An impressive feature of these studies is that they have used indwelling dipolar fine wire electrodes (Basmajian and Stecko 1962). This technique involves the actual insertion of the electrode through the skin and into the body of the muscle itself. Surface electrodes could have been used for the electromyographic pickup of individual muscle activity, but are less effective in regions of multiple muscles, and a single electrode over one muscle can easily suffer indiscriminate pickup from other facial muscles in close proximity. Insertion of fine wire electrodes however, affords a high degree of functional separation. The really surprising thing is that the investigators were able to persuade such a large number of professional performers to submit to the use of this invasive technique, which often required the implantation of several electrodes at once to allow simultaneous measurement.

Isley and Basmajian (1973) conducted a study on eight performers. They took measurements from ten different facial muscles in varying patterns of four muscles at a time, while the subjects performed specific test notes at given dynamic intensities. They found considerable intersubject variations for the same task. Some players recruited particular muscles strongly while others relaxed the same muscles for the identical exercise. Isley and Basmajian noted moreover that no particular pattern of recruitment produced demonstrably better results than another. Two points need to be made about this study. Firstly, only eight subjects were used, which makes the data all the more vulnerable to the influence of individual differences. Secondly, the implications of the observed intersubject variance are in themselves of great theoretical and practical interest. The nature and effect of this variance is probably very similar to the situation already discussed in the context of the X-ray research.

The most extensive E.M.G. study of the embouchure was conducted by White (1972). White made electromyograms from four muscles simultaneously during the performance of fifty-one test items by eighteen trumpet players. The muscles studied were the orbicularis oris superioris (O.O.S.) and the orbicularis oris inferioris (O.O.I.) which are the upper and lower lip muscles respectively, and the levator anguli oris (L.A.O.) and depressor anguli oris (D.A.O.) which are two of the major out-of-lip muscles. The choice of muscles was made on the basis of the work by Isley (supra) as being the muscles most actively involved in trumpet performance. White's study is of very great utility for several reasons. Firstly, the experimental design was formulated to have direct bearing on the most controversial theories as to the function of the embouchure musculature. Secondly, and most important of all was the inclusion of proficiency as an experimental variable in the design. This factor had rarely been systematically investigated in the literature and the omission to do this weakened the implications that the earlier work had for teaching and the improvement of performance skills.

White found that register and intensity both positively affected the embouchure's muscular activity, thus negating the theory that there should be static tension levels during normal performance. The most illuminating finding however was that advanced trumpeters have greater tension levels out of the lips (L.A.O. and D.A.O.) than in the lips (O.O.I. and O.O.S.), ($p < .025$), while the beginning players showed no significant difference. This throws great doubt on the popular theory of brass playing (described by Farkas 1962), called the "drawstring" theory. This theory proposed that correct performance required the simultaneous application of two patterns of muscle activity; the puckering of the lips and the contraction of the cheek muscles into a sort of smiling configuration. These two actions were thought to be balanced in a "tug-of-war" in such a way that, as pitch or intensity increased, the tension in the two sets of contracting muscles mounted, but acted in perfect opposition, the result being a motionless embouchure. White's data effectively disconfirmed this theory, and supported the alternative theory of Stevens (1971) and others, that muscle activity should be concentrated in the out-of-lip muscles. This is also consistent with the common sense observation that the out of lip muscles are in any case larger and inherently stronger than the orbicularis oris.

White also found that the advanced players showed a smaller ratio of upper to lower lip potentials than the beginners, which supports Steven's view that learners should employ more muscle activity in the lower lip, and shows that despite being anatomically a single unit, the lips must be considered functionally as two seperate muscles.

Although one might quaere White's assumption of bilateral symmetry (the out-of-lip muscles of only one side of the face were subjected to electrode implantation) and he might have looked at more muscles than the four eventually chosen, (there are ten pairs of muscles which are potentially active in the embouchure), the study has generated several extremely useful suggestions towards the improvement of brass performance.

Taken together, the E.M.G. studies show that it may eventually be possible, after taking individual differences into account, to indentify specific postures and patterns of facial and jaw muscular activity as being the most efficient embouchure for particular players. This leads in turn to the question; can a player actually be taught in the intentional usage of scientifically accepted patterns of muscle activity anyway? It is conceivable for instance that players are limited by their anatomical makeup such that some would never ever be physically able to perform the appropriate manipulations. Moreover, successful training methods may well depend on some sort of feedback training. It is appropriate therefore that we now consider the issue of feedback in some detail.

Some E.M.G. studies have attempted to use extrinsic feedback to control the tension of the muscles in trumpet playing. However, none of these involved the muscles of the embouchure. These studies were attempts to control and reduce abnormally high levels of tension in muscles away from the critical facial area, and are more akin to general Biofeedback Relaxation Training (see e.g. Chen 1976), rather than training in the skilled use of groups of fine muscles. Henderson (1979) for example tried to control general levels of tension in the sternocleidomastoid muscle during blowing of the trumpet, and concluded that E.M.G. biofeedback was a feasible method of achieving this. Levēe et al (1976) reported on the successful clinical application of electromyographical feedback to a particular muscle dysfunction in a wind player, again by promoting the relief of excessive tension. Feedback, however has never been used as a method of controlling the muscles of the embouchure. How successful might such a method of training be? Unfortunately, most techniques involving the provision of extrinsic and augmented information about underlying physiological changes face severe problems when applied as an aid to learning. For example, the subject can come to rely on salient and augmented feedback to the exclusion of the more subtle intrinsic cues to performance. The distraction of the trainee from attention to proprioceptive and other cues may thus be counterproductive in the end. Ultimately, the trumpet player has to have learned to rely on these intrinsic cues entirely. Thus, the full value of the E.M.G. work will only be realised if it leads to the development of higher levels of discrimination in the proprioceptory channels. The work of Sime (1975) provides some optimism in this respect. He examined the relationship between a subject's perception of muscle tension and the actual tension as measured by an E.M.G. meter. He found that the subjects demonstrated a greater awareness of muscle tension at the end of biofeedback training than at the start.

The practical drawbacks of E.M.G. biofeedback methods would equally apply to attempts to train players to regulate other internal physiological structures using ultrasound, fluorography or any other investigative tools as a means of providing knowledge of results not present in the ordinary performance of the task.

In truth, the various practical attempts to use feedback as an aid in the improvement of musical skills have not achieved the sort of success which have recommended them to the instrumental tutors themselves. Seashore, (1906) used the "voice tonoscope" as a feedback aid in the improvement of pitch discrimination. He claimed that there was a remarkable permanency in the transfer of the training from the tonoscope sessions to the later musical situation. However, the tonoscope was never used as a recognised method of training singers. Despite such failures, it must be hoped that some types of feedback training, if properly conceived and developed, will eventually produce demonstrably beneficial results.

There are other factors which have interested researchers into muscular aspects of trumpet performance. An obvious area of importance is that of muscular fatigue. Trumpet playing places great strain on the muscles of the embouchure, both from the force exerted by the mouthpiece on the lips and from the tension levels which must necessarily be used to sustain a tone. Indeed the role of stamina in expert performance cannot be overemphasised. There are various general theories as to what causes fatigue; general failure of the contractile mechanism, (Merton 1954), neuromuscular block (Brown and Burns 1949), or perhaps central fatigue in the motoneurones (Reid 1928). Unfortunately, it is not presently known which mechanisms are at work in the embouchure of trumpet players. The great potential for research in this area is shown by the study of Briggs (1968). Briggs investigated the variation in oxyhaemoglobin levels among trumpet players and non-trumpet players. The phenomena of oxygen debt and muscle fatigue are obviously closely related. Briggs observed that there was a definite relationship between performance proficiency and oxyhaemoglobin saturation levels. The "outstanding" players were found to have below normal arterial oxyhaemoglobin saturation levels, while there was no difference in these levels between the average trumpet players and non-players. Briggs speculated that the principle of the oxygen-conserve reflex may therefore be an important influence on performance proficiency. As a complement to this work, the medical study by Faulkner and Horvath (1967a) shows dramatically the deleterious effects on trumpet players of the failure to reduce the oxygen debt, some of these effects being very long term in nature. Thus are raised fascinating possibilities for trumpet teaching, because there may therefore be certain definite physical organismic constraints on expert trumpet performance which, if identified in individuals, could provide a basis for predicting aptitude specifically for the trumpet. The importance of achieving this objective has already been discussed above.

Lip Vibration Studies

The next group of studies look at the manner in which the vibratory portions of the upper and lower lip behave during trumpet performance. The research done in this field is almost entirely of academic interest alone, lacking in practical application largely because proficiency related effects were not considered. Martin (1941) photographed the lip action of trumpet players through a specially constructed transparent (lucite) mouthpiece. He found (not surprisingly perhaps) that the amplitude of the lip vibration decreased as frequency increased. Henderson (1942) in a few simple experiments claimed to demonstrate that trumpet tones are originated by the upper lip vibrating as a single reed against a relatively fixed body, namely the lower lip. The main function of the lower lip was merely to control the vibration rate of the upper lip. Weast (1963) observed the vibration of the lips with the aid of a square plexiglass mouthpiece and a stroboscope. Forty-two brass players (including horns, trumpets, tubas and trombones) were studied. Weast confirmed Henderson's finding that the upper lip is the prime vibrating mechanism, with the lower lip either apparently inactive or inconsistent in behaviour. Weast does not however, propose what the precise function of the lower lip is. The E.M.G. studies (see especially White,1972) show that there is a degree of tension in the lower lip (orbicularis oris) during trumpet performance, so this must mean that some manoeuvre is being performed.

The picture being built up by these three early studies was thrown into confusion by the work of Leno (1970). Using high speed motion analysis cameras on a small group of trombone performers blowing through transparent mouthpieces, Leno observed a consistent <u>vibratory</u> action in the lower lip. Moreover, he noted that the vibration of the lips could not be described by simple analogy to single/double reeds or vibrating strings at all, describing the vibrational motion producing the sounds on a brass mouthpiece as uniquely complex. Clearly, more work is necessary to clarify the position. It would be very interesting for future research to concentrate on the relationship between the motion of the lips in vibration and the muscle behaviour as revealed by the E.M.G. technique.

External Observation Studies

Some researchers have explored the relatively limited information provided by the outward physical appearance of trumpet players, in an attempt to look for correlates of skilled performance. The first attempt was by Hall (1954). From photographs of professional players performing long tones, measurements were made between points on opposite sides of the mouth. One pair of points was located in the red of the lips, the other pair being at the corners of the mouth. Hall then attempted to relate the change in distance between these points and the various selected pitch levels. Unfortunately, no consistent pattern emerged as to the skilled manipulation of these anatomical variables. The lack of significant results can probably be attributed to individual differences.

Gibson (1972) studied the position of the upper and lower lips inside the mouthpiece using a cutaway rim device. Although there was large interindividual variation among the twelve performers as to the proportion of upper and lower lip in the mouthpiece generally, Gibson noted a striking regularity in the direction of <u>change</u> of the lip ratio between the low and the high register. Table 1.1 shows the marked tendency for the performers to move more lower and less upper lip into the mouthpiece as pitch ascends.

Table 1.1 Percentage of subjects using each category of upper/lower lip ratio inside the mouthpiece at each of five selected pitch levels. (Gibson, 1972)

	Even Placement	More Upper Lip	More Lower Lip	Total
C_4	0	75	25	100
C_5	0	66.6	33.3	100
C_6	8.3	25	66.6	100
G_6	8.3	16.6	75	100
C_7	0	33.6	66.6	100

Another startling result of Gibson's study was the relationship that was found to exist between the direction of the air flow from the lip aperture and differences in pitch. Eleven out of the twelve players changed the direction of their air stream in an upward direction as pitch ascended. Between C_4 and C_7 the mean upward shift was 15°. Overall, there are three notable features of Gibson's study. Firstly, the pitch range performed was extreme, with all twelve players encompassing three and a half octaves up to a full octave above top C. Secondly, the twelve players were top American artists, all of truly exceptional ability. These aspects enhance the validity of the conclusions as to the inherent elements of the "proficient embouchure". However, despite the tendencies observed in this group of highly skilled players, the third crucial feature to note is that two or three of the players demonstrated a reversed pattern of lip ratio and air flow direction from the general trend, without this being detrimental in any way to the quality of the performance. The experiment is thus another example of the tenacity of individual adaptations in the face of statistically valid optimum modes of performance.

It is to be hoped that skills researchers interested in the patterning of skilled movements will use some of the more modern techniques developed in Bioengineering. For example, lip and jaw motion has been measured using a videoscanning system (Mc Cutcheon et al 1977), which measures continuously the positions of small reflecting points on the face, allowing complex movement to be recorded and analysed. This type of arrangement would be ideal for trumpet performance research.

The final observational study to be reviewed concerns the survey and description of the superficial facial anatomy of trumpet players by Malek (1954). He looked at the physical characteristics of fifty-two "expert" players, classifying them as to lip thickness, evenness of teeth, and the rest position of the lower jaw, in an attempt to relate these variables to proficiency. Malek concluded that neither thinness nor thickness of the lips was detrimental to performance. Also, neither irregularity of the teeth nor recession or protrusion of the lower jaw were found to be bars to fine performance, unless these abnormalities were present in the extremest degree.

This completes the review of the physical investigations into brass playing.

THE ROLE OF THE PRESENT STUDY

This study has two main aims; to fill a gap in the literature on trumpet performance and to attempt a novel empirical approach to the exploration of certain performance variables.

The gap which needs to be filled is the lack of knowledge about a potentially crucial variable in trumpet playing, namely mouthpiece force. There are several reasons why the current study chose to investigate this particular phenomenon. Firstly, the correct use of mouthpiece force has always been a great bone of contention amongst trumpet teachers, and there is even now very little agreement between these authorities. Secondly, medical knowledge gives very strong prima facie grounds for believing that forces acting on the lips may have decremental effects on performance. The adverse effects of pressure applied to any surface of the human body are well known. Local tissue blanching oedema and anaesthesia are likely results, and the particularly delicate nature of the lip tissue is likely to magnify these effects, (Barbenel et al 1985). Briggs (1968) gives an excellent description of some of the probable effects of mouthpiece force:

> "Physiologically, when the mouthpiece is placed against the lips, and a hermetic seal is formed, pressure from the rim causes ischemia with resultant passive hyperemia due in part to the Di Palma effect. Initial blowing of the instrument appears to induce vasodilation which accounts for the passive hyperemia, and as blowing continues hypoxia causes various metabolic changes that result in labial deformation and a loss of sensitivity.
>
> When the mouthpiece is placed against the lips normal hemodynamics are interrupted. As the player produces a pitch the lip-reed mechanism necessitates a continuous flow of air from the lungs and this inhibits normal ventilatory function. No doubt, both ischemia and hypoxia are responsible for the fatigue effects which occur in the vibratory portion of the labial structure during trumpet performance".

Although there is no empirical evidence relating the behaviour of lip tissue and vascular structures under pressure from the trumpet mouthpiece to the factors of tone quality and endurance, many commentators (see e.g. Weast and Hake 1965) believe that there must be a definite association. The third reason for studying mouthpiece force derives from the subjective report of many trumpet players. There is a tendency among players to attribute short term loss of sensitivity, labial swelling and loss of control to the excessive use of force on the lips.

Despite these circumstances, very little research has been done into the measurement of forces exerted on the embouchure, (see chapter 3 for a review of the few reported studies). This was therefore a relatively uncharted area, ideal for exploratory research.

This study however always intended to be more than a purely physical analysis of mouthpiece force usage. The aim was to relate the obtained physical data to certain psychological constructs in a very direct way, giving a phenomenological framework to the experiment as a whole. Very few other musical performance studies have attempted to cut across the division between the psychological and the physical approaches in this way. For example, despite the availability of the methods of traditional psychophysics, which are an obvious means of approaching phenomenological variables, only Briggs (1968) has used these techniques to study brass performers. Briggs looked at psychophysical force judgments of labial strength (the forces which can be exerted by pressing the upper and lower lip together). He found no relationship between the force judgments and proficiency level.

It could be argued that the selection for study of a single attribute from the complex of factors involved in trumpet performance is an arbitrary way of proceeding. The "Unitarist" theory of aptitude and achievement for instance, denies the validity of breaking musicality up into component parts, (see Mursell 1937). However, the atomistic approach has gained in respectability since Seashore's (1919) intervention, and receives support from both psychologists (see e.g. Barrett and Barker 1973) as well as from musicians (many of whom consider the existence of several independent factors in musical ability as self evident). This thesis maintains that as long as sufficient care is taken with interpretative conclusions drawn from such data, it is empirically permissible to isolate individual elements for analysis.

THE THESIS PLAN

The work reported in this thesis takes the form of four separate experimental studies. The idea is that all of these discrete studies should be self contained enough to be read and understood on their own, yet there should be a logical conceptual progression linking them all together. This has been achieved at the expense of occasional repetition.

The first study (chapter two) is a report of the design, development and implementation of a unique method of measuring mouthpiece force using electrical strain gauges. As this application of transducer technology was completely new and quite complex, it was necessary to devote an entire chapter to a description of the measurement device.

The bulk of the second experiment (chapter three) involved the measurement of the mouthpiece force used by trumpet players across a wide range of performance material, paying special attention to the factors of subject proficiency and performance style. The remainder of chapter three explores some of the phenomenological issues raised by the initial exploratory work. The third and fourth experiments (chapters four and five) develop some of these issues, focusing on the nature and implications of two particular perceptual problems involved in judgments of mouthpiece force.

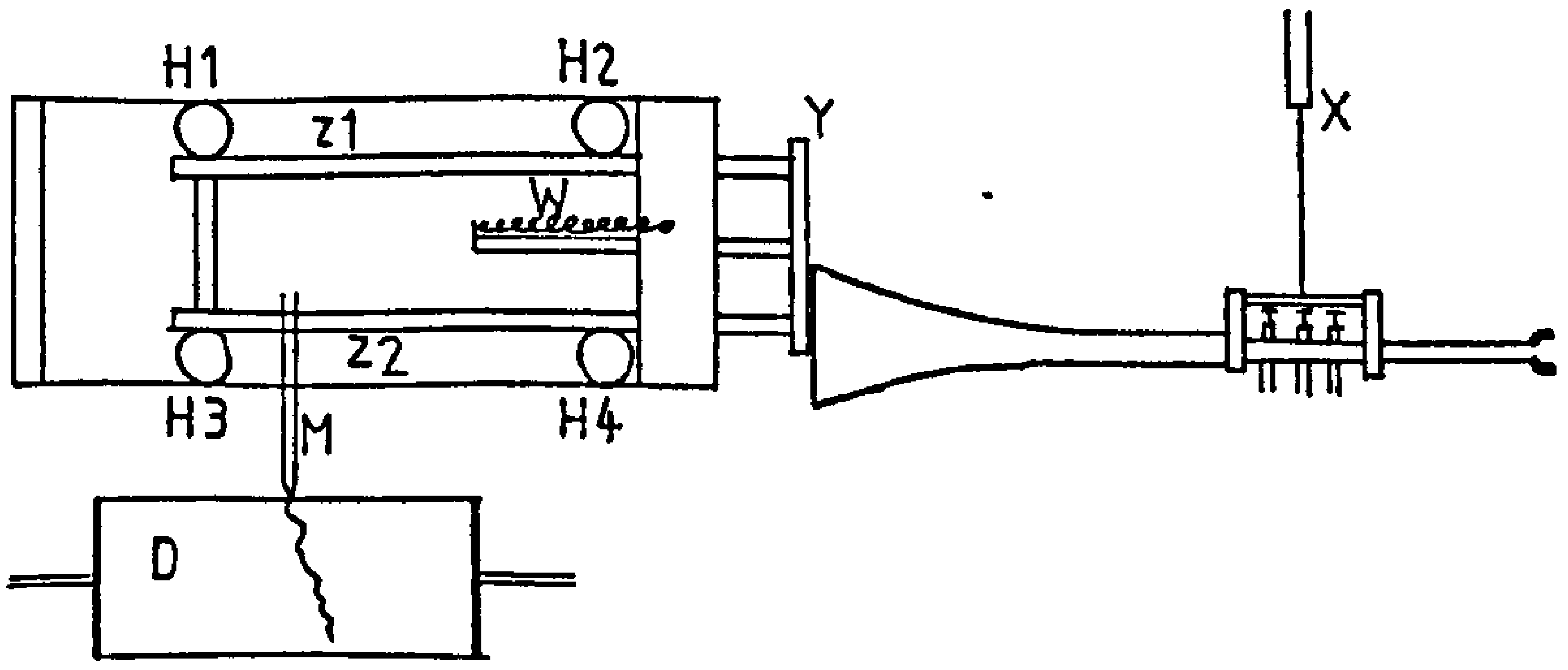
CHAPTER TWO

INSTRUMENTATION

Despite the potential importance of mouthpiece force as a major variable in trumpet performance, there have been very few attempts to investigate the phenomenon empirically. These early studies made use of relatively crude mechanical apparatus all using basically the same principle of measurement.

The first such study was undertaken by Henderson (1942). This gave a kymographic representation of force on the mouthpiece (see figure 2.1) The trumpet was suspended from a fixed point X by a wire. The subjects had to produce the notes by setting their lips against the mouthpiece of the trumpet and applying the force by moving their whole upper trunk forward.

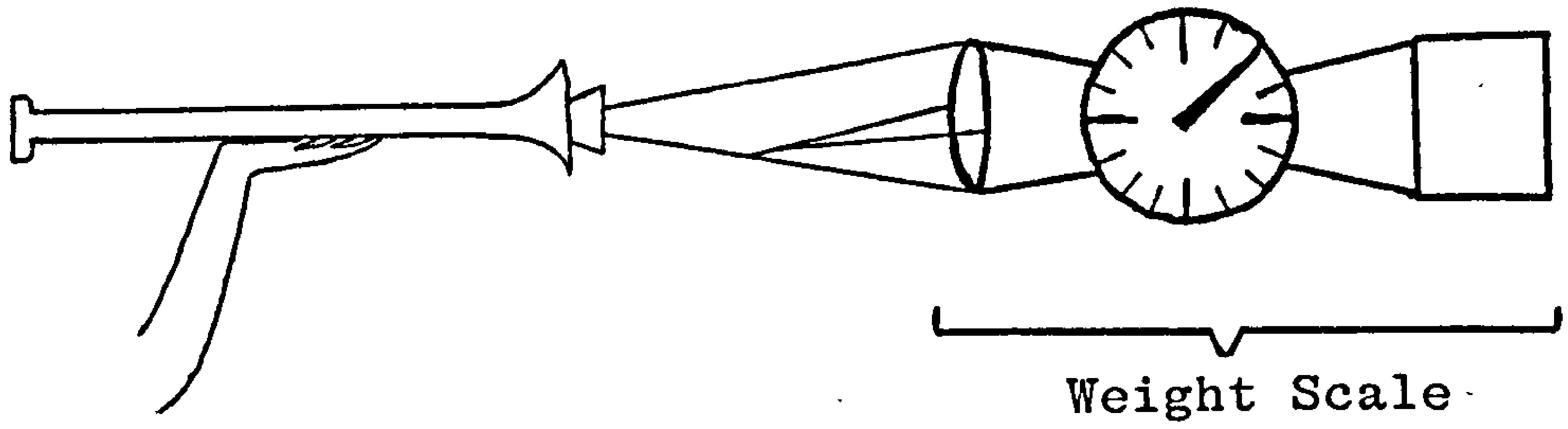
Figure 2.1 Henderson's (1942) kymographic apparatus.



This action forced the trumpet against a plate Y whose motion was resisted by a spring W. This resulted in the slides z1 and z2 of the plate Y moving over the bearings H1 - 4. A pointer M which was attached to the slide made contact with a revolving drum D. Thus, changes in the force on the mouthpiece were represented by etchings on the lightly waxed paper revolving on the drum. The instrument was calibrated with known weights giving an approximately linear scale.

A similar technique was employed by Weast and Hake (1965). In this case the apparatus consisted of a simple weight scale which was mounted with a set of felt padded rods which fitted into the bell of the trumpet. The trumpet was held by each subject in the flat palm of the hand (see figure 2.2). The readings taken from the weight scale itself were held to be a direct measure of the force being applied.

Figure 2.2 Weast and Hake's (1965) weight scale apparatus.



White (1972) used the same technique in his replication of Weast and Hake's study.

To the knowledge of the author these three experiments involve the only three systems which have been used so far to measure mouthpiece force. There are however several methodological weaknesses associated with these methods.

The major problem is that the subjects are obliged to play in a highly restricting way. As they are not allowed to manipulate the instrument they are very unlikely to employ the same mouthpiece forces as they would under normal playing conditions. It is clearly unsatisfactory for the means of measurement of an independent variable to radically alter the value of that variable. Secondly, the rudimentary nature of the equipment adds a further restriction on the type of experimental material which can be used in testing the subjects. As the valves of the instrument could not be used the subjects were limited to the performance of isolated notes in a single harmonic series, and it was not possible to test players with any extended sequences of notes or musical phrases. The artificial nature of the stimulus material makes it unlikely that the results obtained will be of any general validity. Thirdly, mouthpiece force is not a simple unitary concept. The previous studies used a single force reading as a measure of mouthpiece force. However, the situation is more complicated in that there are a number of different forces and moments acting on the mouthpiece. Whilst a single gross measure might have some heuristic value, it seems likely that an accurate picture can only be gained by a more detailed consideration of the relevant component forces. To overcome these three major problems it was decided to use a completely different system of measurement.

The chosen method involved use of the transducer. This is a device which changes energy from one form to another. The output is usually electrical as this is the most useful in terms of further signal processing. Transducers were originally developed for use in the physical sciences. They have more recently been used in the modern science of Bioengineering, for example in the design and testing of orthotic structures (Kirkpatrick et al 1969), and it is this line of research which suggested their potential for the present study.

The principle of transduction in wire strain gauges by which deformations can be measured by changes in electrical resistance is briefly as follows.

In stretching a wire, the resistance of the wire varies because of the changes in diameter, length and resistivity. This is given by the equation;

$$R = \frac{pL}{A}$$

where R = Resistance (Ohms)
p = Resistivity (Ohms/Metre)
L = Length (Metres)
A = Cross sectional area (Sq Metres)

Stretching the wire will increase its length to an extent which is inversely proportional to its modulus of elasticity. Also, the cross sectional area of the wire will decrease as its length increases. This decrease will conform to Poisson's ratio of the rod material. Taking into consideration the decrement in the wire's linear dimensions, it can be seen that there is a resistance change which is proportional to the amount of extension (strain) or the amount of force applied (stress). There are also other factors which vary with linear dimensional changes, for instance the forces may change the crystalline and mollecular structure of the material.

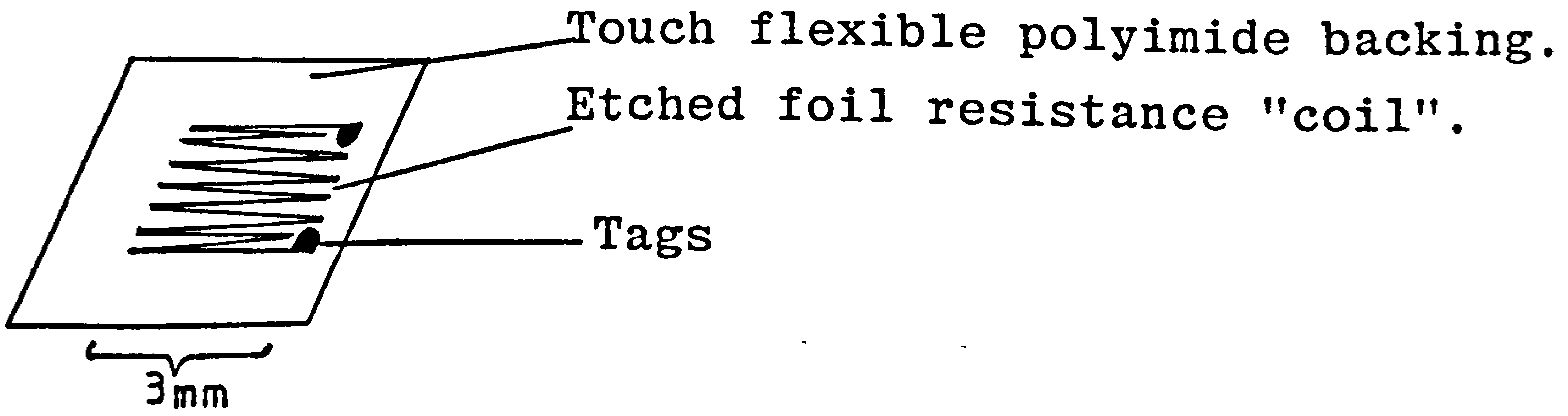
Taken together the overall behaviour of the wire under stretch is characterised by a designated gauge factor. This is the change in electrical resistance per unit change in length.

$$F = \frac{\Delta R \; \Delta L}{R \quad L}$$

where ΔR is the decremental change of resistance
ΔL is the decremental change of length
R is the resistance of the unstretched wire
L is the length of the unstretched wire
F is the Gauge Factor

A "bonded" wire strain gauge consists of a zig-zag pattern of wire bonded onto a plastic backing (see figure 2.3) This is then cemented onto the surface of the metal in which the strain is to be measured. Once cemented, the gauges become effectively part of the surface such that any strain occurring in that surface is transferred through to the gauge element unchanged. By mounting such gauges between the end of the trumpet mouthpiece and the point at which the instrument is held, it should be possible to measure the strain produced, and from this obtain the force exerted by the lips on the mouthpiece.

Figure 2.3 Schematic diagram of typical foil bonded resistance strain gauge.



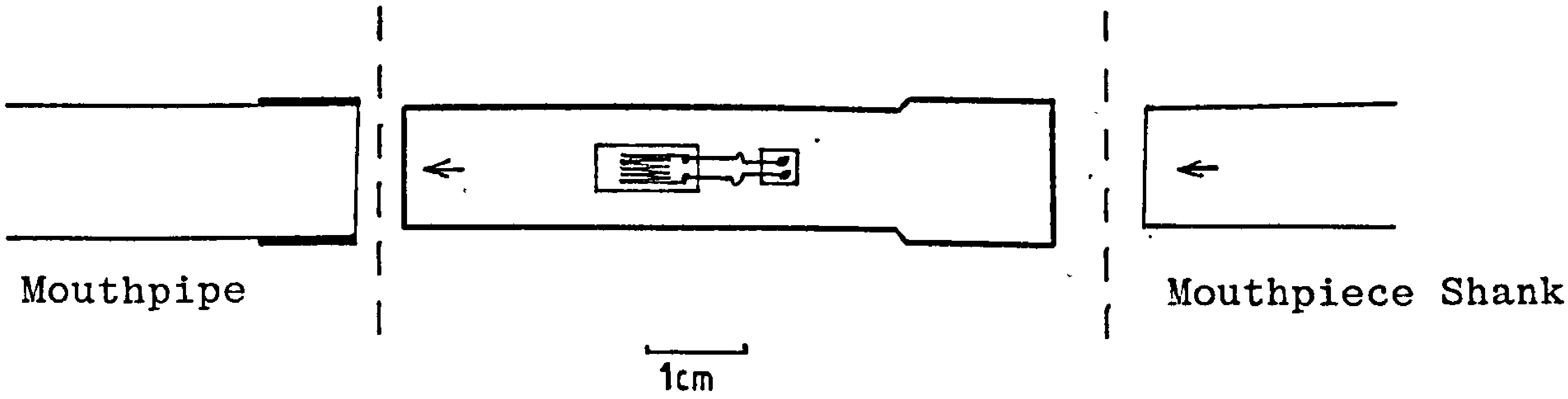
It was decided not to mount the gauges on the trumpet or mouthpiece itself for two reasons. Firstly, being at the early stage of this novel transducer application it was considered unwise to risk employing expensive musical instruments before the system was tried and tested. Secondly, it was preferable to allow each subject to use his or her own trumpet and mouthpiece to which they were accustomed.

The best compromise seemed to be to have a cylindrical connector which could be placed between the mouthpiece and mouthpipe.

TRANSDUCER NO 1: Feasibility Study

The first step was to carry out a feasibility study to determine a suitable combination of experimental materials and structural dimensions, as there was no precedent to follow in this case. A cylindrical piece of tubing in brass was constructed, whose male end fitted the tapered female mouthpipe of a standard Bb trumpet and accommodated at the female end a standard trumpet mouthpiece which is also tapered. (see figure 2.4)

Figure 2.4 The first transducer



With the connector in place, the pitch of the trumpet was lowered. This was compensated for by adjusting the tuning slide. The maximum compensation determined the length available for the transducer tube.

Two 120 Ohm copper nickel gauges 5mm long with polyester backing (Type N11-FA-5-120-11) were cemented longitudinally on opposite sides of the tubing. The output of the bridge was conditioned by a strain gauge bridge amplifier. A chart recorder then processed the signals. This initial arrangement of the apparatus showed good sensitivity to force on the mouthpiece. It was possible to select suitable amplifier gains for the resultant bridge supply potentials to produce signals of a magnitude convenient for the pen displacements of the chart recorder.

However this initial study revealed a particularly acute problem with temperature fluctuations. The signals from the gauges are very sensitive to temperature. This makes it impossible to give a baseline reading from which meaningful forces can be measured. Two sources of temperature variability can be identified. Firstly the ambient temperature is not constant. Secondly, blowing through the instrument appreciably raises the temperature of the brass. Even if it were possible to control the ambient temperature, the latter effect would completely confound the true strain in the bridge output.

There are two effects of temperature variation, both of which have to be eliminated. Firstly there is the differential thermal expansion between the grid conductor and the test part of substrate material to which the gauge is bonded. With temperature change the substrate expands or contracts, and since the gauge is firmly bonded to the substrate, the gauge grid is forced to undergo the same expansion or contraction. To the extent that the thermal expansion coefficient of the grid differs from that of the substrate, the grid is mechanically strained in conforming to the effects of the free expansion or contraction of the substrate. The resultant resistance change appears to the strain indicator as strain in the substrate. The second source of apparent strain derives from the electrical resistivity of the grid conductor which itself is temperature dependent, over and above the differential thermal expansion factor. Although these may seem to be small influences in themselves, they assume considerable importance because the change in resistance with actual stress or strain itself is very small. It was decided to attempt to remove the spurious influence of temperature by designing a second transducer.

TRANSDUCER NO 2

Firstly 'self temperature compensation' was attempted. This meant using gauges which had been manufactured with a thermal coefficient of expansion equivalent to that of brass, the material of the substrate to which they were to be bonded. Secondly the gauges used (Constatan alloy) had a low and controllable temperature coefficient which further minimised the error. However there were still several sources of error which made the output from any one gauge susceptible to extraneous influences. Usually the best way to minimise such effects is by using two or four active arm bridges in a Wheatstone Bridge configuration. (Cook and Rabinowicz 1963). In this way, the temperature induced variations in the bridge arm resistance will tend to cancel each other out.

Because the tubing had to have a certain minimum thickness to withstand the forces applied to it and there were stringent accuracy requirements, the sensitivity of the device was increased by using four 90^{o} rosettes equispaced around the tubing, with four single gauges all wired into a 1 x 8 gauge configuration. The gauges are connected into the bridge circuitry in such a manner as to make use of the Poisson's ratio (i.e. the ratio between the relative expansion in the direction of the force applied and the relative contraction perpendicular to the force); this increases the effective gauge factor and hence the sensitivity. (See figure 2.5, a & b overleaf).

Because of the small transducer signals involved, it was necessary to use a bridge amplifier with very low noise and drift characteristics. The output from the gauges was therefore conditioned by special strain gauge bridge amplifiers designed and built at the University of Strathclyde, and incorporating an SGA 100 Amplifier (C I L Electronics; Worthing). The strain gauges were attached to the metal using a heat curing epoxy adhesive (M - Bond 610) and were coated with Gaugecoat G to give impact and humidity protection. Both the adhesive and the gauge coating were manufactured by Micro Measurement, (Welwyn Strain Measurement, Basingstoke). The temperature compensation mechanism was

Figure 2.5a Photograph of the second transducer.

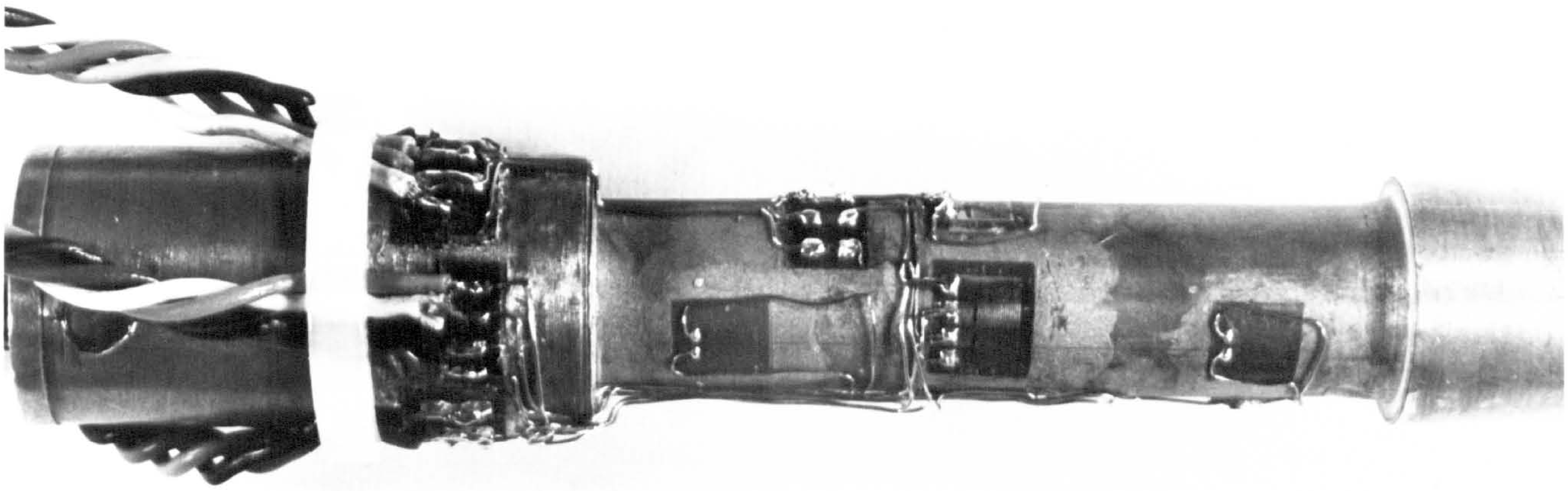
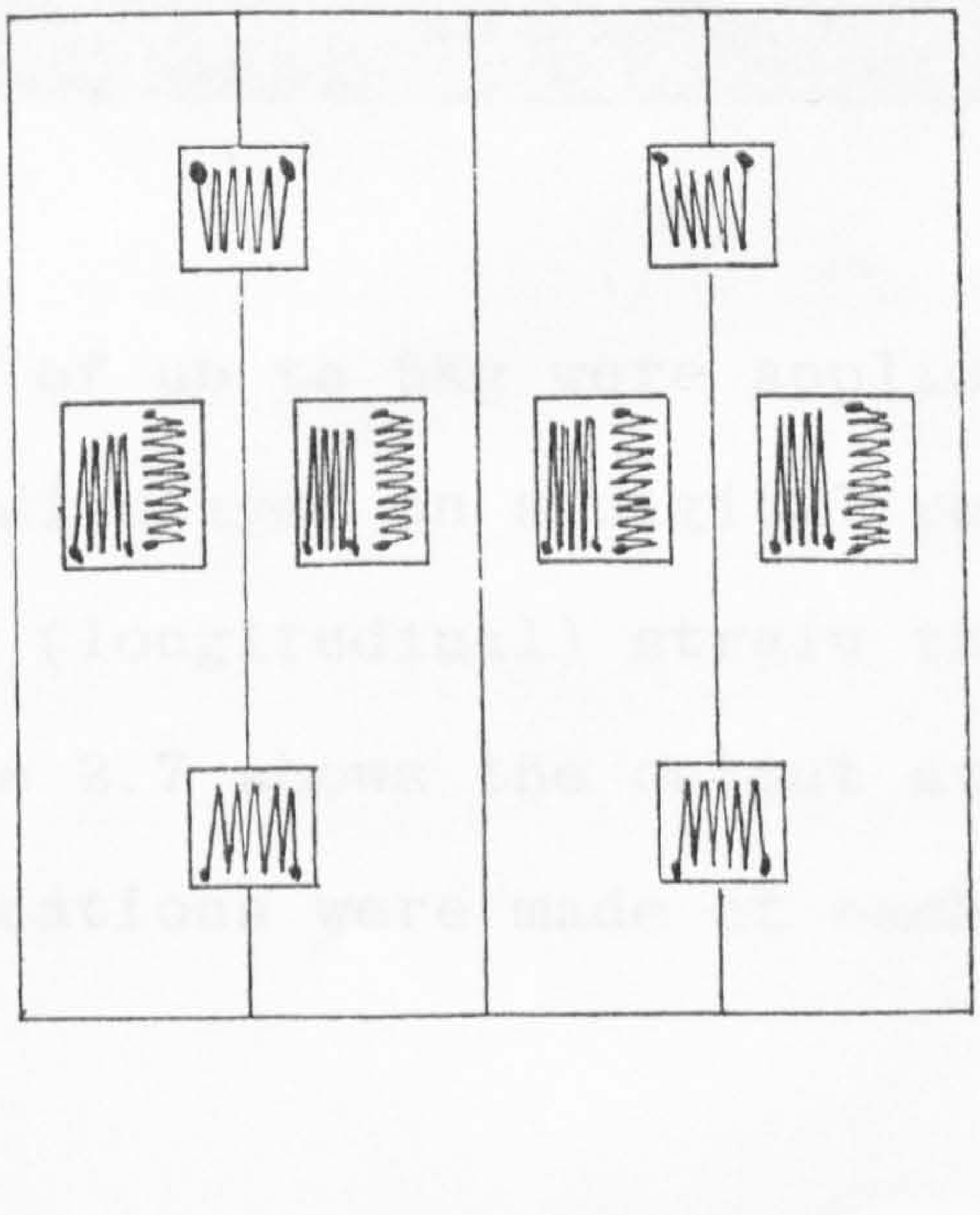


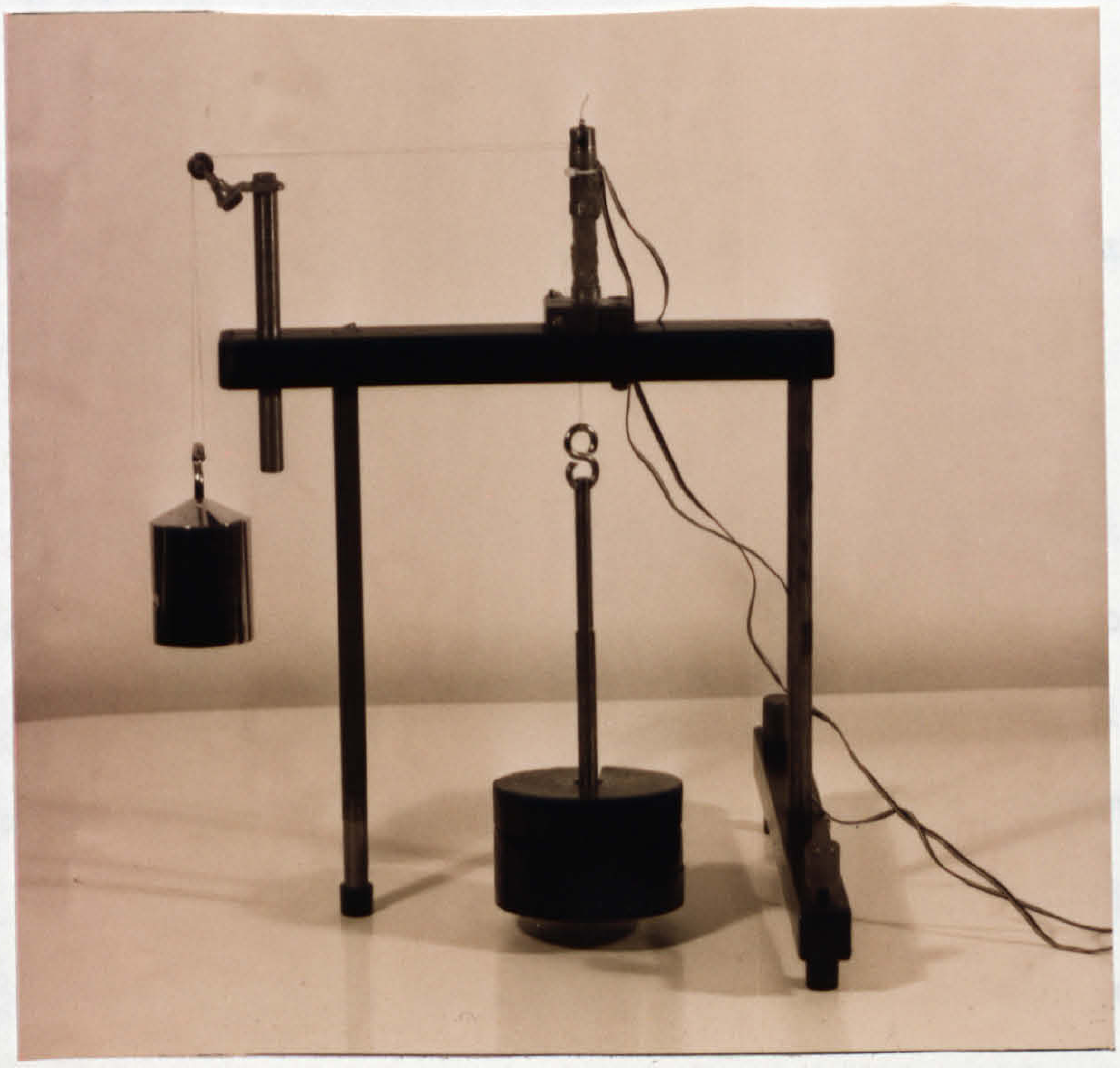
Figure 2.5b Exploded layout of the 1 x 8 gauge configuration of the second transducer.



found to be very effective, with drift tests showing the output from the gauge assembly to be very stable. This stability was also partially attributable to the elimination of noise by the amplifier.

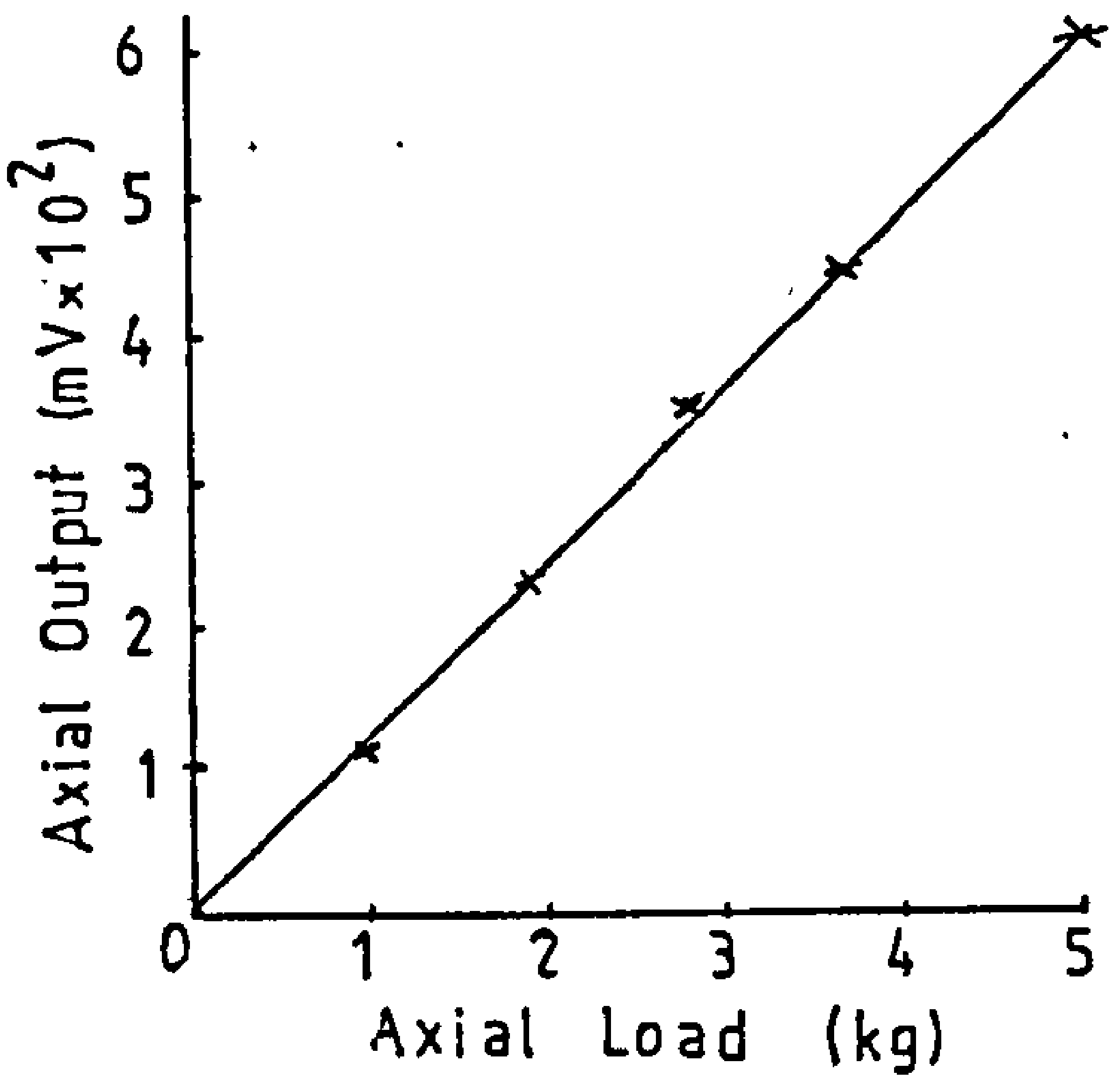
Calibration was accomplished by applying known loads to the transducer by means of a specially constructed jig, (see figure 2.6).

Figure 2.6 Calibration jig



Loads of up to 5kg were applied, and the channel outputs were displayed on a digital voltmeter. When calibrated for axial (longitudinal) strain the output was clearly linear. Figure 2.7 shows the output at 6 levels of load. Two applications were made at each level and the order of testing was randomised. Each point on the graph represents the arithmetic mean of the voltage output at each load level.

Figure 2.7 Output of transducer in response to axial load application

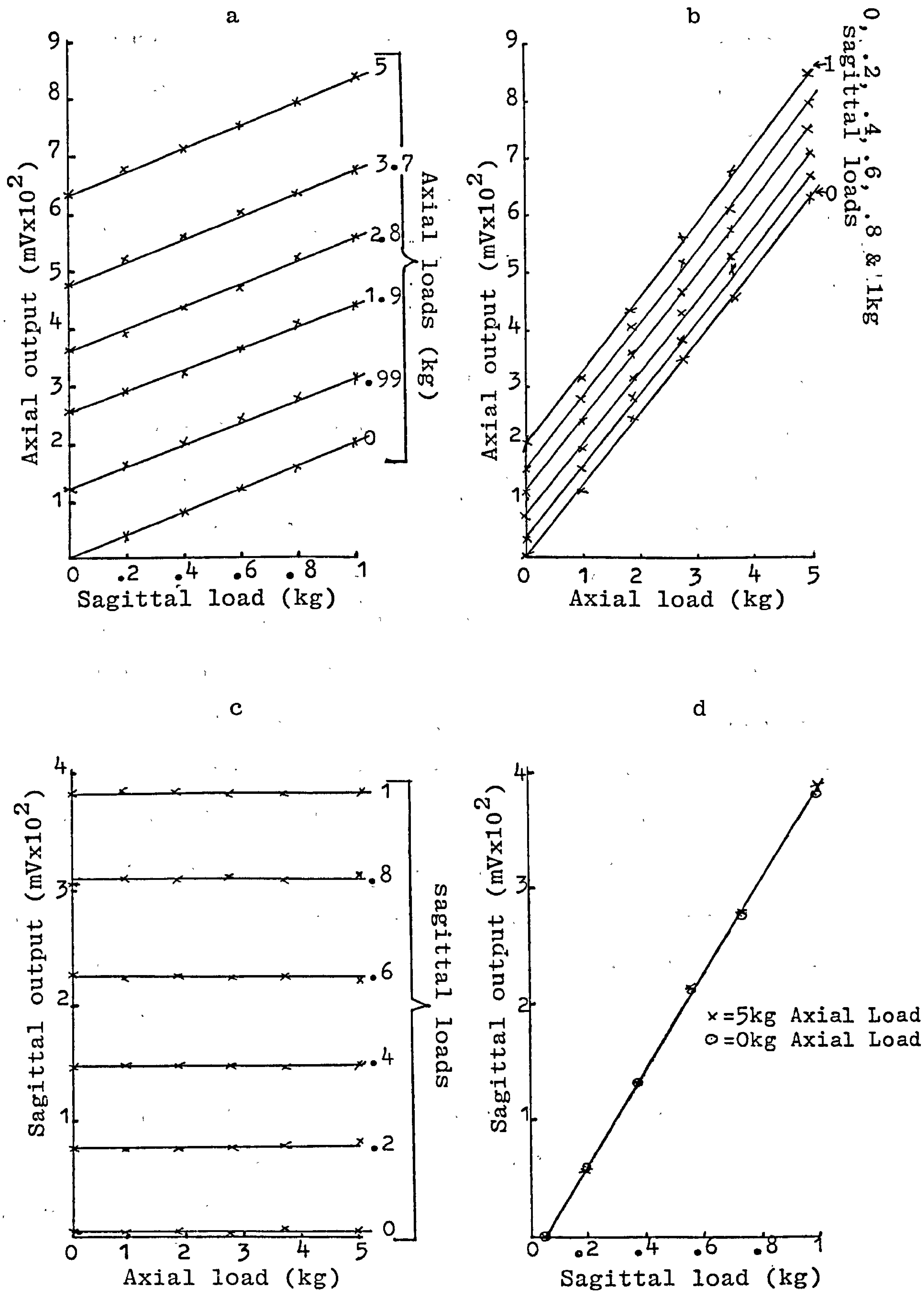


Unfortunately, the problem of cross effects arose which complicated the course of the experiment.

Ideally, the longitudinal strain channel should exclusively process the axial strain. The positioning of the gauges ought to achieve this. However, it is possible for strains in other dimensions to show apparent loads in this channel. After testing many of the subjects with the above apparatus it became apparent that the axial output channel had become extremely cross sensitive to forces acting laterally at right angles to it. This may have resulted from the excessively large forces which some subjects had occasionally applied to the transducer (over 10kg axially) As each gauge can only be extended to a small percentage of its length without permanently changing its elastic properties, it is possible that the thinness of the walls of the tubing had allowed the exceeding of the elastic strain limits of the gauges upon application of the largest forces.

In order to measure axial strain, it became necessary to look in detail at the cross effects. The predominant non-axial lateral force is that acting in the sagittal plane. (See figure 2.10). To measure the influence of such a force on the axial output, the calibration jig was adapted to allow the simultaneous application of axial and sagittal loads. Figures 2.8 and 2.9 show the influence of these loads on the output of the different channels. Six levels of axial load were used, (0, 0.992, 1.89, 2.80, 3.70, 5kg), and there were 6 levels of sagittal load, (0, 0.2, 0.4, 0.6, 0.8, 1kg). At each combination of axial and sagittal load four readings were taken, two for lateral and two for axial output. The order of application of loads was randomised to take any drift effects into consideration. The points on graph 2.8 a - d represent the arithmetic mean of the readings at each load combination. Figure 2.8a shows how, for each level of axial load, the sagittal load influences the axial output. It can be seen that a one kilogram load acting sagittally displaces the axial output by approximately + 200 mV at each level of axial load. The same displacement can be seen plotted a different way in figure 2.8b Figures 2.8c and 2.8d show how on the other hand the sagittal output is conversely

Figure 2.8 The effects of axial and sagittal loadings on axial and sagittal output. The second transducer.



much less influenced by axial load.

The problem of cross sensitivity, which had complicated the instrumentation of the project demonstrated the possible importance of the sagittal force variable in trumpet performance itself, and not just as a spurious factor in the measurement process. Firstly however, it was necessary to validate a new calibration procedure for the second transducer.

Looking at figure 2.8a and 2.8b it seems that, within the range tested, the increment in axial output attributable to sagittal loads is fairly constant over the axial load range and proportional to the sagittal load. For example at OKg, 0.992kg, 1.89kg, 2.80kg, 3.7kg, and 5kg axial loads, a 1kg sagittal load will add a 200 mV displacement. At these axial loads a 0.5kg sagittal load would add a 100 mV displacement. As long as the axial load does not conversely interfere with the sagittal output itself (as it seems not to from figure 2.8c and 2.8d), it would theoretically be possible to determine the true axial load from any axial output by subtracting the known cross acting sagittal component. This is only possible if both axial and sagittal channels are monitored simultaneously.

To verify that such a procedure is permissible, it was necessary to perform a statistical analysis on the Bridge Output data.

Table 2.1 The effect of axial and sagittal loads on axial output. The second transducer.

Source	S.O.S.	Df	Mean Squares	F Ratio	Sig
Sagittal Load S	343913.0	5	68782.6	18074.2	< .01
Axial Load A	3281425.16	5	656285.0	172479.7	<.01
Interaction A x S	79.34	25	3.17	.83	N.S.
Within	137.00	36	3.80		
Total	3625554.50	71	51064.10		

Table 2.2 The effect of axial and sagittal loads on sagittal output. The second transducer.

Source	S.O.S.	Df	Mean Squares	F Ratio	Sig
Sagittal Load	310711.5	5	6214.2	5345.6	<.01
Axial Load	156.07	5	31.2	2.68	N.S.
Interaction A x S	677.8	25	27.1	2.33	N.S.
Within	418.5	36	11.6		
Total	311963.9	71	4393.0		

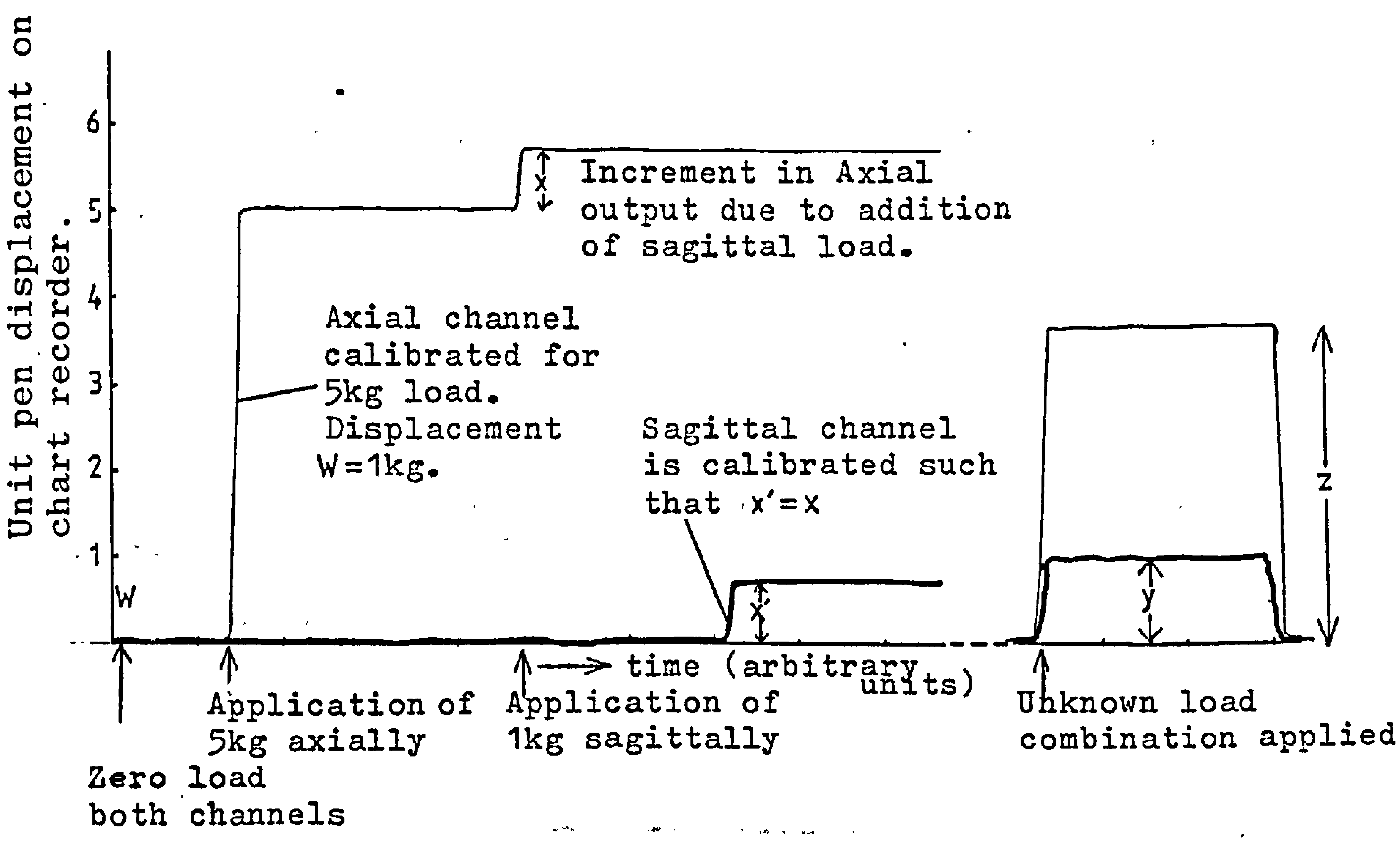
A two-way analysis of variance (Winer 1962) was carried out on both the axial and sagittal outputs, (see tables 2.1 and 2.2 respectively). The fact that the axial output is determined by both the axial and sagittal load is reflected in the very high F Ratios for these factors. (See table 2.1). We are really interested in the interaction term however which turns out to be not significant, (see table 2.1). The absence of cross sensitivity in the sagittal channel is shown by table 2.2. The output here is uninfluenced by axial or interactive terms but is highly significantly influenced by the corresponding sagittal load. This is important because it means that the sagittal output is a pure measure of sagittal load. The existence of these conditions enable the transducer to be calibrated in the following way to provide a valid measure of axial load.

A 5kg load is suspended axially from the transducer. A convenient unit/load displacement is selected by adjusting the sensitivity on the chart recorder. (See figure 2.9). Then a 1kg load is added sagittally. This results in a given increment to the axial output (x). Next the sagittal channel is itself calibrated to give a 1kg unit displacement equal to the aforesaid increment in the axial output. Thus, the true axial load can be calculated for any combination of loads which go to produce a given axial output, by simply subtracting the sagittal output from the corresponding axial output. For example after calibration the axial load in the case of the load combination applied at K is calculated as:

$$\text{Axial Load} = \frac{z - y}{w} \text{ Kg}$$

and the sagittal load is simply $\frac{y}{x}$ Kg

Figure 2.9 Calibration procedure



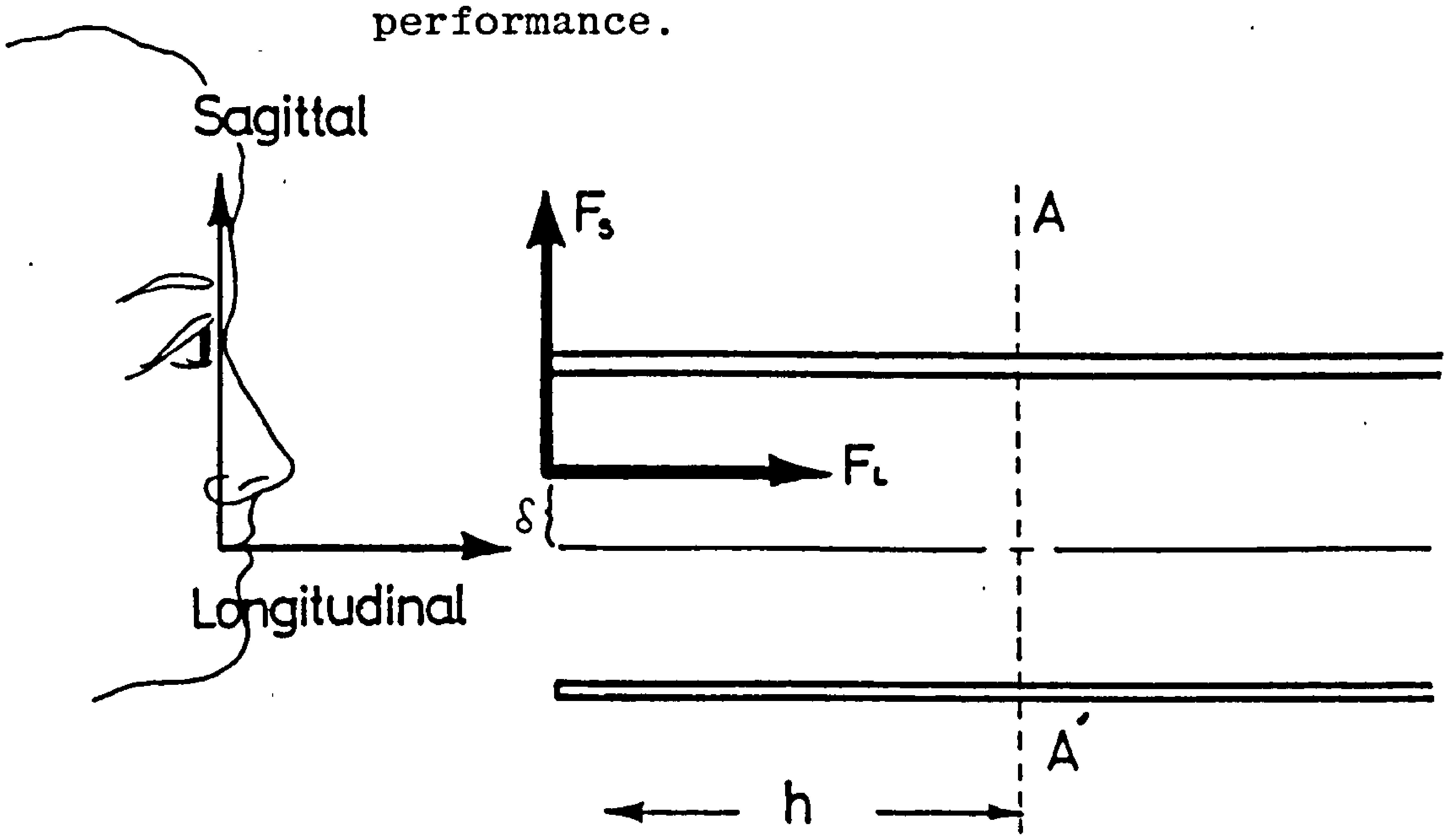
Employing this procedure it was possible to make use of this transducer in testing many subjects. However, the inherent structural deficiencies of the transducer had raised fears that further extreme deformations might eventually introduce unnoticed some interactive cross effects. Moreover an intermittent electrical fault had developed in the gauges. It was therefore decided to build a final transducer incorporating the design modifications which had been suggested by experience with the previous models.

TRANSDUCER NO. 3

We are now in a position to summarise the final theoretical position which has been reached regarding the possible forces applied to the mouthpiece in trumpet performance. This is as follows.

Three forces and moments can be applied to the mouthpiece, (see figure 2.10). The largest will be that being applied longitudinally, (parallel to the axis of the mouthpiece).

Figure 2.10 Forces about the mouthpiece in trumpet performance.



However, the mouthpiece is not necessarily played at a right angle to the lips and there may be a force acting perpendicular to the axis of the mouthpiece in the sagittal (vertical) plane. Although there is also the possibility of the presence of another lateral force acting transversely (sideways), this is likely to be negligible. The upper and lower lips may apply unequal forces to the mouthpiece, and the point of application of the longitudinal force resultant will be displaced in a sagittal direction from the axis of the mouthpiece, producing an additional bending moment. There is also the possibility of a sideways displacement of the axial force producing an additional lateral moment; this however is also likely to be negligible. Finally it is possible but highly unlikely that there may be a moment about axis of the mouthpiece.

This theoretical model was applied in the design of the final transducer. Where a transducer is symmetrical about the lateral (transverse and sagittal) planes through the longitudinal axis, the total strain at two symmetrical points on the upper and lower surfaces A and A' (see figure 2.10), will have four components. There will be strains due to the longitudinal (F_L) and sagittal (F_S) forces as well as the moment due to the sagittal displacement (δ) of the longitudinal force. There will also be a thermal strain due to the temperature changes in the transducer during playing (e_T). The strains at A (e_A) and A' ($e_{A'}$) are equal to the sum of these four strains.

$$e_A = C_1 F_L + C_2 h F_S + C_3 \delta F_L + e_T \quad \ldots\ldots\ \text{Equation 2.1}$$

$$e_{A'} = C_1 F_L - C_2 h F_S - C_3 \delta F_L + e_T \quad \ldots\ldots\ \text{Equation 2.2}$$

C_1 C_2 and C_3 are constants which define both the substance from which the device is made and the geometry of the transducer and the loads acting on it. By connecting the strain gauges in a suitable Wheatstone Bridge configuration, an output O can be obtained that is proportional to the strains $e_{A'} + e_A$. By adding the equations above, the sagittal displacement terms should cancel and leave, in theory, the output sensitive only to the axial loads and thermal strains. The temperature sensitivity as described before is eliminated by using two pairs of gauges, one pair each at A and A'. One of the pair is aligned longitudinally and the other at right angles to it. Figure 2.11 shows the transducer. This is a departure from the last

Figure 2.11 The third transducer.

design however in that there are fewer gauges. This is because of a change in design of the connector tube itself. The transducer had to be sensitive enough to the loads applied without being susceptible to the excessive structural deformation which damaged the previous transducer. The middle section of the new tubing therefore consisted of two symmetrically opposite flat surfaces which were thin enough to undergo measurable deformities. The opposing rounded surfaces however were much thicker. This had the dual function of protecting the gauges from excessive strain, and minimising the effect that any transverse (sideways) force being used might have had by way of cross sensitivity with the longitudinal and sagittal channel outputs. The connector was machined as one component from a brass bar with the inside of the tubing cylindrical and narrow; this reduced the internal volume, and produced the least alteration in playing characteristics. There were knurled ridges at both ends. These were essential to ensure that the tubing could be inserted and removed without difficulty from the many different trumpets and mouthpieces used, without having to grip the gauged surfaces themselves. Such manipulation might have harmed the gauges underneath the protective gauge coating. As the forces used on the mouthpieces were occasionally very large and the connecting tapered parts can get stuck, this small design modification saved much potential embarrassment.

An initial calibration of the transducer was performed.

The jig was again used to simultaneously apply axial and sagittal forces. Six axial load levels (0 - 5kg) and five sagittal load levels (0 - 2kg) were used. (The design of this transducer with its stronger walls permitted the application of larger sagittal calibration loads). Each load combination was applied twice in random order, and the output of both load cell amplifiers measured using a digital voltmeter. Figure 2.12 shows the results of these tests. Each point represents the arithmetic average of the particular load combination. The outputs are clearly all a linear function of the loads. Despite the careful design and construction of the latest transducer, there is still some cross sensitivity between the bridges. This is probably because it is extremely difficult to manufacture transducers of perfect symmetry, and almost impossible to align and secure them in exactly the desired position. There may also still be some elastic deformation in load application.

The sagittal load in this case adds an increment of around 5mV to the axial output per kilogram of sagittal load. As can be seen from figures 2.12a and 2.12b the effect is much less than with the second transducer but not negligible. Inspection of figures 2.12c and 2.12d show that the sagittal output itself appears not to demonstrate cross sensitivity. A two-way analysis of variance was carried out on the data (see tables 2.3 and 2.4). Table 2.3 confirms that the longitudinal output responds to both axial and sagittal

Figure 2.12 The effects of axial and sagittal loadings on axial and sagittal output. The third transducer.

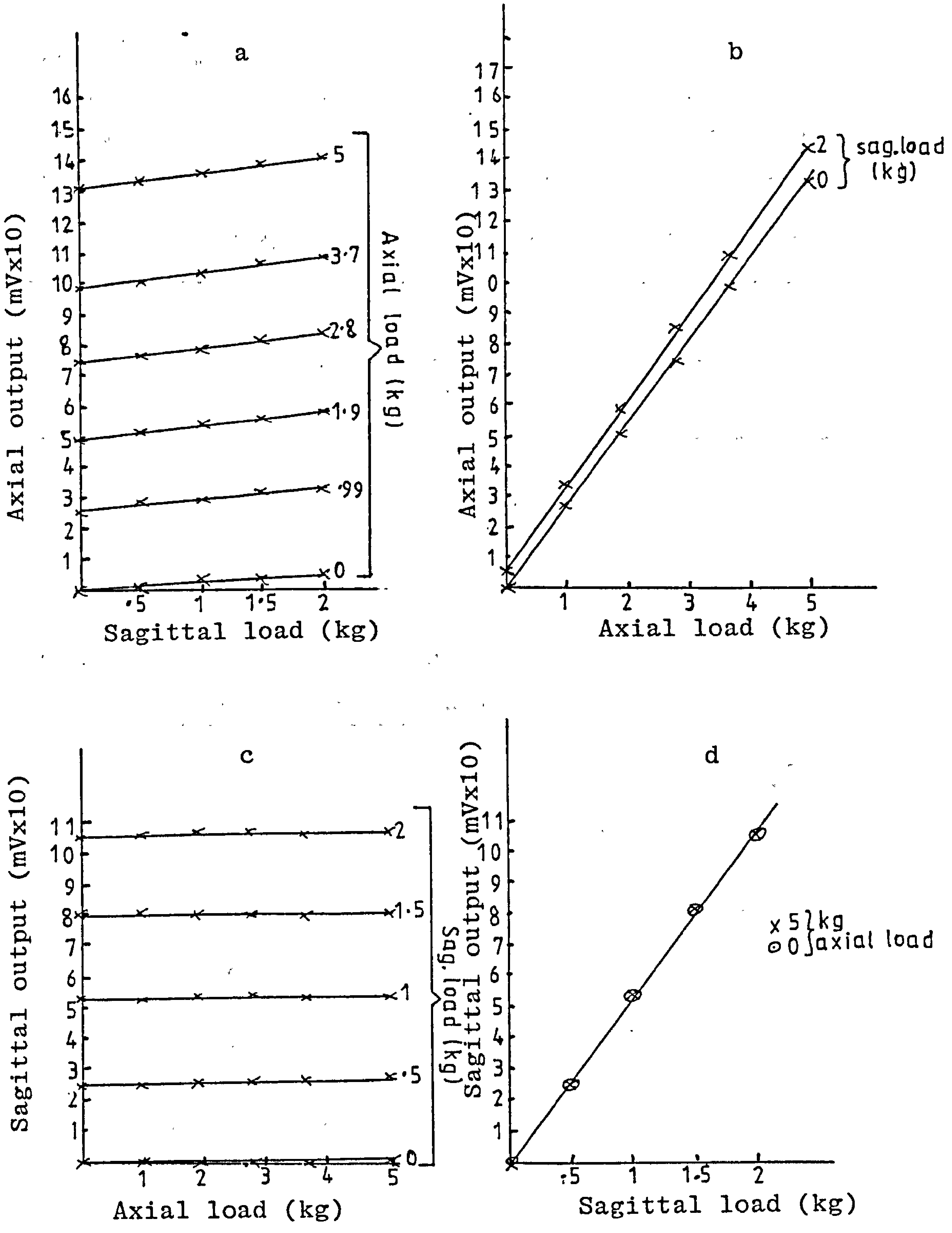


Table 2.3 Axial bridge output. Third transducer

Source	S.O.S.	Df	Mean Squares	F Ratio	Sig
Axial Load	118653.8	5	23730.76	10062.50	<.01
Sagittal Load	538.9	4	134.74	57.13	<.01
Interaction A x L	73.92	20	3.69	1.56	N.S.
Within	70.75	30	2.35		
Total	119337.50	59	2022.66		

Table 2.4 Sagittal bridge output. Third transducer

Source	S.O.S.	Df	Mean Squares	F Ratio	Sig
Axial Load	18.05	5	3.61	0.26	N.S.
Sagittal Load	332301.50	4	83075.38	6205.40	<.01
Interaction A x L	126.26	20	6.31	.47	N.S.
Within	401.62	30	13.38		
Total	332847.50	59	5641.48		

loads. Table 2.4 indicates that there is no demonstrable effect on sagittal output of axial loading. The outputs of the bridges can be described by the equations:

$$(O_L)_{ij} = (F_L)_i + (F_S)_j + (F_LF_S)_{ij} + z_{ijk} \ldots \text{ equation 2.3}$$

$$(O_S)_{ij} = (F_S)_j + (F_L)_i + (F_SF_L)_{ij} + z_{ijk} \ldots \text{ equation 2.4}$$

Equations 2.3 and 2.4 represent the dependence of the longitudinal $(O_L)_{ij}$ and sagittal $(O_S)_{ij}$ outputs on longitudinal (i) and sagittal (j) load levels. The first terms on the right hand side of the equations are the true loads. The second terms are the cross effects, the third terms are the interactions. z_{ijk} is the residual for the k th replication. The crucial values in tables 2.3 and 2.4 are therefore the interaction terms. It can be seen that these interaction terms were non-significant for both the longitudinal and sagittal outputs. So, in order to arrive at a measure of axial load which is not contaminated by the sagittal load it is necessary to <u>measure</u> the sagittal load. With two gauges (at A and A', see figure 2.10) aligned to measure the sagittal strains and connected in the neighbouring arms of a Wheatstone Bridge, a bridge output may be obtained which is proportional to the <u>difference</u> of the strains e_A and $e_{A'}$ (see equations 2.1 and 2.2), that is:

$$C_2hF_S + C_3\delta F_L$$

In order to fully define the bending moment, it would be necessary to measure the stress at two points with different values of h. However, to *compensate* for cross effects the measure need only be taken at a single value of h, at a site close to the gauges used to measure longitudinal strains. This does not mean the information from the sagittal output is without use. This output nevertheless provides a *relative* measure of sagittal strain, if it can be assumed that the distances from the gauges on the tubing to the end of the different trumpet mouthpieces used by the subjects are not too dissimilar. Fortunately, trumpet mouthpieces have to be very similar in length to be compatible in pitch with standard Bb trumpets. Hence although it is not possible to give the sagittal strains in absolute force units, it is valid to compare this factor in a relative way.

During use the sensitivity of the bridge amplifiers was set so that an increment in the sagittal load produced equal output changes for each bridge. The relationship between the output and the loads could be written as the matrix equation:

$$\begin{bmatrix} O_L \\ O_S \end{bmatrix} \begin{bmatrix} C_{LL} & C_{LS} \\ C_{SL} & C_{SS} \end{bmatrix} \begin{bmatrix} F_L \\ F_S \end{bmatrix}$$

where C_{LS} are due to cross effects. The value of the four elements in the matrix were calculated using a multiple regression analysis. (Winer 1962).

Table 2.5 Multiple regression on an AXIAL output. Third transducer.

Source	Df	S.O.S.	Variance	F Ratio
Regression	2	118746.50	59373.23	4797.62
Remainder	57	705.40	12.37	
Total	59	119451.90		

Least square coefficients 4.13; Sagittal load:confidence band .642 (95%)
26.60; Axial load:confidence band .271 (95%).

Table 2.6 Multiple regression on a SAGITTAL output. Third transducer.

Source	Df	S.O.S.	Variance	F Ratio
Regression	2	252409.40	126204.70	43.57
Remainder	57	165098.70	2896.47	
Total	59	417508.10		

Least square coefficients 91.3; Sagittal load:confidence band 9.8 (95%)
3.67;Axial load:confidence band 4.17 (95%)

This gives the best calibration matrix values with the load cell amplifiers at maximum sensitivity as:

$$\begin{bmatrix} 266 & 41 \\ 37 & 913 \end{bmatrix} \text{mVn}^{-1}$$

Reducing the longitudinal load factor to unity the values become:

$$\begin{bmatrix} O_L \\ O_S \end{bmatrix} = 266 \begin{bmatrix} 1 & 0.154 \\ 0 & 0.154 \end{bmatrix} \begin{bmatrix} F_L \\ F_S \end{bmatrix}$$

where the output is in mv and the force in N. A sagittal load of 1N was equivalent to a bending moment of 61 mNm.

CHAPTER THREE

FORCES ON THE EMBOUCHURE DURING TRUMPET PERFORMANCE

Introduction

The raison d'etre of the second experimental study is largely heuristic. It is an explanatory investigation of some of the most important variables which influence mouthpiece force usage, and the relations and interactions among those variables.

At the moment it is not possible to give a valid general description of the way mouthpiece force is used by trumpet players, let along make any teaching recommendations regarding its use. The principal reason for this lack of real knowledge is that there is a shortage of empirically based research in this area. Despite exhaustive literature searches, the author has been able to uncover only three published experimental studies providing truly scientific information on this issue. Furthermore, these earlier studies are extensively flawed by weaknesses in methodology which make for interpretative complications.

Common to all the studies, for instance, is the rudimentary nature of the instrumentation employed. The implications for experimentation utilizing such restrictive equipment are fully discussed in the previous chapter, (see chapter 2). Equally detrimental are deficiencies in the control of several important variables within the experimental design. The latter problem will be discussed in detail now in conjunction with a review of the major findings from the three reported studies.

Previous Studies

Following a chronological structure we will first consider the earliest study, that by Henderson (1942). He provided kymographic representations of the mouthpiece force used in playing a group of notes in the harmonic series on C. On a purely descriptive (i.e. non inferential) statistical level, the results were held to show that mouthpiece force increased with ascending pitch. With no control for intensity however, and data from only three subjects, this conclusion is not to be accepted without some reservations. For the same reasons, Henderson's finding of an association between player proficiency and the range of mouthpiece force usage observed with given pitch changes must be seriously questioned.

Weast and Hake's (1965) study also involved the purely descriptive treatment of experimental data but has the merit of incorporating some control of intensity, and there were more subjects (30 in all). Their results confirmed Henderson's finding that mouthpiece force varies positively with pitch. Also, mouthpiece force was found to vary directly with intensity. At given pitch levels, louder notes were associated with higher forces. No attempt was made to quantify the magnitude of these effects.

Weast and Hake found no support for Henderson's contention that better players have smaller ranges of force usage over the given range of pitch levels tested. The data appeared to show, though, that at the fixed pitch level for the analysis (G_5), there was a difference in the range of mouthpiece force used for the experimental intensity levels ("piano" and "forte") between the two proficiency levels. The so-called "advanced" group showed a smaller range of mouthpiece force variation for the given intensity change than the "beginner" group. It is difficult to gauge the general validity of this finding. The criteria actually used to assign the subjects to the respective proficiency groups are not reported; nor are stated the absolute intensity levels that were used. These shortcomings, together with a lack of control over important extraneous independent variables makes it hard to interpret the real meaning of this postulated interaction.

The latest study to date was conducted by White (1972). This investigation entailed a more rigorous test of the effects of pitch (3 levels) and intensity (3 levels; 85, 95 & 105 dB) on the application of mouthpiece force. Only 18 subjects were tested however. External validity was improved with a wider variety of test items. There were 54 different musical elements in the test. With mouthpiece force as the dependent variable, statistical tests revealed main effects for pitch and intensity ($p < .005$ in both cases). No attempt however was made to quantify the strength of these associations, or to explore the interaction between pitch and intensity.

As for the variable of performer ability, White cites two pieces of evidence against the existence of any relationship between mouthpiece force usage and performer proficiency standard. Firstly, White computed a Spearman correlation of $r_s = .06 (t=1.6; p > .4)$ between rank order proficiency and overall use of mouthpiece force. This was achieved by calculating an index representing the total force used throughout all 54 test components by each subject, and correlating this with a rank ordering of the subjects according to a proficiency index which was a composite of three selected criterion measures. Secondly, White found no difference between the "advanced" group and the "beginner" group with respect to the change of mouthpiece force for the given experimental ranges of register or intensity.

Although the statistical treatment of the data in this study lends more confidence to its conclusions, there are still several methodological weaknesses which the design of the present study aims to overcome.

To begin with, the operational definitions of the "advanced" and "beginner" groups for the two levels of the proficiency factor are too ad hoc. Every subject is scored on each of three criteria: (1) Number of years of private study; (2) Years played; (3) Age. These yield an aggregate score or single index for each subject. The subjects were then divided into two groups for analysis, the advanced group consisting of subjects ranked 1 - 9 on this score, and the beginning group being composed of those subjects ranked 10 through to 18.

This arrangement entails a crucial sampling inadequacy which derives from two identifiable sources. Firstly, the classification advanced/beginner on the basis of the three stated criteria biases the proficiency measure towards experience in terms of the amount of time spent playing the instrument rather than skill per se. Although experience in itself may be of some interest in the development of mouthpiece force usage, it has little to do with proficiency in the sense of skilled musical performance, where it is a sad fact that practice alone does not make perfect. Secondly, White used the expedient of incidental sampling by testing any conveniently available local players. This post hoc assignment of subjects to the criterion groups leads to the presence in both the beginner and advanced samples of what White himself refers to as "a fairly large percentage of subjects representing a medium level of performance". These circumstances would inevitably make is harder to demonstrate any difference in the use of mouthpiece force between proficiency levels.

Another problem with the White study is the possible loss of valuable information by overreduction of the data. Mouthpiece force is calculated as a gross measure by collapsing across one or more independent or assigned variables. For example, the criterion measure for determining differences between registers was each subject's mean score on two notes at three intensities for each register. The effect of extensive pooling of variances could be to mask any intrinsic differences between the proficiency groups, particularly of an interactive nature. The exact nature of the interaction between pitch and intensity is hereby also obscured.

The order of presentation of the test material is also a matter of some concern. White's preliminary tests apparently demonstrated "negligible order effects", and led to the order of presentation of test items being dictated by convenience alone. However, this decision was taken on the basis of very few subjects, and does not agree with the common belief amongst trumpet players that the amount of mouthpiece force used will increase over a playing session due to fatigue. The very question of fatigue effects deserves systematic investigation, and will form an integral part of the current study.

Another problem with White's design concerns the nature of the test material. White notes that only ten out of the eighteen players were able to play notes above high C (C_6), and simply omits data collected above this pitch from the analysis. As the ability to play in the extreme upper register is itself a mark of highly skilled trumpet performance, it seems unwise not to explore this upper region.

This completes the review of the three available studies on mouthpiece force. In sum, it can be seen that the various interprocedural differences, together with a failure of the three investigators to assess the impact of extraneous confounding influences have prevented the emergence of a compelling consensus view of the role of mouthpiece force application in trumpet performance.

We are now in a position to outline the main features of the present experiment which are directed towards enhancing the validity of this descriptive investigation of the variables influencing mouthpiece force.

INDEPENDENT VARIABLES

(1) Subject Variables

(i) Proficiency

Proficiency or expertise on the trumpet is an explanatory variable whose partitioning presents great difficulty. When attempting to specify nonmanipulated variables in research of an exploratory kind it is often impossible to avoid using ad hoc definitions. Unfortunately, experimenters have frequently ignored the logical and empirical consequences of employing classificatory systems which are "operationally convenient but conceptually questionable", (Hadden, 1969). The present study however aspired to compare groups of players differing by proficiency in a meaningful way. This was attempted in accordance with Kerlinger's (1973) Maxmincon principle of designing, planning and conducting research to maximise the systematic experimental variance. In this case it was achieved by having only two levels of the proficiency factor, and making these "high" and "medium" proficiency conditions as different as possible. There is an important proviso however. Mention has already been made of the problem of distinguishing between proficiency and experience. Fully professional performers will tend ipso facto to have had more experience than amateurs on the instrument. To compare professional players with amateurs on their use of mouthpiece force would entail the possibility of confounding skill with experience.

The overall requirement therefore was to find the best combination of minimum experience difference and maximum skill difference between two referent groups. This was achieved in the following way. The "high" proficiency group consisted exclusively of full-time professional performers from major British symphony orchestras, radio bands and recording studios. The "medium" proficiency group consisted of highly experienced amateur players of intermediate ability with a negligible professional performance history.

The stringent criteria for inclusion in the high proficiency group reduced the number of subjects available for this sample, although it was possible to test thirty of these players in all. This also restricted the study to male players only, as these represent the large majority of the profession. The difficulty of obtaining high proficiency subjects also meant that it was not feasible to test all these players on separate occasions. Most of the professional performers were tested when on tour and often available in Glasgow for only a few hours. The necessity to carry out all the main testing in a single session in turn introduced further methodological constraints.

(ii) Performer Style

In brass performance much is often made of the difference in playing techniques between "straight" i.e. classical players and their "dance" or "jazz" counterparts. It was decided to introduce performance style itself as a factor in the search for variables determining mouthpiece force. Although similar definitional problems exist with this categorisation as with that of proficiency, it seems to be the case that very few players are active in both classical and popular styles to any degree. This exclusivity made assignment of subjects to the criterion groups relatively unambiguous.

(2) Performance Variables

(i) Pitch

It was decided to explore initially the full normal gamut of the trumpet. Orthodox trumpet range is G_3 to C_6. As all the subjects could manage up to C_6 in most circumstances, there were few playing problems to present limitations to the accurate manipulation of this variable.

(ii) Intensity

The regulation of this parameter is much more problemmatic. A sound level meter is required to ensure that the subjects are playing at the stipulated intensity. A pilot study with this revealed that the dB range 87 to 103 (measured at 1m from the bell of the trumpet) approaches what is reasonably playable on the instrument, and this fixed the limits for the study of intensity.

(iii) Trumpet Techniques

The advantage with the present equipment is in the wide variety of performance material which can be used as test items. It was thus proposed to investigate whether the use of mouthpiece force varies with the performance of items representing different trumpet techniques.

There were 5 chief catagories of representative test material, covering a wide range of performance variables.

I Long Note Study

This data provided the basis for the primary analysis of the effects of pitch and intensity, and is directly comparable to previous research in this area.

II Diverse Musical Exercises

These consisted of: (1) Scales; (2) Arpeggios; (3) Lip Flexibilities. All exercises were to be performed both tongued and slurred.

III Tests Eliciting the Limits of Mouthpiece Force Usage

(i) A separate extreme upper register study.
There are three main justifications for a separate high pitch study. Firstly, it is physically extremely difficult to play in this register. It is here where the mouthpiece force used is likely to be at its greatest, and where the deleterious effects of such large forces will be most evident.

Secondly, Legge and Barber (1976) on skills research in general, make the point that it may only be by investigating the upper limits of skilled performance that results may be obtained which reveal aspects of the processes underlying that performance; and upper register playing is generally taken to be evidence of a high level of skill mastered by only a small percentage of players.

Thirdly, even if a player is able to perform in the extreme upper register, it is still inordinately difficult to play over the full 87 to 103 dB range of intensity. In order to avoid the confounding of pitch and intensity, a different design to study the extreme upper pitches had to be effected.

(ii) An ascending scale exercise.

(iii) Active application by the performer of maximum and minumum force.

IV An Endurance Test

This test was included to assess fatigue effects of an acute nature, elicited by performance of an extended exercise.

V Judgmental Tasks

This category of test material was devised to throw light on the phenomenology of the use of mouthpiece force.

DEPENDENT VARIABLE

The response variable of primary interest is axial mouthpiece force, i.e. that directed longitudinally down the mouthpiece. Minor attention will be paid to the data on sagittal forces, although these can only be measured in a relative way (see chapter two).

SEQUENCING EFFECTS

Although previous researchers did not consider order or carryover effects to be important, it was decided that this experiment would take the possibility of these into account. Indeed this study follows the recommendation of Cochran and Cox (1957) to protect all experiments from unusual events even when the experimenter expects no systematic bias from extraneous variables.

Given the relatively small number of subjects and the large amount of stimulus material, the extent of counterbalancing possibilities was extremely limited. Hence it was decided to test the five chief categories of experimental material in the same order for all subjects, while randomizing or couterbalancing the performance of the various items within each category.

STABILITY OF MOUTHPIECE FORCE

(1) Short Term Stability

The question of the extent of persistance over time of mouthpiece force in the short term, that is, within the experimental session, is inseperable from that of fatigue. As already mentioned this could be an important variable from the point of view of sequencing effects. The variable of fatigue was therefore deliberately built into the research design both to achieve control and give additional information about its effect as an independent variable per se in the determination of mouthpiece force usage.

(2) Long Term Stability

It was possible to obtain some longitudinal data by retesting some of the subjects at regular intervals over a six month period.

NATURE OF THE GENERAL DESIGN

The research design employed in this study would be described as quasi-experimental (Campbell 1968), mixing ex post facto with experimental modes of analysis and interpretation. The distinction between these types of research rests on the question of the direct manipulation of independent variables. In this study, many of the independent variables to be investigated are not amenable to direct control because they have already occurred or are inherently not manipulable. These attribute or measured variables are chosen "after the fact" (Badia & Runyon, 1982). In this study the trumpet player subjects are assigned to groups on the basis of their initial difference in proficiency and performer style, to determine if they respond differently in a common situation.

It is the particular nature of the chosen research problem which dictates the use of ex post facto methods with their concomitant problems. Subject proficiency is clearly not manipulable, nor is performer style. Without direct investigation of these factors however, it might be impossible to develop theories of mouthpiece force usage which could lead to useful recommendations for teaching, which is the overall aim of this treatise.

It is vital therefore to recognise the inherent limitations of ex post facto research. For example, the selection procedure for the dichotomous proficiency and performer style samples may result in the referent groups differentially possessing traits or characteristics extraneous to the research problem which make it difficult to attribute any discovered differences in the level of the dependent variable unequivocally to the specified attribute variable, (the problem of dissociating experience and skill has already been discussed). Indeed taking the sampling criticism to its extreme it could be objected that the whole collection of subjects in this experiment was drawn from a population of relatively advanced trumpet players.

As it was impossible to avoid non random selection and assignment of subjects, the study aimed to achieve maximum control of the extraneous influences commonly associated with research of an exploratory nature, and to take particular care to avoid unwarranted casual inferences drawn from such data.

HYPOTHESES

The few and discrepant findings from the available literature are unfortunately not conducive to a conceptualization of the research problems in terms of definite hypotheses to be tested. In any case Christenson (1977) points out that substantive hypotheses serve little purpose when engaged in exploratory work in a relatively new area where the important variables and their relationships are not known.

GENERAL METHOD

The experiment consists of several discrete studies. Each study employs different test material representing some of the most important techniques of brass performance. The specific rationale and analysis for each of these parts is different, and will thus be considered separately. However, the features common to the whole experiment will be dealt with first to avoid repetition and to show the overall experimental format of the study.

Subjects

The subjects were 60 white male trumpet players between the ages of 18 and 55. The "high proficiency" players were 30 full time professional trumpeters. This group was subdivided further into two smaller groups of 15 subjects each, representing the "performer style" subgroups. The "classical" group contained 15 leading symphonic players from the major orchestras, while the "popular" group consisted of 15 top jazz, pop and big band players.

The "medium" proficiency sample consisted of 30 trumpet players of intermediate performance standard, with a minimum of six years experience each on the instrument. Likewise, this group was divided into two subgroups of 15 subjects each according to the performer style dichotomy. The intermediate classical players were mainly from local amateur orchestras and brass bands while the intermediate "popular" group were largely local jazz and big band trumpeters.

This gives four overall subgroups with 15 players in each which are the basic units for the experimental analysis.

Because the rationale of the study necessitated the use of performers of an exceptionally high standard for the "high" proficiency group, it was only possible to test these subjects as and when they were on tour and performing in the locality, and could be persuaded to give up their time to take part in the experiment. No payments were offered for acting as subjects in the study, but most who were approached were keen to render their services if at all possible.

Apparatus

The subjects were tested in an acoustically isolated chamber; 2.0m high, 1.8m long & 1.9m wide (see photograph 3.1).

Photograph 3.1

Acoustically isolated chamber.

A monaural audiotape recording of the entire experimental session for each subject was made with a Uher M 516 unidirectional prefessional microphone, and a Revox A77 reel to reel tape recorder.

The transducer, custom strain gauge amplifier and potentiometric chart recorder used to record axial and sagittal mouthpiece forces have been described in chapter two. The two channels of the chart recorder gave out continuous force readings.

In order to allow the exact identification of the points in the test material to which these readings correspond, it was necessary to devise a metronome which was synchronised with the chart and which generated audible metronomic pulses for the subject to follow. The metronome also ensured that all the subjects would perform the test exercises at precisely the same tempo, which was an important methodological consideration.

To achieve this, a device was constructed which used a light source (12 Volt M.E.S. Bulb) and a photodetector, (O.R.P. light dependent resistor) from which were derived the trigger pulses for the metronome. These were fitted into the Bryans 2800 chart recorder such that the sprocket holes of the chart paper ran continuously between them, (see photograph 3.2 and figure 3.1).

Photograph 3.2

Chart recorder with light sensing metronome.

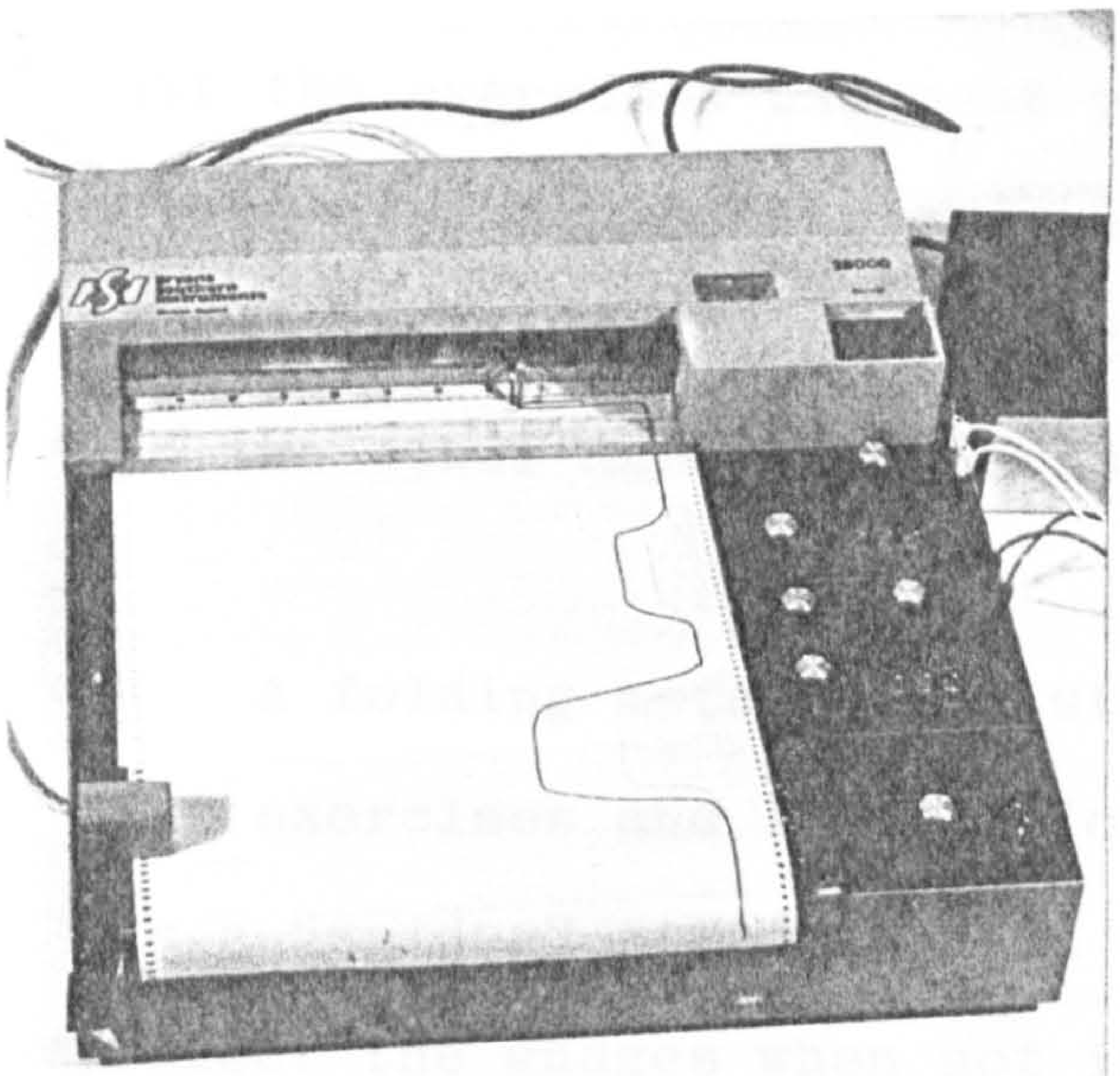
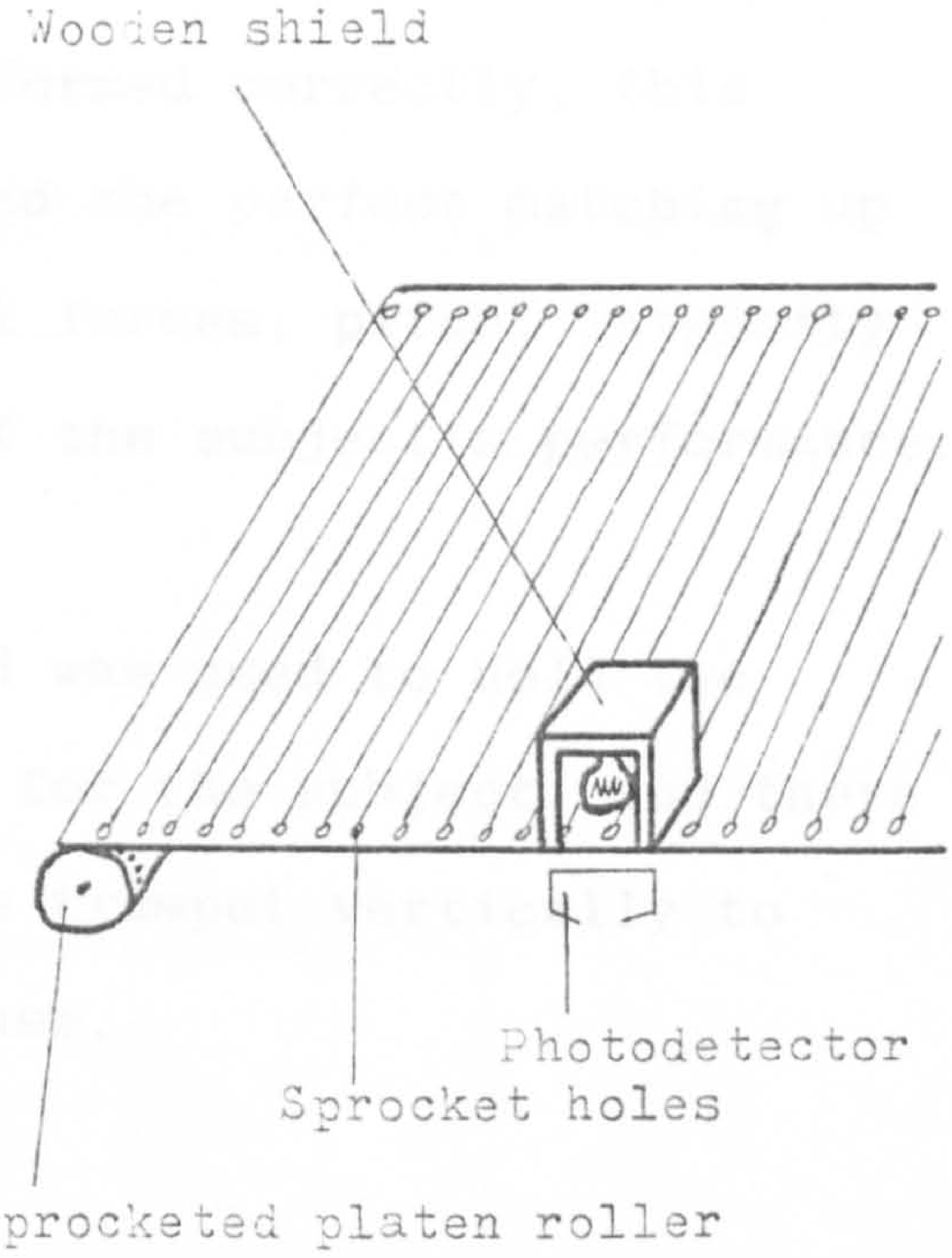


Figure 3.1

Schematic diagram of light sensing metronome.



As the chart paper was run at a constant speed of 5mm/sec, and the sprocket holes were exactly 5mm apart, it was possible to generate trigger pulses at a rate of precisely one per second. A pizoelectric transducer then converted the signal into an audible pulse which served as the tempo for the test exercises. A circuit diagram of this light sensing mechanism is given in the appendix (see appendix 3.1).

The sound intensity level was monitored by a Dawe decibelmeter (type 1400 F). The intensity was recorded by another chart recorder (a single channel Servoscribe type RE 5111). This was also synchronised with the chart recorder by means of pulses from the photoelectric device to give an exact timing of the attained volume.

With the aid of the reel to reel recorder to check that the exercises had been performed correctly, this method of synchronisation enabled the perfect matching up of the longitudinal and sagittal forces, pitch, intensity and the other musical aspects of the subject's performance.

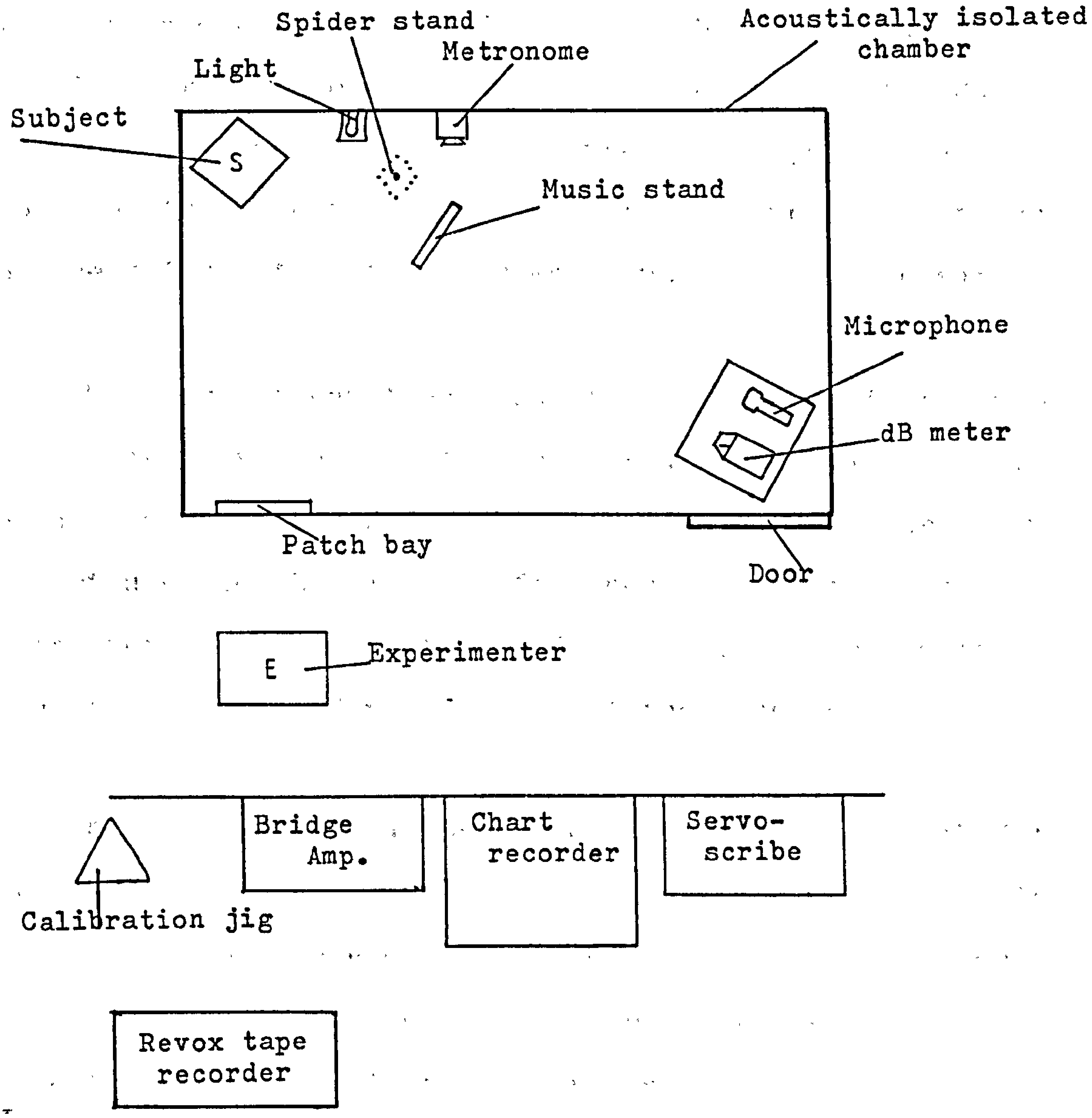
A folding metal music stand was used to hold the test exercises and instructions for the subject, and there was a "spider" stand to hold the trumpet vertically to protect the gauges when not in use.

Procedure

Figure 3.2 shows the layout of the experiment. All the equipment was switched on ten minutes prior to each session to allow the output of the gauges to stabilize. Before each subject arrived the transducer was calibrated according to the procedure detailed in chapter two.

The seat for the subject was positioned in the far left hand corner of the sound proof room. The metronome was mounted on the wall above and to the left of the subject when seated. The music stand, elevated to a height of approximately two-thirds of a meter was situated one meter in front of the subject. The dB meter and microphone were placed opposite the subject, approximately on axis and one and a half meters from the bell of the instrument in such a way that the subject could clearly see the dial of the dB meter while playing. To ensure that the placement of the various elements of the apparatus were uniform for each session, permanent positioning markers were fixed to the floor and walls of the room.

Figure 3.2 Experimental Layout



On arrival the subjects were only told that the experiment involved looking at "various aspects of trumpet performance", although they were also told to expect a fuller explanation at the end of the experiment. The subject was then shown into the sound proof room and seated in the chair. Each subject had been told to bring their own trumpet and mouthpiece. The transducer was then interposed between these. Vaseline had been applied to all the tapers to prevent jamming the transducer. The main tuning slide of the trumpet was pushed all the way in to give maximum compensation for the pitch lowering which resulted from the extra length that the transducer added to the mouthpipe section of the trumpet. This action also obviated the need for the players to tune up as the pitch of the different instruments was thus rendered the same.

The subject was then requested to warm up in the same manner as they would before any normal performance. Five minutes were allotted for this. This also permitted them to get used to the slight differences in the playing characteristics of the modified trumpet, and to become accustomed to the dry acoustics of the sound proof room.

During the adaptation phase the players were asked to examine and rehearse the more difficult test exercises in time with the metronome. They were also able to become familiar with the different sound pressure levels that were to be required in the experiment proper.

The experimenter was seated at a table just outside the sound proof room. The electrical connections between the instrumentation in the sound proof room and the experimenter's monitoring equipment were made via a DIN plug and socket patch bay set in the wall of the chamber itself. The subject and the experimenter remained seated in their respective positions throughout the entire experiment.

The five chief sections of test material were carried out in the same order for all the subjects. The instructions to the subjects were given before each section which allowed time to rest from the previous exercises.

The subjects were tested at the time of day suited to their own convenience, although they were asked to avoid practising on the day of the experimental session.

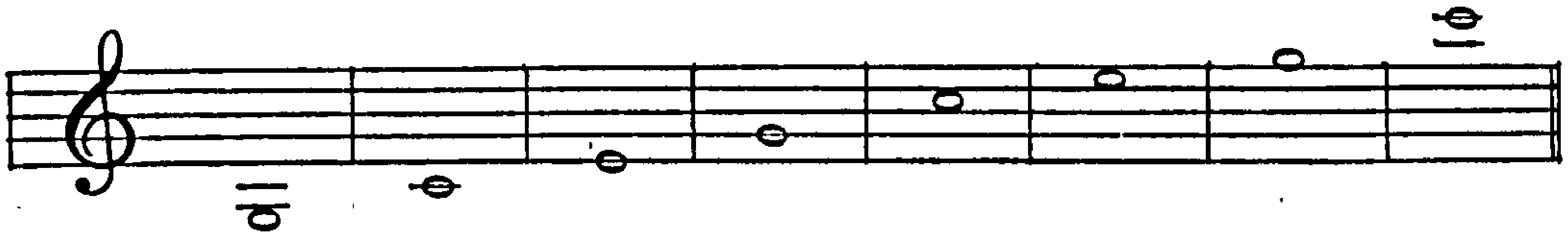
The five sections will be treated as discrete studies forthe purposes of analysis, although there will be a concluding discussion to present an overview of the complete experiment. The order of reporting of each of the individual studies in this experiment follows the order of their presentation to every subject.

I LONG NOTE STUDY

Design

There were eight levels of pitch (see figure 3.3) representing the basic harmonic series on the trumpet between G_3 and C_6.

Figure 3.3
Experimental pitch levels.



Each pitch was performed at three levels of intensity; 87, 95 and 103 dB. With a single replication of each note at these three intensities, 48 separate notes were played by each subject.

Fifteen different randomizations of the 48 tones were devised using random numbers tables. With each subgroup containing 15 subjects, all the subjects in the same subgroup received different randomizations.

The 48 tones were played as semibreve, (where crotchet = 60) with a space of four crotchet beats rest between each note. There was a break of fifteen seconds after every 6 notes giving eight groups of six notes.

The subjects were required to play the notes at "piano" (87 dB), "mezzoforte", (95 dB) and "forte" (103 dB), and adjust the loudness of the note until they could see by the dial of the dB meter that the desired sound level had been reached. They had become used to the levels required from the warmup period, and most of the necessary accommodation was managed without difficulty.

The study proposed to address the following general questions;

1) What is the nature of the relationship between pitch, intensity and mouthpiece force ?

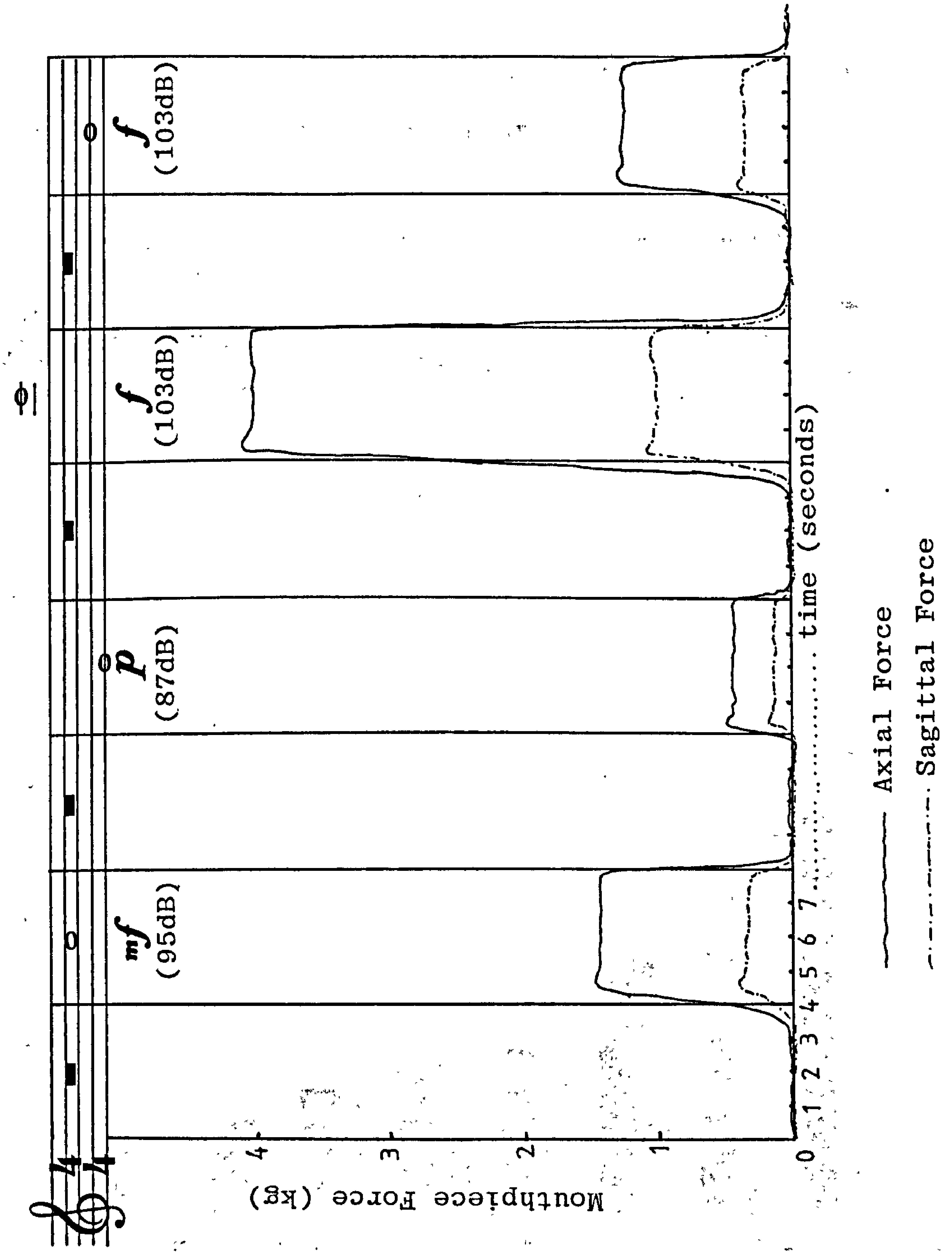
2) Are there any differences in the use of mouthpiece force between levels of proficiency or performance style?

Results

Figure 3.4 shows a typical mouthpiece force trace for the experiment, taken from subject B.B. After an initial adjustment to reach the required intensity, the mouthpiece force was held fairly constant until the end of the note.

The mean of the two resultant forces at each pitch-intensity combination gave the criterion score for the subject's mouthpiece force measure. The research design had been structured originally to provide data amenable to analysis

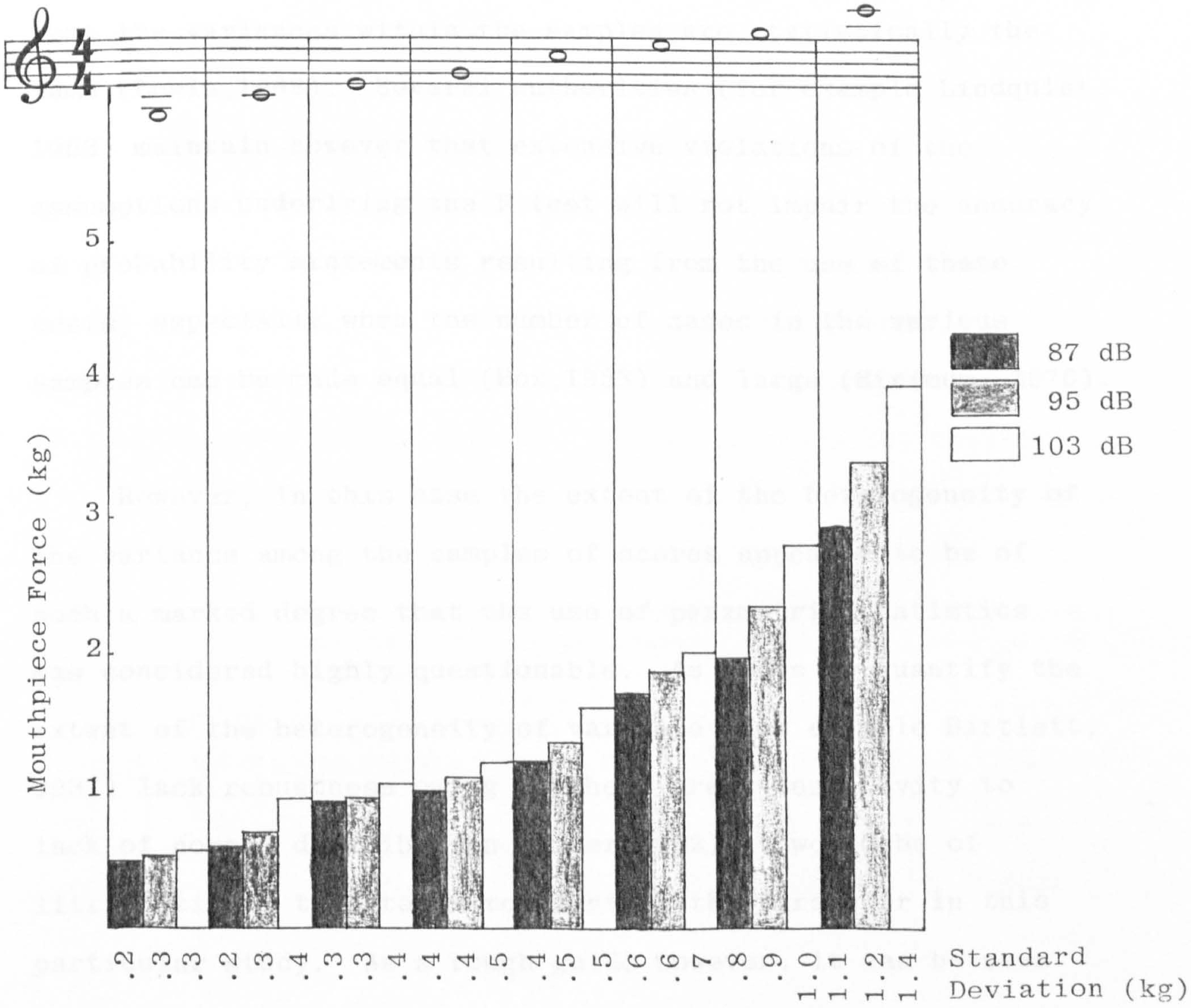
Figure 3.4 Mouthpiece force tracing for subject B.B.; long note study.



by a mixed model multifactor analysis of variance (Hays 1963), with repeated measures on the random effects factors of pitch and intensity, and with proficiency level and performance style as the fixed factors. However, the nature of the data which emerged made this type of overall test inapplicable. The reasons for this are as follows.

Figure 3.5 shows the summary table for the pooled data of all 60 subjects in bar graph form.

Figure 3.5 Long note study results for 60 subjects.



The height of each bar gives the arithmetic mean of the mouthpiece force used at the given pitch and intensity. The standard deviation of the mean is also given beneath each bar. It can be seen that the variance of the mouthpiece force scores increases with both pitch and intensity. Furthermore, frequency polygons of the data show a distribution which has a marked departure from normality. These conditions represent major violations of the two most crucial assumptions behind the use of the analysis of variance. These are firstly that the samples which are analysed have been drawn from populations which are normally distributed, and secondly that the variances within the samples are statistically the same (Lewis,1968). Several authorities (for example Lindquist, 1953) maintain however that extensive violations of the assumptions underlying the F test will not impair the accuracy of probability statements resulting from the use of these tests, especially when the number of cases in the various samples can be made equal (Box,1953) and large (Minimum, 1970).

However, in this case the extent of the heterogeneity of the variance among the samples of scores appeared to be of such a marked degree that the use of parametric statistics was considered highly questionable. As tests to quantify the extent of the heterogeneity of variance (for example Bartlett, 1937) lack robustness owing to their great sensitivity to lack of normal distribution (Winer,1962) it would be of little utility to attempt to quantify the parameter in this particular study. As a rough guide however, it can be seen from figure 3.5 that the sample variances differ up to a factor of 25 times between pitch levels alone. It was therefore decided to use primarily distribution free statistics where the homogeneity of variance assumption was violated to a marked degree.

Intensity

It was not possible for the above reasons to give a simple direct demonstration of the effect of intensity by way of a <u>main effect</u> as this would have involved using means derived by averaging over the totality of the pitch factor levels. The general effect of intensity is thus shown indirectly by a series of sixteen Wilcoxon matched pairs signed ranks tests (Siegel, 1956) with two for each pitch level. These tests show that, at every pitch level of the experiment, the increase in mouthpiece force is significant for the increase in intensity between 87 and 95 dB, and between 95 and 103 dB. Table 3.1 shows the results of these individual tests.

Table 3.1 Wilcoxon tests on mouthpiece force differences for two intensity changes at eight fixed pitch levels.

87-95 dB								95-103 dB								
4.5	5.8	5.3	5.6	6.3	6.6	6.3	6.7	5.7	5.7	6.2	6.3	6.3	6.7	6.7	6.7	z score
			P<.0001								P<.0001					p =
G_3	C_4	E_4	G_4	C_5	E_5	G_5	C_6	G_3	C_4	E_4	G_4	C_5	E_5	G_5	C_6	pitch

It was noted that the size of the difference scores (the criterion statistic of the Wilcoxon test) for the effect of intensity level increased systematically as pitch increased. This suggested the existence of an interactive effect between pitch, intensity and the dependent variable. The size of the effect, although evident from the difference scores is unfortunately not adequately reflected in the increasing z scores in table 3.1. The effect of changes in intensity on mouthpiece force is to produce almost universally like-signed ranks at all the test pitches. This leads to a sort of statistical ceiling effect which although unequivocally demonstrating the intensity effect, obscures the extent of the interaction which exists. With reference to table 3.1, 6.74 is the maximum possible z score for the Wilcoxon test where n = 60.

Pitch

The effect of pitch on mouthpiece force is even more marked. For the same reasons as given earlier with intensity, the influence of pitch on the dependent variable cannot be measured as a simple main effect. As with intensity therefore, a series of Wilcoxon matched pairs signed ranks tests were carried out. This involved comparing, at each intensity level, the subjects' mouthpiece force scores for adjacent pitch levels. A total of 21 tests were executed and each showed that the effect of pitch on mouthpiece force was highly significant. (Table 3.2 shows the results of these analyses).

Again the increase in the size of the difference scores with ascending pitch suggested another interactive effect of importance, although this once more is not adequately shown by the z scores in table 3.2.

Table 3.2 Wilcoxon tests on mouthpiece force differences for seven pitch changes at three fixed intensity levels.

	G_3-C_4	C_4-E_4	E_4-G_4	G_4-C_5	C_5-E_5	E_5-G_5	G_5-C_6	Pitch changes
z	6.45	6.61	6.50	6.62	6.58	6.62	6.70	87 dB
p			$p < .0001$					
z	6.21	6.45	6.21	6.72	6.59	6.67	6.74	95 dB
p			$p < .0001$					
z	6.45	6.56	6.73	6.73	6.60	6.62	6.74	103 dB
p			$p < .0001$					

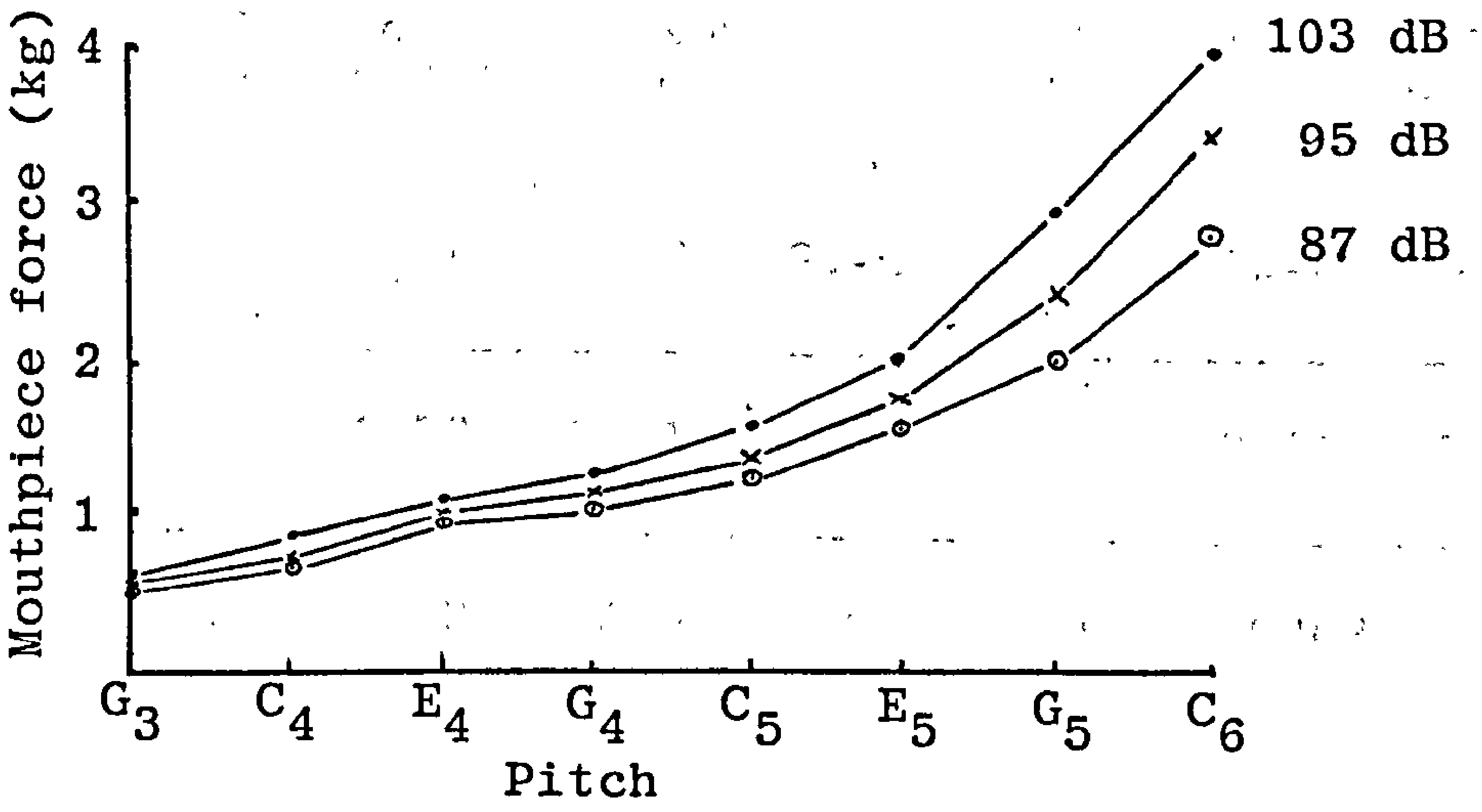
(The maximum value of z where n=60 is 6.74)

Interactions

Figure 3.6 shows how the magnitude of the effect of intensity on the dependent variable changes with the experimental levels of pitch. At G_3, a 16 dB difference in intensity is accompanied by a mean change of .1kg mouthpiece force. At C_6 however the same intensity change results in a mean force difference of 1.2kg.

Given the restrictions on the use of parametric analyses of variance with this data, it was decided to use the combined S test (Jonckheree, 1954) to demonstrate this interaction statistically. The combined S method tests ordered predictions about comparison samples.

Figure 3.6 The interactive effects of pitch and intensity on mouthpiece force.



The repeated measures design in this experiment gives eight related samples where each sample is composed of scores which are the <u>increment</u> in mouthpiece force associated with the 87 - 103 dB intensity increase for the sixty individual subjects at each pitch level. The combined S test with the normal approximation for large samples enables us to reject the null hypothesis ($J = 1107$, $z = 17.6$, $p < .0001$) and confirm that the mouthpiece force differential for the given intensity change increases with ascending pitch. However, figure 3.6 shows that the ostensibly large effect in the pitch region C_5 to C_6 may have chiefly contributed to the overall significance finding. To establish the extent to which the upper pitch comparisons are responsible for the result, a series of multiple comparisons were carried out using the Wilcoxon test. To avoid Type 1 errors which are always a danger in post hoc tests of this kind (see Leach, 1979), the per experiment error rate was fixed at .01. This gives a per comparison rate of 0.0004, as calculated by Leach's procedure.

Table 3.3 Post hoc Wilcoxon tests comparing the intensity related increments between contiguous pitch levels.

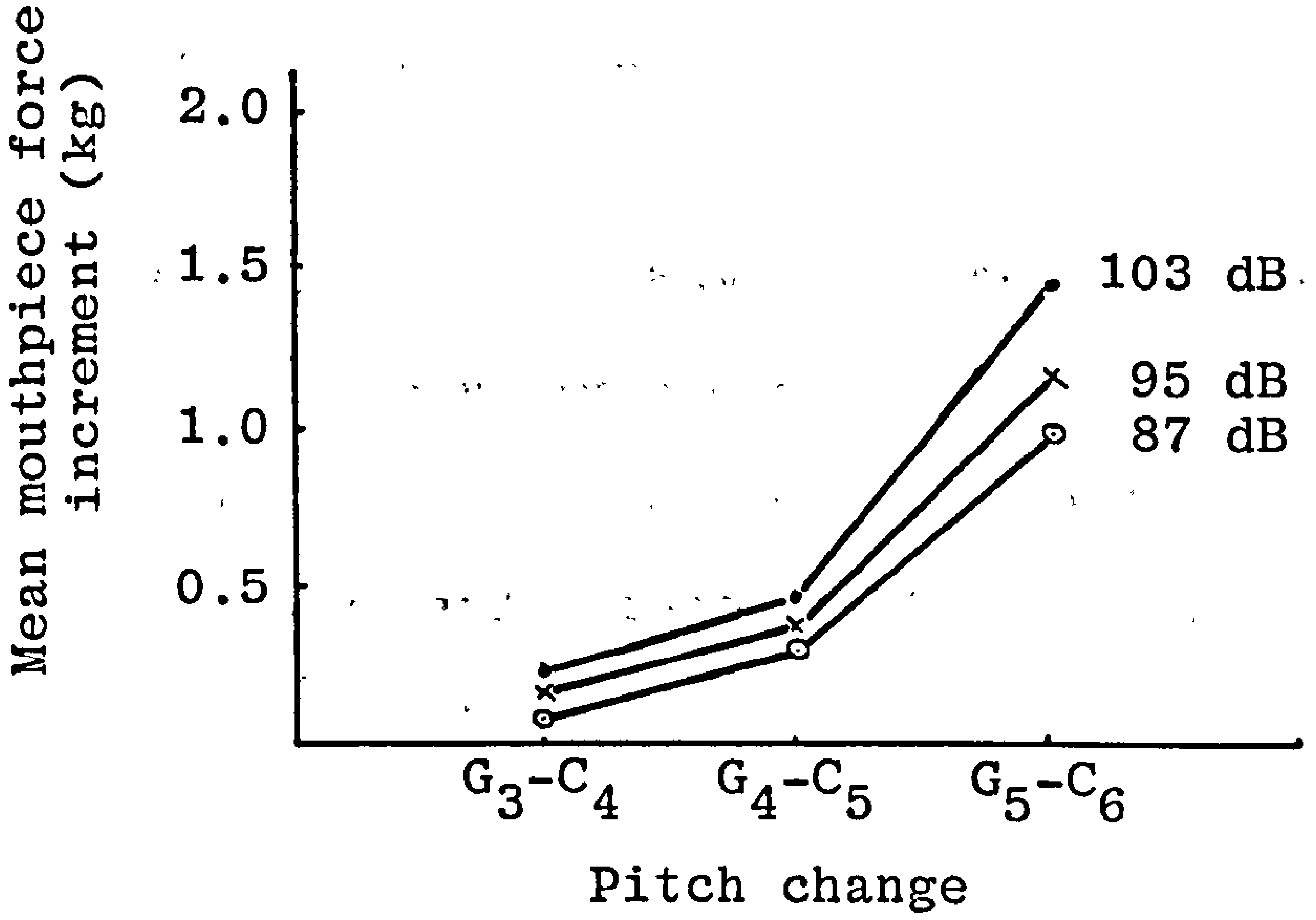
Comparison	G_3-C_4	C_4-E_4	E_4-G_4	G_4-C_5	C_5-E_5	E_5-G_5	G_5-C_6
z score	2.72	1.53	3.21	3.37	3.89	3.61	5.81
probability	.003	.063	.0007	.0004	.0002	.0002	.0001

Table 3.3 shows the probabilities (one tailed) associated with the z score for comparisons between contiguous pitch levels. This shows that on this basis only the comparisons G_4 - C_5 and above demonstrate significance. However, a less conservative per experiment error rate might have been justifiable in view of the nature of this data, and would have resulted in significant findings at lower pitch levels. Nevertheless, table 3.3 does give a more precise account of the disposition of the observed interaction.

A further interaction related to pitch level is shown in figure 3.7. This reveals that a local rise of pitch has a differential effect on the magnitude of the concomitant mouthpiece force increment depending on the position in the register where the local pitch difference lies. Figure 3.7

shows the equivalent local pitch change G - C for the three adjacent registers with the resultant effect on mouthpiece force at the three experimental intensities. In this diagram it can be seen for example that at 103 dB, a change of pitch from G_3 to C_4 results in an increase of mouthpiece force of 0.2kg (mean), whereas the equivalent pitch difference two octaves higher (G_5 to C_6) increases the mouthpiece force by 1.4kg (mean).

Figure 3.7 Comparison of mouthpiece force increments for an equivalent pitch change in three different registers at three intensity levels.



This effect is evident at each intensity level. As the data once more precluded the parametric approach of analysis, the combined S test was again used, with three tests being made (one for each intensity level), each testing ordered predictions about three samples. Each sample contained the sixty scores representing the individual increments in mouthpiece force from G to C for the low (G_3 - C_4), medium (G_4 - C_5) and high (G_5 - C_6) register conditions. With the normal approximation for large samples, the null hypothesis of no difference in the magnitude of the increment of mouthpiece force for the equivalent local pitch changes between different registers was rejected at 87 dB ($J = 38.2$, $z = 2.6$, $p < .005$); 95 dB ($J = 46.4$, $z = 3.14$, $p < .001$) and 103 dB ($J = 50.3$, $z = 3.4$, $p < .0005$).

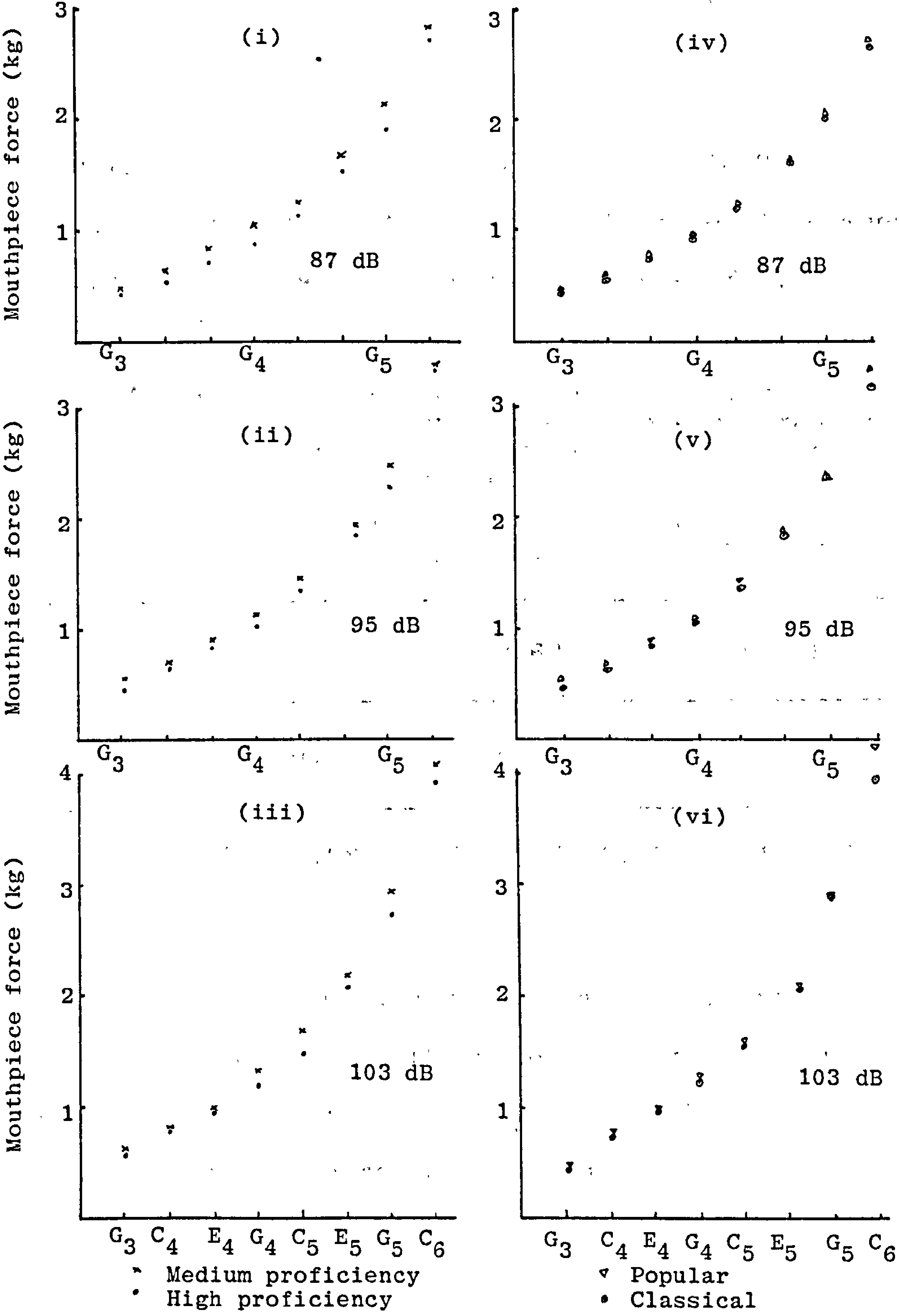
Performer Variables

Firstly we are interested in the effect of performer attribute variables on the absolute magnitude of mouthpiece force in the long note study.

Figure 3.8 (i) - (vi) shows the effect of proficiency level (i - iii) and performance style (iv - vi) on the dependent variable. The data seem to show that, at every combination of pitch and intensity, there is a very small but consistent difference between the referent groups.

To maximise the likelihood of finding any existing effects of performer proficiency or performer style, the mouthpiece force scores were collated into the four subgroups and submitted to a series of single factor analyses of variance, one for each pitch-intensity combination. The use of the parametric test is justified in this case because the homogeneity of variance assumption holds good between the four samples of scores within a single fixed pitch-intensity level; it is only when computing across pitch or

Figure 3.8 The effect of proficiency level and performance style on mouthpiece force.



intensity levels when the problem arises. Although in this case there is a question as to the normality of the distribution of the scores, the test should be robust enough where the other assumptions are fulfilled (Ferguson, 1971).

The results of the analyses are given in table 3.4. With 3,56 d.F. a significant effect at the 0.05 level between the groups would require an F value in each case of 2.78.

Table 3.4 Analyses of variance comparing the four performer sub-groups at all the experimental pitch-intensity combinations for absolute mouthpiece force usage.

dB level	G_3	C_4	E_4	G_4	C_5	E_5	G_5	C_6
87	1.01	1.64	1.22	1.18	0.63	0.43	0.40	0.48
95	0.85	0.88	0.91	1.01	0.68	0.17	0.17	0.64
103	0.88	0.55	0.68	0.56	0.67	0.15	0.44	0.66

As none of the F values approach significance, the null hypothesis of no difference in the absolute magnitude of mouthpiece force usage between proficiency level or performance style cannot be rejected at the 0.05 level of confidence.

In the light of the failure of these tests to uncover any differences in the absolute use of mouthpiece force among the referent groups, it is perhaps unlikely that tests to compare differences in the range of individual mouthpiece force usage would result in significant effects. However, as an attempt to replicate the findings of previous studies, this analysis is warranted. Two further series of single factor analyses of variance were carried out to determine whether the groups differed in the ranges of mouthpiece force used for (1) differences in intensity and (2) differences in pitch. Again, because each analysis was carried out at single fixed pitch-intensity combinations, it was considered reasonable to use the parametric test.

Table 3.5 shows the results of analyses of variance at each of eight pitch levels for the range of mouthpiece force used between 87 and 103 dB by the four groups of subjects. With 3,56 d.f. a significant difference between the samples would require an F value of 2.78 in each case (at the .05 level).

Table 3.5 Analyses of variance at eight experimental pitch levels comparing the four performer subgroups for range of mouthpiece force used for the intensities 87 - 103 dB.

Pitch	G_3	C_4	E_4	G_4	C_5	E_5	G_5	C_6
F value	.98	.82	.33	.08	.48	.49	1.22	.66
p	n.s.							

Table 3.6 shows the results of analyses of variance testing the differences in forces associated with pitch changes for

the four criterion groups. Tests are made at each intensity level and for three equivalent pitch changes in different registers.

Table 3.6 Analyses of variance at three intensity levels comparing the four performer sub-groups for range of mouthpiece force used between G - C in three registers. F values are given.

	87 dB	95 dB	103 dB
G_3-C_4	0.72	0.35	0.18
G_4-C_5	0.12	0.97	1.36
G_5-C_6	1.43	1.87	.28

Once more, none of the computed F values reached the critical value for significance.

The data therefore show convincingly that neither proficiency level nor performance style affect the individual range of mouthpiece force variation for both (i) differences in intensity and (ii) differences in pitch.

Discussion

Although the effects of pitch and intensity on the dependent variable conformed generally to expectations, the value of this study was in providing a more exact quantitative description of the extent of the association between these variables.

The long note data show that pitch is the major determinant of mouthpiece force, followed by intensity. There are two points which are important in relation to this however. Firstly, this finding is truly representative of the individual subject's force usage, and is not purely a statistical summary statement arising from the data reduction process. Inspection of the raw data reveals that the postulated effects of both pitch and intensity are unequivocally reflected at the level of each individual player's pattern of force application. As far as every player is concerned, higher pitches or intensities were almost invariably associated with increased use of mouthpiece force. The only exceptions to this occurred occasionally at the lowest pitches and intensities. This is where the force used is lowest, and where in addition the effects of the small degree of random variability in the transducer output are most acute. The universal effect of pitch and intensity on mouthpiece force resulted in the statistical ceiling effect observed in tables 3.1 and 3.2.

Altogether, these findings show that although all players vary the magnitude of mouthpiece force in effecting register or intensity changes, it seems to be the case that each player differs appreciably in general personal overall force levels. By way of illustration, table 3.7 shows how players' use of mouthpiece force at low pitch levels correlates with their performance at high pitch levels. The high observed correlations suggest that it would be reasonable to classify players according to their overall force usage, perhaps as low, medium or high category players.

Table 3.7 Correlation at three intensity levels of mouthpiece force usage for individual players at G_3 and C_6.

Intensity	Mean mouthpiece force		Spearman rho	t	p
	G_3	C_6			
87 dB	0.45	2.80	0.792	9.9	$<.001$
95 dB	0.50	3.35	0.811	10.5	$<.001$
103 dB	0.57	3.99	0.807	10.39	$<.001$

These circumstances reveal how the proficiency level and performer style sub-groups (see figure 3.8) could yield consistent differences at each pitch-intensity level without these differences attaining statistical significance. The small differences between the group means are attributable simply to individual differences within the respective samples which operate systematically across the experimental pitches and intensities.

These results suggest two factors which might account for the reason why performance on the trumpet becomes progressively harder with ascending register and higher intensities. Firstly, the absolute levels of force application will be highest here, and the deleterious effects of large forces at their greatest. Secondly, the interaction analysis showed that intensity changes and localized pitch changes in the upper register involve the greatest variation in mouthpiece force. Throughout the low and middle register, smaller mouthpiece force change seems to be necessary to achieve given changes in pitch and intensity.

This study examines very broad limits of performance. Certain reservations however must be stated. The use of long notes was mainly for convenience, as there are problems with the precise regulation of intensity in more complicated musical contexts. The effect of different trumpet techniques and musical exercises on mouthpiece force is of great importance, however and will be examined in the next section.

Also, it must be understood that the experimental pitch levels are not intended to be equispaced, being the notes of the natural harmonic series of the trumpet. This arrangement was again dictated by convenience.

II DIVERSE MUSICAL EXERCISES

The second part of the experiment aims to see how the use of mouthpiece force varies with some of the major techniques involved in brass performance, and particularly to attempt to relate the use of mouthpiece force in different musical contexts to the player attribute variables under consideration.

At the pilot stage of the experiment it was observed that the subjects all had great difficulty in maintaining the required intensity levels with any consistency while playing extended exercises. This is partly because the exercises were too complicated to be quickly memorized and the subjects were unable to read the music and monitor their sound level simultaneously. In any case, performance of a complex passage at an exactly uniform intensity level is extremely difficult. The task was therefore simplified by using the two intensity levels, 87 dB and 103 dB. The implications of the observed deviations from the criterion will be discussed as they affect each analysis.

Design

The three different musical exercises to be performed by each subject were (i) An arpeggio on the basic harmonic series on C.

(ii) A two and a half octave scale in C.

(iii) A lip flexibility exercise on the harmonic series in C.

To maximise the validity of comparisons between these exercises, they were so constructed to be as similar as possible with regard to variables extraneous to the analysis. They were of exactly equal length. All could be performed in one breath. The full two and a half octave trumpet range (G_3 to C_6) was encompassed in each case, and the distribution of notes within the passages was as closely similar as possible.

Figure 3.9 shows the exercises designed specially for the experiment. As a test of tonguing and slurring techniques each exercise was performed in the following ways:

1) Tongued - 87 dB (2) Tongued - 103 dB (3) Slurred - 87 dB (4) Slurred - 103 dB.

With a single replication of each of the above conditions, each exercise was performed eight times, giving 32 items to be

Figure 3.9 Musical exercise material.

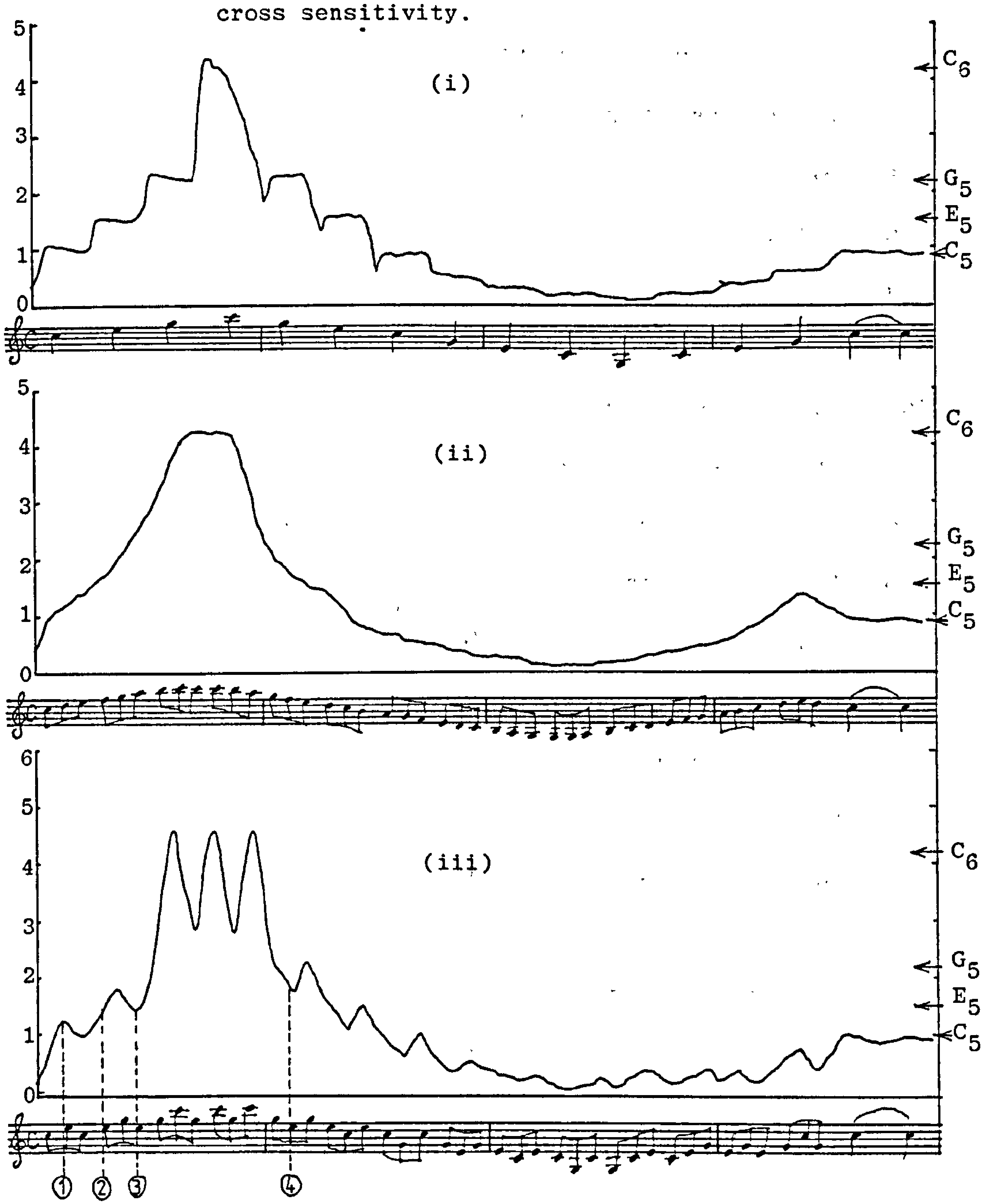
played by each subject. To control for order and carryover effects here it was decided to use a mixture of the intra-group counterbalancing techniques suggested by Myers and Grossen (1974). As the primary interest lies in the effect of the different trumpet techniques, the order of presentation of these was partially counterbalanced within each sub-group. Within each exercise, the presentation order of the four conditions was determined by random counterbalancing for each subject. There was a minute rest between each of the three exercise blocks and ten seconds rest between each item.

Results

Figure 3.10 (i), (ii) and (iii) show typical mouthpiece force records for the three exercises, in this case each performed tongued at 103 dB by subject C.M. The important point to note is that the variation in mouthpiece force is characteristically different in each exercise, and can be seen to relate precisely to the contour of the sequence of notes being played. The trace of mouthpiece force seems to follow the profile of the musical passage very closely. Moreover, the performer seems to adjust the mouthpiece force for almost every single note played. This is seen to best effect in exercise (iii), the lip flexibility exercise, where mouthpiece force is changed systematically at the rate of three adjustments per second in accordance with the triplet character of the exercise.

An indication of the consistency of the adjustments in force can be given by comparing the forces used for given notes between the exercises. The ordinate axis to the right shows the mouthpiece force results for the same subject at the same intensity level from the long note study. To look more closely into the question of consistency in mouthpiece force application, the preliminary analysis examined the variation in force application associated with the same test notes in different contexts.

Figure 3.10 Mouthpiece force trace for subject C.M. for the three experimental exercises each at 103 dB. The left hand vertical axis gives the axial force (kg) after correction for sagittal cross sensitivity.



The right hand vertical axis gives the forces used for the given pitches in the long note exercise (I)

Contextual Influences on Mouthpiece Force

1) Within each exercise.

To determine whether within individual exercises contextual factors significantly affected force usage, several non parametric tests for correlated samples were carried out at a representative selection of pitch levels between G_3 and C_6. Obviously, tests were only applicable where more than one occurrence of the test pitch-intensity combination was observed in the particular exercises. Also the mouthpiece force score for each occurrence of the pitch - intensity level is actually the mean of two observations given that each exercise condition is repeated once. Table 3.8 shows the results of the analyses carried out. These show that within the arpeggio and scale exercise, there is no evidence to suggest that internal contextual factors systematically affect the mouthpiece force used to achieve the required notes. Again this is a summary statement about a lack of a group trend. Inspection of the individual mouthpiece force profiles reveal that there is a certain amount of interindividual variation in the fine detail of the force contour, but this does not affect the present analysis.

The lip flexibility exercise however demonstrates a systematic effect for contextual influences. Five of the six pitch levels yield a significant overall result. By way of clarification, the mouthpiece force used to obtain

Table 3.8 Statistical tests for the contextual analysis of mouthpiece force.

Exercise	Pitch	Samples of test pitch	Test		103 dB		87 dB	
					Tongued	Slurred	Tongued	Slurred
(i) Arpeggio	G_5	2	Wilcoxon	z=	0.26	0.93	0.78	1.36
				p=	n.s			
	E_5	2	Wilcoxon	z=	1.40	0.18	0.64	0.96
				p=	n.s			
	C_5	3	Friedman	Q=	3.50	2.26	0.83	1.43
				p=	n.s			
	C_4	2	Wilcoxon	z=	0.42	0.28	0.37	0.93
				p=	n.s			
(ii) Scale	G_5	2	Wilcoxon	z=	0.24	0.86	0.58	0.36
				p=	n.s			
	E_5	3	Friedman	Q=	0.36	0.91	3.32	1.68
				p=	n.s			
	C_5	4	Friedman	Q=	0.62	0.74	0.38	1.92
				p=	n.s			
	C_4	2	Wilcoxon	z=	1.23	0.67	0.46	0.93
				p=	n.s			
(iii) Lip flexibility	C_6	3	Friedman	Q=	7.0	7.5	7.6	7.3
				p=	$<.05$	$<.025$	$<.025$	$<.05$
	G_5	6	Friedman	Q=	21.2	18.7	19.3	20.2
				p=	$<.001$	$<.005$	$<.005$	$<.005$
	E_5	6	Friedman	Q=	19.4	21.2	20.3	20.6
				p=	$<.005$	$<.001$	$<.005$	$<.001$
	C_5	5	Friedman	Q=	13.5	12.6	13.8	13.3
				p=	$<.01$	$<.025$	$<.01$	$<.01$
	C_4	4	Friedman	Q=	8.2	8.7	7.7	8.8
				p=	$<.05$	$<.05$	$<.1$	$<.05$
	G_3	2	Wilcoxon	z=	1.32	1.13	0.82	1.46
				p=	n.s			

the various occurrences of E_5 will be examined in detail. With reference to figure 3.10 (iii) it can be seen that the mouthpiece force associated with the first occurrence (①) of E_5 (as the second triplet in the first group of three triplet quavers) is smaller in magnitude than the force associated with both of the two occurrences of the same note in the second triplet group (② & ③). Post hoc Wilcoxon matched pairs tests show these differences to be significant (see table 3.9). The next occurrence of E_5 (④) is in the fifth triplet group, where the mouthpiece force is greater than that associated with the two observations of E_5 in the second triplet group. Table 3.9 shows that these differences are again significant. Setting a per experiment error rate of 0.05 the per comparison rate becomes 0.008 as there are six possible pairings.

Table 3.9 Post hoc comparisons of pairs of occurrences of E_5 in the triplet exercise.

	87 dB				103 dB			
Comparison	Means		z	p	Means		z	p
① vs ②	1.41	1.72	2.59	<.005	1.66	2.13	2.86	<.0025
① vs ③	1.41	1.76	2.84	<.0025	1.66	2.19	3.48	<.001
② vs ④	1.72	2.02	3.35	<.001	2.13	2.72	3.33	<.001
③ vs ④	1.76	2.02	3.13	<.001	2.19	2.72	3.27	<.001

The data reveal a particular pattern of force application which might be explained as follows. The player appears to "set" the force for a particular triplet at a level commensurate with the main pitch level about which the smaller pitch fluctuation occurs. In our example, the predominant pitch centre for the first triplet is C_5, for the second triplet it is E_5 itself and for the fifth triplet it is G_5. Thus, the player tends to approach E_5 by a small increase from C_5 in the first triplet, and by a small decrease from G_5 in the fifth triplet. A fine analysis of the intensity printout for this exercise provides corroboration for this theory in that the intensity associated with the "pivotal" pitches tends to be higher than with the subordinate pitches. This intensity effect, which amounts essentially to what would be called a musical accent is not on its own large enough to account for the distinctive mouthpiece force patterns obtained with the lip flexibility exercise.

As far as the tonguing and slurring modes of performance were concerned, comparisons were made at five representative pitch levels. The criterion measure of mouthpiece force was an index computed by averaging the forces associated with every observation of that pitch and intensity in the same mode for the totality of the test exercise items. This gives two correlated t tests per pitch level, one for each intensity, comparing the forces used. Table 3.10 shows the computed probabilities (two tailed) associated with the computed t scores. These analyses show that, as far as the modes of tongued and slurred performance are concerned, there is no evidence to suggest that mouthpiece force is differentially influenced by these techniques.

Table 3.10 A comparison of the mouthpiece force used between tongued and slurred performance.

	87 dB		103 dB	
Pitch	t	p	t	p
G_3	0.847	n.s.	1.334	n.s.
C_4	0.688	n.s.	0.271	n.s.
C_5	0.921	n.s.	1.142	n.s.
G_5	0.262	n.s.	0.615	n.s.
C_6	0.342	n.s.	0.108	n.s.

where n=60

Intraindividual deviation from the required intensity tended to be consistent between the different modes of the same exercise. From the point of view of the effect of tongued versus slurred modes of performance, the subject in effect acts as his own control for these departures from criterion intensity.

2) Contextual Influences - Between Exercises

As well as exercises II(i), (ii) and (iii), the long note study was included in this analysis. For each exercise, every subject received scores at each of five representative pitch levels, one for each intensity level. In each case, these scores were composite measures. They were derived as follows.

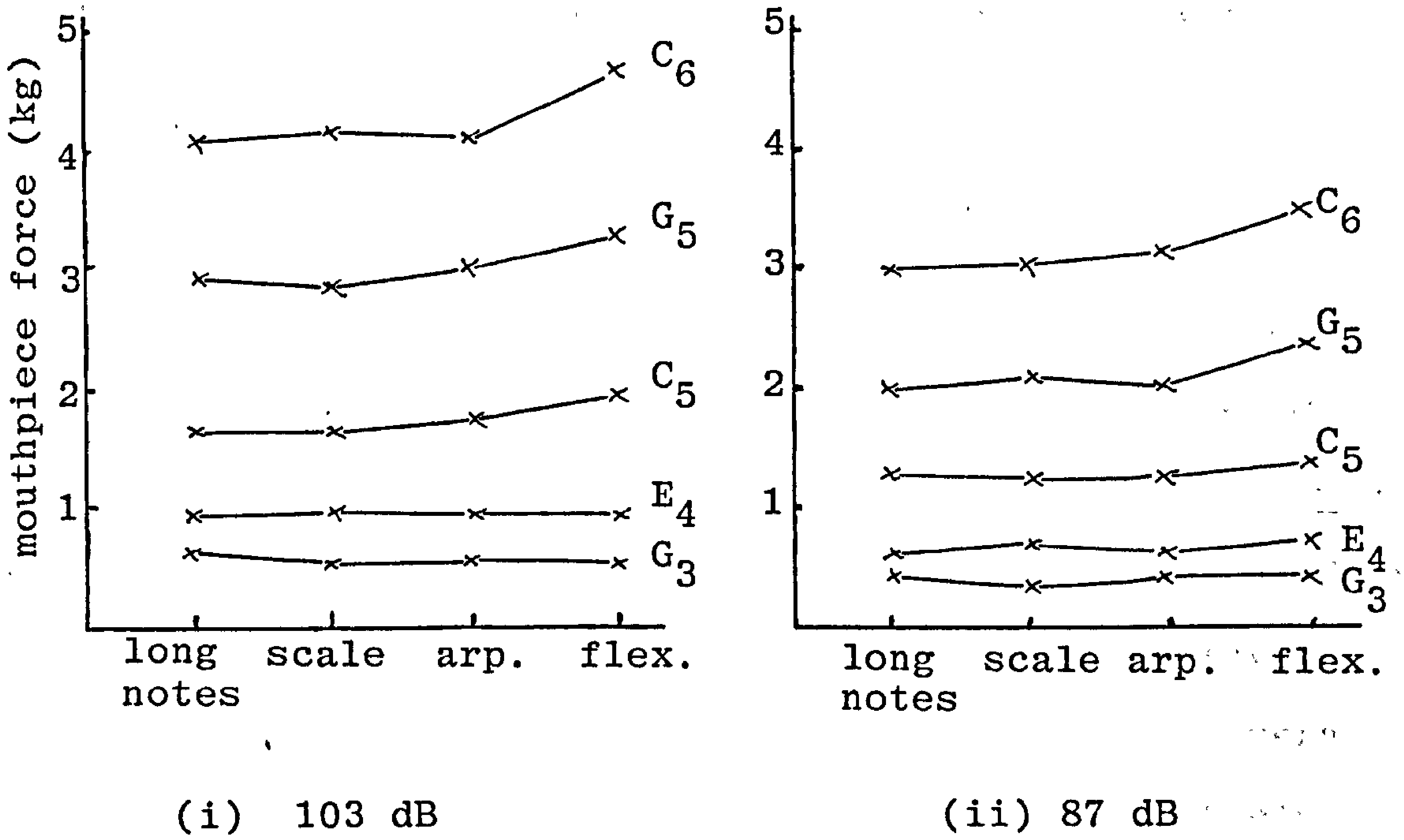
Long notes; the mean force of two observations at each pitch-intensity level.

Arpeggio; at each intensity, the score for the test pitch was the arithmetic mean of the forces applied for each observation of the sample pitch across the totality of the conditions and replications of this exercise. This pooling of data was considered permissible because the previous analysis had failed to find any significant effect for either context or mode of performance on mouthpiece force within this exercise.

Lip flexibility; as the previous analysis found that the nature of the lip flexibility technique led to systematic variability in the amount of force used at the experimental pitch intensity levels within the exercise, it was legitimate to calculate a score in essentially the same way as for the scale and arpeggio exercises, only dropping from the calculation the force scores associated with occurrences of the test pitches where these were not the pitch centre of the triplet in which they occurred.

Figure 3,11 shows the data plotted for the four exercises. The exercise type is given on the absissa, and the ordinate gives the average of the sixty individual composite scores for each exercise.

Figure 3.11 A comparison at two intensity levels of the mouthpiece force used for selected pitches in different test exercises.



The data seem to show that the use of mouthpiece force is very similar for the long note, scale and arpeggio studies. There does appear to be some difference with the lip flexibility exercise, although this effect is evident chiefly at the higher pitch levels.

The Friedman test was chosen as an omnibus test of whether the location of the four related samples differed in any way. The results of this test carried out at the selected pitches and intensities are shown in table 3.11. The level of significance of the resulting Q values are given using the chi-square approximation to the exact distribution which is the appropriate procedure with samples of this size. The order of the exercises in terms of the rank sum for the statistical test (see Table 3.11 column 5) gives essentially the order, from least to most, of the average mouthpiece force associated with each exercise at the specified pitches and intensities.

The omnibus tests carried out show that there are differences in the use of mouthpiece force between the types of exercise for the pitches above and including C_5, and that the differences become more marked with ascending pitch. The various orderings of the rank sums for the four samples suggest that the lip flexibility exercise may be primarily responsible for the significance of each overall test. This possibility was explored using the multiple comparison technique described by Leach (1979). Selecting a per experiment error rate of 0.05 and with six possible dyadic comparisons, a per comparison level of 0.008 is recommended. Successive pairwise comparisons with the Wilcoxon matched pairs test where the overall Friedman test was significant revealed that, for each such test, the lip flexibility exercise involved higher mouthpiece forces than each of the other three exercises. As none of the other pairwise comparisons approached the critical value for significance, it could not be concluded that the use of mouthpiece force was differentially affected by these particular contexts, (i.e. long notes, scales and arpeggios).

Table 3.11 Friedman tests comparing the mouthpiece force used at selected pitch-intensity levels between the different experimental exercises.

Pitch	Q	d.f.	p	Rank sum of mouthpiece force (least→most)	intensity (dB)
C_6	15.2	3	0.005	L S A F	87
	19.3	3	0.001	L A S F	103
G_5	11.6	3	0.01	L A S F	87
	12.5	3	0.01	S L A F	103
C_5	8.2	3	0.05	A S L F	87
	4.9	3	n.s.	L S A F	103
E_4	1.4	3	n.s.	L A S F	87
	2.1	3	n.s	F A L S	103
G_3	4.8	3	n.s.	S L A F	87
	3.7	3	n.s.	F A S L	103

where L = Long note study

S= Scale exercise

A= Arpeggio exercise

F= Flexibility exercise

The problem with the intensity regulation is relevant here. The subjects found it difficult to maintain precisely the required dynamic level throughout each exercise. Inspection of the recording from the dB meter showed that departures from the desired intensities were probably minimal for the scales and arpeggios, but in the case of lip flexibility sometimes quite marked. The nature of the technique required here made performance of the higher register notes at a higher intensity level very difficult, causing a tendency of the subjects to play under the designated intensity level.

However, in view of the general relationship between intensity and mouthpiece force observed in this study, it can be seen that the finding that greater mouthpiece force was exerted in the flexibility exercise is actually given even stronger confirmation. This is because in this case the direction of the observed departures from the required intensity would generally be associated with lower applied forces, and would tend to minimise the extent of the effect demonstrated with the flexibility technique.

Performer Variables

As an overall test of whether the consistency in the use of mouthpiece force across different contexts was related to proficiency level, the variance of the force usage for the given intensities and pitches was calculated for each subject. The individual variance scores were used as the sample measures for a series of Mann Whitney U tests comparing the independent samples of high and medium proficiency performers. Table 3.12 shows that, for pitches C_5 and above, the high proficiency group tend to be more consistent in the use of mouthpiece force throughout the various experimental musical contexts.

Given the results of the previous contextual analysis, it was decided to analyse the flexibility exercise specifically. Mann Whitney tests were effected at each

Table 3.12 Comparison between high and medium proficiency performers of the consistency of mouthpiece force application for the experimental pitch-intensity levels.

Pitch	G_3		C_4		E_4		G_4		C_5		E_5		G_5		C_6	
Intensity (dB)	87	103	87	103	87	103	87	103	87	103	87	103	87	103	87	103
z value	.3	.5	1.0	.4	.9	1.1	1.3	1.5	1.7	1.9	2.1	1.9	2.3	2.5	2.6	2.8
p	← p = n.s. →								← p<.05 →				← p<.01 →			

high proficiency n=30
medium proficiency n=30

pitch and intensity above C_5, with absolute force as the dependent variable. Table 3.13 shows that the medium proficiency group tend to use more mouthpiece force for a given note in the flexibility exercise than do the high proficiency group.

Table 3.13 Comparison between high and medium proficiency performers of the mouthpiece force used for selected pitches above C_5 in the lip flexibility exercise, (at two intensity levels).

Pitch	C5	E5	G5	C6	
Means (high v med)	1.7 v 2.1	2.3 v 2.6	3.0 v 3.4	4.4 v 5.0	
z	1.67	1.82	2.56	2.88	103 dB
p	< .05	< .05	< .01	< .01	

Pitch	C5	E5	G5	C6	
Means (high v med)	1.3 v 1.5	1.4 v 1.9	2.0 v 2.6	3.2 v 3.7	87 dB
z	1.35	1.86	2.23	2.67	
p	n.s.	< .05	< .05	< .05	

high proficiency n=30
medium proficiency n=30

Once more, as the medium proficiency group tended to perform the upper register region of the exercise at a lower intensity than the high proficiency group, this proficiency related effect is given even stronger confirmation.

Discussion

The finding that there is no significant difference in the amount of force used for the long note, scale and arpeggio exercises is an overall statement of the lack of a trend for the whole group of sixty performers. The no significance result may nevertheless conceal heterogeneity among the subjects which a more detailed analysis might reveal. In order to give more substance to the suggestion that each individual subject tends to use forces indistinguishable between contexts, it was necessary to consider the correlation between the forces used in the separate exercises. Table 3.14 gives the intercorrelation matrix where each correlation is a mean of the Spearman r indices calculated at G_3, E_4, C_5, G_5, and C_6 for the sixty subjects, for both 87 dB and 103 dB.

Table 3.14 Averaged intercorrelations of mouthpiece force usage by individual subjects for selected pitches and intensities in four different contexts.

	Long notes	Scales	Arpeggios
Scales	.947	-	-
Arpeggios	.920	.918	-
Flexibilities	.757	.793	.824

The uniformly high series of correlations demonstrates that the large degree of consistency in the use of mouthpiece force found in this study operates at the level of the individual subject.

The degree of accuracy and regularity in the application of mouthpiece force revealed in this experiment confirms that the consistency of mouthpiece force may be a very important variable in the acquisition of skills in brass performance. This is especially evident in regard to the finding that the more highly skilled performers tend to be more consistent in the manipulation of these forces across different contexts than intermediate players.

The finding that the lip flexibility exercise was associated with higher force usage is interesting (the high correlation in the relevant column of table 3.14 shows that this effect again operates at the level of the individual subject and is not merely a statistical group tendency). This particular technique is one of the hardest to master on the instrument. It is possible that a performer will sometimes resort more readily to the aid of higher mouthpiece force levels in dealing with tasks of a greater technical complexity. Although this effect was evident at both the high and medium skill levels, it was noteworthy that the higher skilled performers were less susceptible to this tendency.

III TESTS TO DETERMINE THE LIMITS OF THE USE OF MOUTHPIECE FORCE

Design

There were three principal exercises, the order of performance of which was incompletely counterbalanced within each performer sub-group. As the exercises (i) and (ii) demanded performance in the extreme upper register, it was decided to restrict the intensity to the single medium level of 95 dB for these exercises.

(i) Long Notes in the Extreme Upper Register

The notes C_6, E_6 and G_6 were played twice each at 95 dB in an order determined by random counterbalancing. The procedure for this exercise was exactly the same as that for the earlier long note study in the normal register.

Although only 25 out of the 60 subjects were able to perform to the criterion intensity in this range, all the subjects were required to attempt the complete experimental session. This was to ensure that the overall design of the study did not become unbalanced.

(ii) An Ascending Scale in the Extreme Upper Register

A C major scale starting on C_5 and ascending was performed twice by each subject at 95 dB. The subject had to continue up the scale at the required intensity as far into the extreme upper register as possible, (see figure 3.13). The criterion score for each subject in this exercise was the average of the maximum force attained in the two trials, this being irrespective of the actual pitch the subject ultimately attained.

Figure 3.13 Ascending extreme upper register scale exercise.

(iii) Maximum and Minimum Mouthpiece Forces Available to the Performer

The subjects performed the note G_5 at 95 dB in the normal way and were required to actively increase the force on the mouthpiece while keeping the note constant in intensity until their maximum tolerable force had been achieved. Two trials for each subject were made, and the average of the maximum attained forces at the threshold of pain was the score for each subject.

The same procedure was applied with the subject performing G_5 in the normal way and reducing the force used to the minimum possible without resulting in a decrease in intensity.

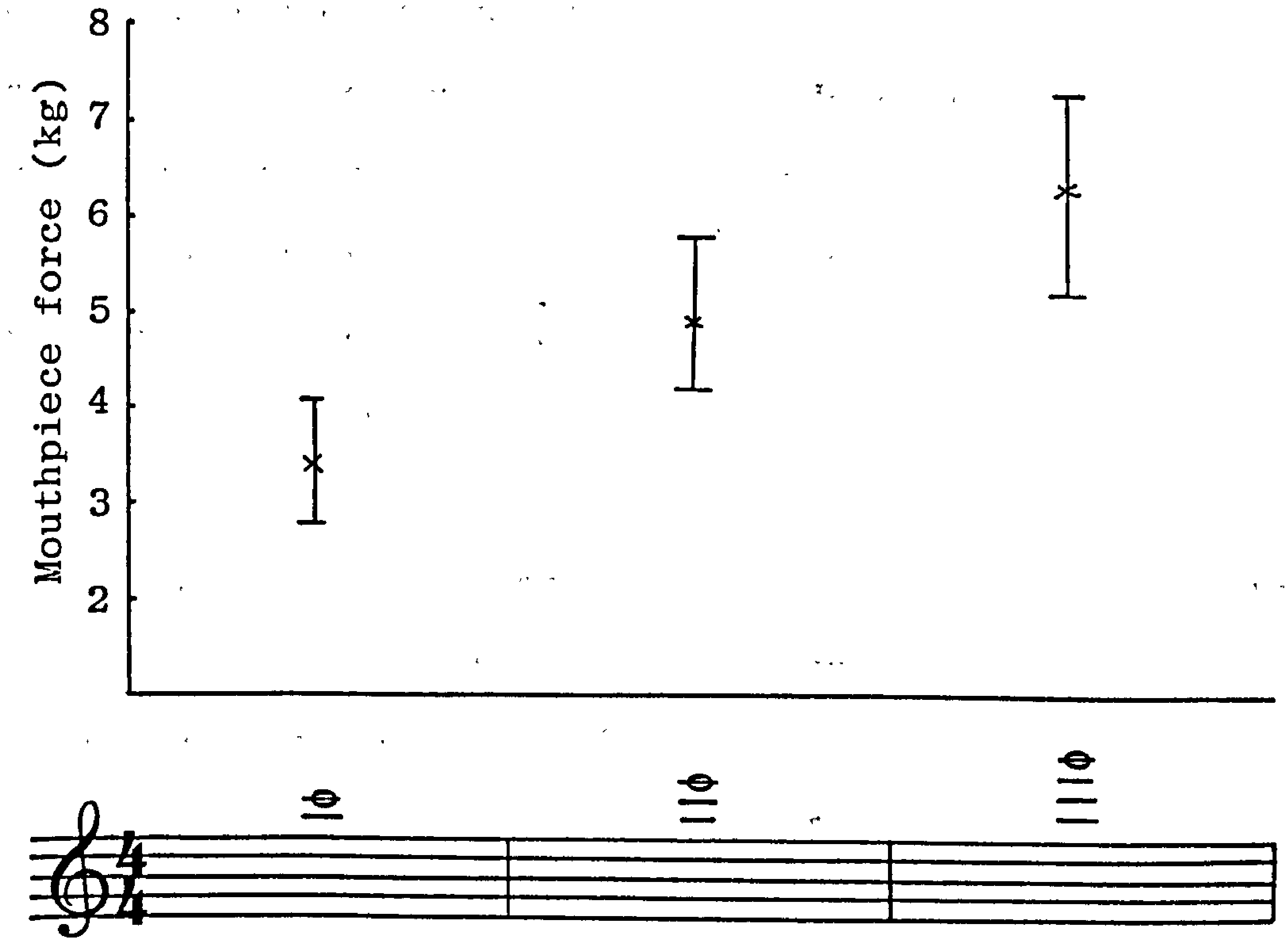
The four trials to determine the available range of forces were performed in an order determined by complete randomization for each subject. The subject was once more required to use the visual display of the sound level meter to maintain the desired intensity.

Results

(i) Long Note Study

Figure 3.14 shows the mean mouthpiece force for the high notes for the 25 players who were able to perform all three notes to criterion intensity.

Figure 3.14 Extreme upper register long note exercise results.



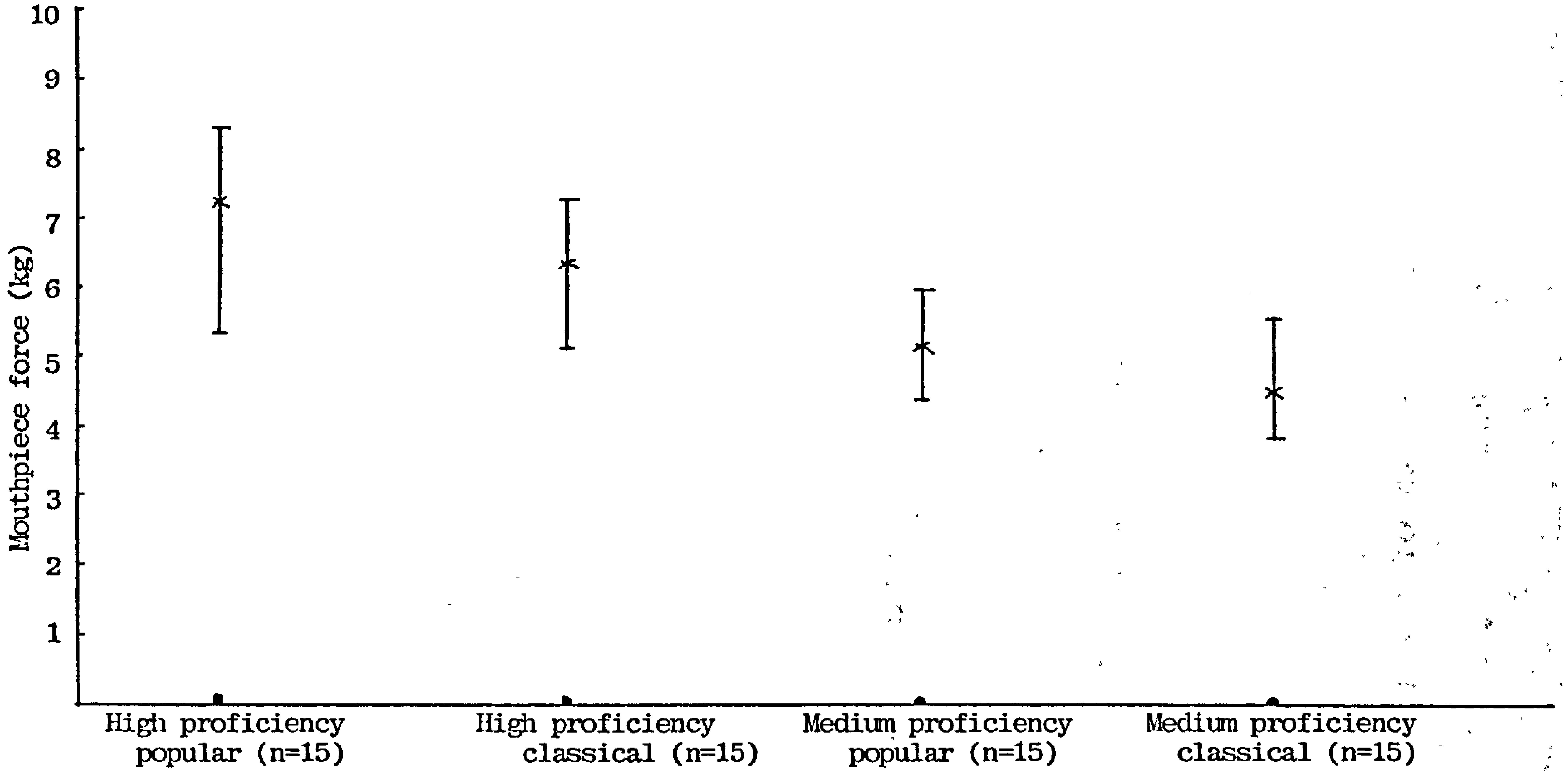
The interquartile ranges are also given. The results show that the increase in mouthpiece force with range found in the normal register is also evident in the extreme upper register. As with the previous long note study, a separate analysis of variance was carried out at each pitch level to look for performer attribute effects. Because different numbers of players in each performer sub-group were successful in performing in this range, it was necessary to use an unequal sample size analysis of variance technique (Winer, 1962). The analysis provided no support for the hypothesis that the mouthpiece force used to obtain these extreme pitches was related to the performer variables. With d.f. 3,21 at both pitches, F values were .48 and .73 (for E_6 and G_6 respectively), both values being well short of significance.

(ii) The Ascending Scale Exercise

For the scale exercise, the maximum recorded mouthpiece force attained by each subject was the dependent variable. The corresponding pitch obtained with this force was not material to the analysis, and varied considerably from subject to subject, (the range for the subjects in this experiment was D_6 to C_8 !).

Figure 3.15 shows the results of the test for the four performer groups, (the means and interquartile ranges are given). An analysis of variance was carried out to look for any systematic differences between the four sample means.

Figure 3.15 Maximum recorded mouthpiece force for the ascending scale exercise by performer subgroup.



The F test revealed a very highly significant effect, (F 3,56 = 10.6; p< .001). To explore the source of these effects it was decided to use an unplanned comparison technique. In view of the insensitivity of the Scheffe test to violations of the homogeneity of variance requirement (Scheffe, 1953), this was the test employed. The pairwise comparison of the means is given in table 3.15.

Table 3.15 Post hoc (Scheffe) tests comparing pairs of performer sub-groups as to the maximum mouthpiece force used in the ascending scale exercise.

	High proficiency classical	Medium proficiency popular	Medium proficiency classical
High proficiency popular	0.9	2.1	2.8
High proficiency classical		1.2	1.9
Medium proficiency popular			0.7

Setting a 5% confidence interval, the computation with the Scheffe procedure yields a critical orthogonal comparison value of 1.9kg. As can be seen from table 3.15, only two of the pairwise comparisons exceed the critical value for significance. This may be due to the relative stringency of the test. However, the nature of the data suggest that comparing combined means might yield interesting results. With respect to table 3.16, the 95% confidence interval for each comparison between the two pairs of means is given by;

$$\hat{\psi}_g - 1.60 \leq \psi_g \leq \hat{\psi}_g + 1.60$$ (Hays, 1963)

Table 3.16 Post hoc (Scheffe) comparisons of pairs of combined sub-group means for the maximum mouthpiece force used in the ascending scale exercise.

			Difference	p
High proficiency	$\frac{7.2 + 6.3}{2}$	=6.75		
vs			1.9	< .05
Medium proficiency	$\frac{5.1 + 4.5}{2}$	=4.8		
Classical	$\frac{6.3 + 4.5}{2}$	=5.4		
vs			0.75	n.s.
Popular	$\frac{7.2 + 5.1}{2}$	=6.15		

The analysis shows that the difference in the mouthpiece force between the high and the medium proficiency level means is significant beyond the .05 level. The absence of a significant finding between the averaged performance style means may again be due to the stringency of the test used.

It was also possible that the circumstances of this test may not have been sufficient to elicit the most extreme forces available to the performer. The last test in this section attempts to do just this.

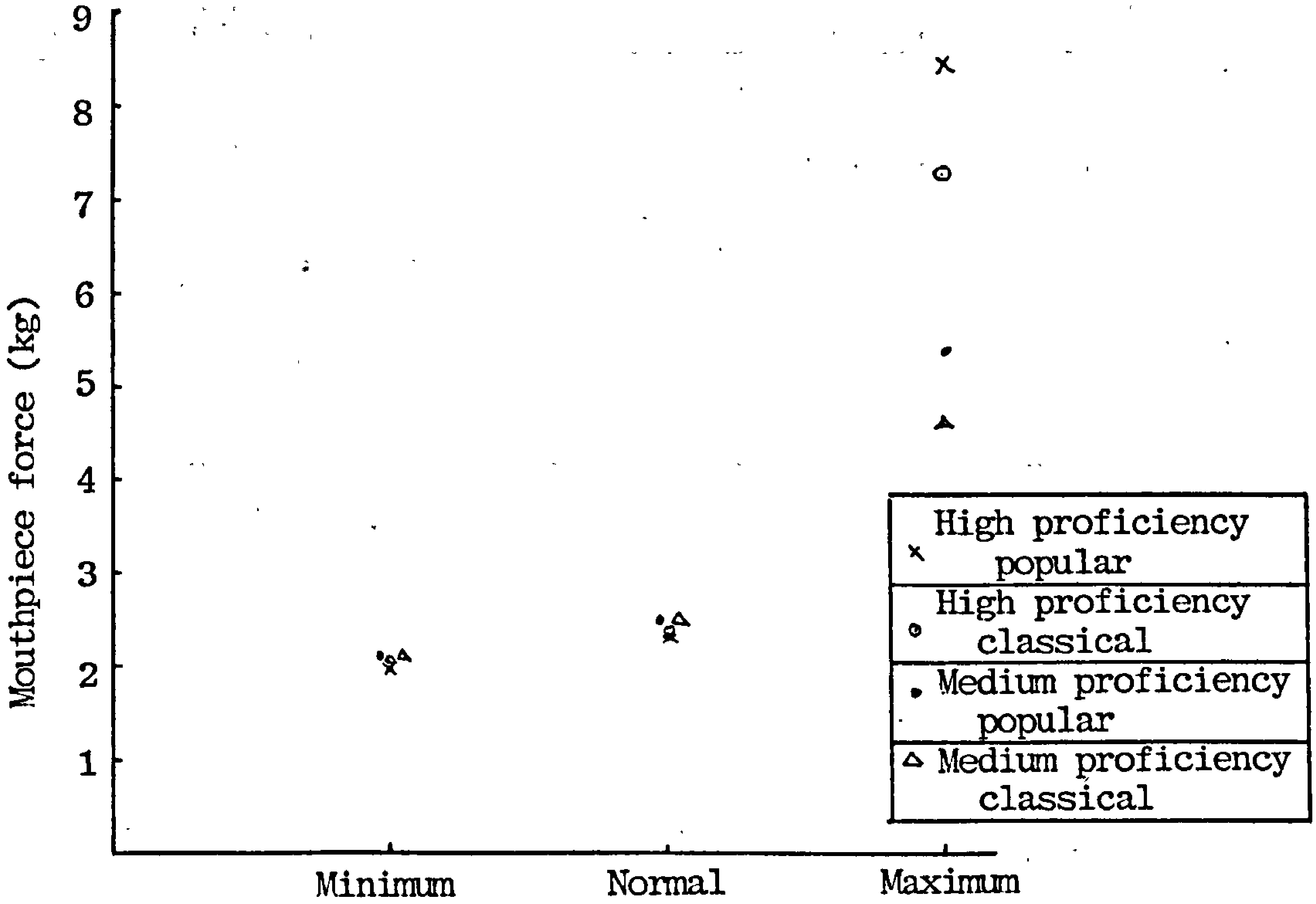
(iii) Maximum and Minimum Forces

Figure 3.16 shows the results for the tests requiring the subjects to actively produce their maximum and minimum possible forces at 95 dB. The data for the normal performance of G_5 is taken from the original long note study.

Maximum Force

An overall single factor analysis of variance for the four sub-groups at maximum mouthpiece force was performed. This gave a highly significant result (F 3,56 = 15.8: $p < .001$)

Figure 3.16 Maximum and minimum mouthpiece forces elicited actively on test III (iii).



This method of eliciting maximum force resulted in a more pronounced differentiation between the referent groups. Once again the Scheffe test was applied to the means. With a reduction in the within groups experimental error in this study, the 95% confidence interval sets the criterion absolute difference between each pair of means at 1.3kg. This results in a greater proportion of significant pairwise comparisons (see table 3.17).

Turning to the main performer variables, table 3.18 shows the results of comparing the average of the sub-group means. With a criterion value of .91 at the 95% level of confidence, both performer proficiency and performer style are found to significantly affect the maximum force produced in this test ($p < .05$ in each case).

Table 3.18 Post hoc comparisons of pairs of combined sub-group means for the maximum force used in test exercise III (iii).

		Difference	p
High proficiency vs Medium proficiency	$\frac{8.5 + 7.2}{2} = 7.85$ $\frac{5.4 + 4.6}{2} = 5.0$	2.8	$< .05$
Popular vs Classical	$\frac{8.5 + 5.4}{2} = 6.9$ $\frac{7.2 + 4.6}{2} = 5.95$	1.0	$< .05$

Minimum Force

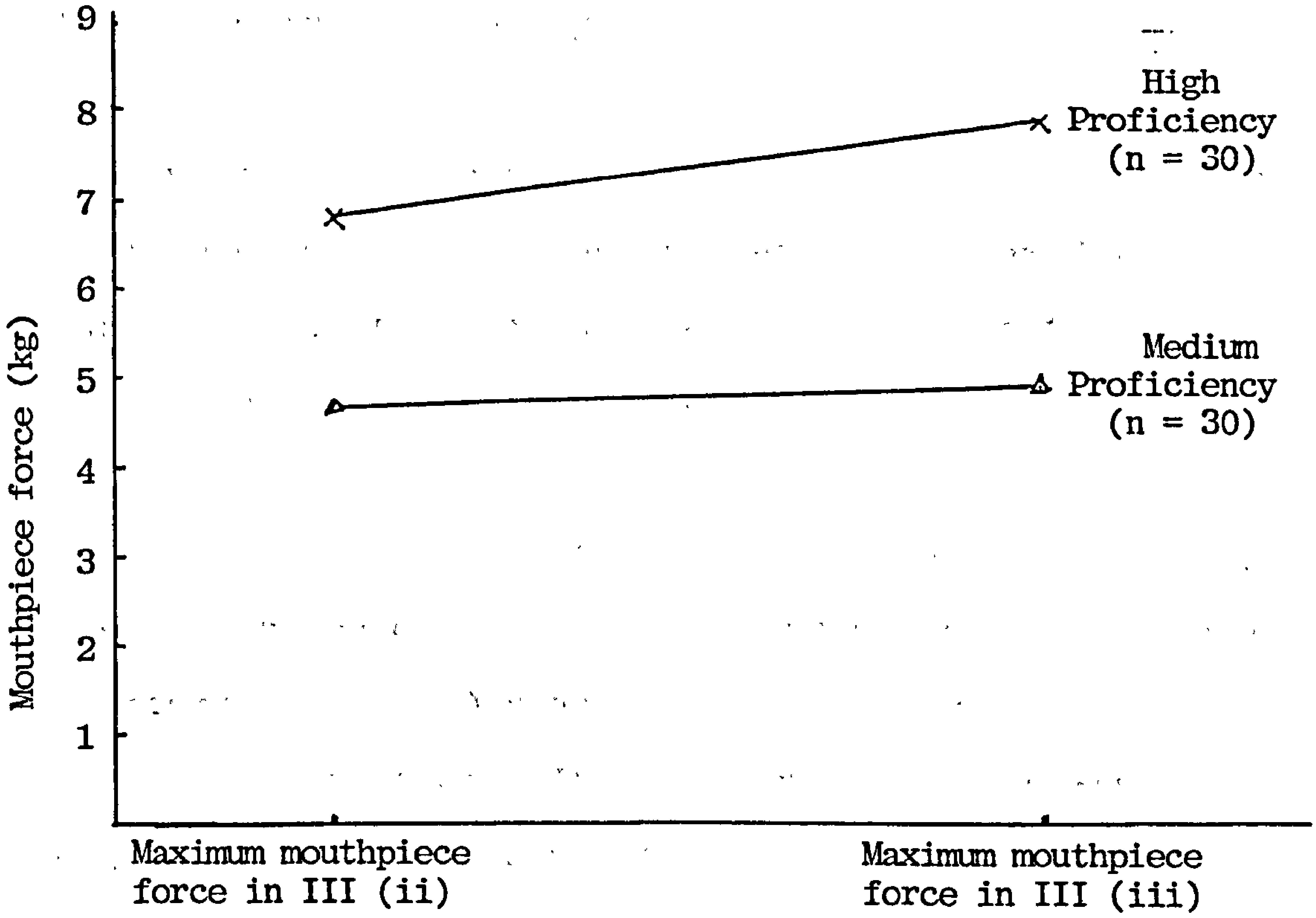
Most subjects were able to reduce the amount of force necessary to produce G_5 at 95 dB. However, the magnitude of the decrements tended to be small. The subjects also evinced great difficulty in maintaining the required intensity with these actively reduced forces. The mean reduction in force at this pitch and intensity was .3kg. A Wilcoxon matched pairs signed ranks test showed this difference to be highly significant ($n = 60$, $z = 6.74$; $p < .0001$)

To look for differences between the sub-groups in the extent to which mouthpiece force could be reduced, a single factor analysis of variance was carried out where the four samples were the scores denoting the difference between the force normally applied and the minimum force which each subject could reduce to for this particular note. The F test failed to show overall significance ($F3,56 = 2.63$; $p > .05$ n.s.). No post hoc analyses were therefore applied to the data.

It is of interest to bring together in a combined analysis the data from tests III (i), (ii) and (iii), all of which deal with the extreme use of mouthpiece force. The long note study in the extreme upper register III (i) provides a basis for discrimination between the performers which transcends the two performer attribute variables used in this study so far. The 25 subjects who were able to perform to the criterion intensity in this study were designated as "successful" upper register players, and the 35 subjects unable to perform these notes were designated as the group of "unsuccessful" high register players. A Mann Whitney U test on the mouthpiece force used by these subjects at C_6, 95 dB reveals no significant difference ($z = 1.17$; $p > .05$ n.s.) between the "successful" group (mean 3.25kg) and the "unsuccessful" group (mean 3.5kg). However, with this same grouping, a Mann Whitney test comparing the maximum tolerable force applied by each subject from test III (iii) as the dependent variable leads to a very highly significant result ($z = 5.01$; $p < .0001$). The means for the "successful" and "unsuccessful" groups are 7.76kg and 4.5kg respectively. This analysis conclusively shows that it is not the case that the successful high register players tend to use lower than average mouthpiece force for their performance of the normal register (i.e. G_3 to C_6). It also shows that the successful high note players are able to tolerate much larger forces than those who are unable to perform in this register.

The scale exercise III (ii) reveals an interesting effect with respect to proficiency level when this data is collated with that from III (iii) where the maximum tolerable force was elicited. The dependent variable in this case is the difference for each subject between the maximum force attained in the scale exercise III (ii) and the maximum tolerable force elicited actively in III (iii). This essentially gives a measure of the extent to which the subjects utilize their maximum available force in the attempt to play the highest notes of which they are capable. Figure 3.17 shows that the high proficiency group used a mean of 1.1kg below their maximum tolerable force, while the medium proficiency group used right up to .3kg from their available maximum. A Mann Whitney U test showed that this difference in the application of force between the proficiency levels was highly significant (z = 3.8 ; $p < .001$).

Figure 3.17 The relationship between proficiency level and maximum mouthpiece force application in tests III (ii) and III (iii).



The same test was used to compare the effect of performance style on this dependent variable, but this failed to reveal any evident difference between the classical and popular performer groups, (z = .53 ; $p > .05$ n.s.).

Discussion

The results show that in the successful performance of the extreme upper register on the trumpet, (i.e. above C_6), extremely large forces are normally applied. The long note study shows that this is the case irrespective of proficiency level or style of performance. The interquartile ranges for the criterion pitches in figure 3.14 show that these extremely large forces are used by most of the subjects in this study. Only two of the twenty five players successful in this register could play G_6 below 5kg at the required intensity.

The successful high note performers were also shown to have available to them a higher potential maximum force, as shown in the test requiring the subjects to actively apply forces up to their tolerable maximum.

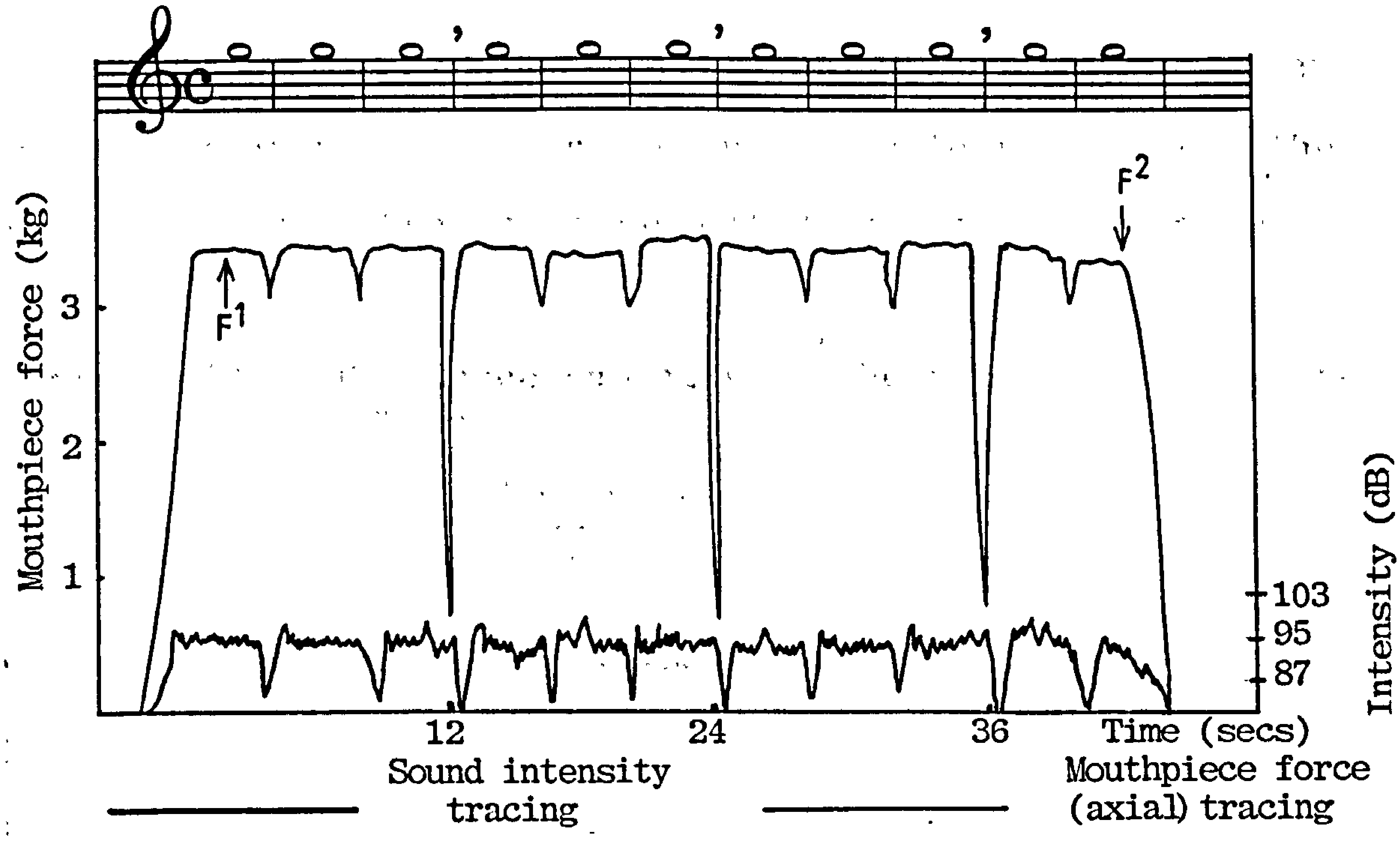
The study revealed a tendency for the medium proficiency performers to respond to an exercise demanding extreme upper pitch performance by applying virtually the maximum force levels available to them. The high proficiency players however responded to the same task by a proportionately lesser increase in mouthpiece force, using less of the total force potentially available to them.

It is also interesting to note that, although players were able to significantly reduce the force they would normally use for a particular note, most showed great difficulty in sustaining the force reduction and reported that abnormal and excessive effort in breath control was necessary to achieve this.

IV ENDURANCE TEST

Each subject was required to play G_5 at 95 dB as a series of semibreves, (see figure 3.18). The subjects' instructions were to continue playing these notes without interruption until they were unable to sustain performance at the required intensity. The subjects performed the test exercise once only each, as the task was extremely tiring.

Figure 3.18 Endurance exercise. Mouthpiece force trace for subject A.K.



Three measures were obtained from each subject.

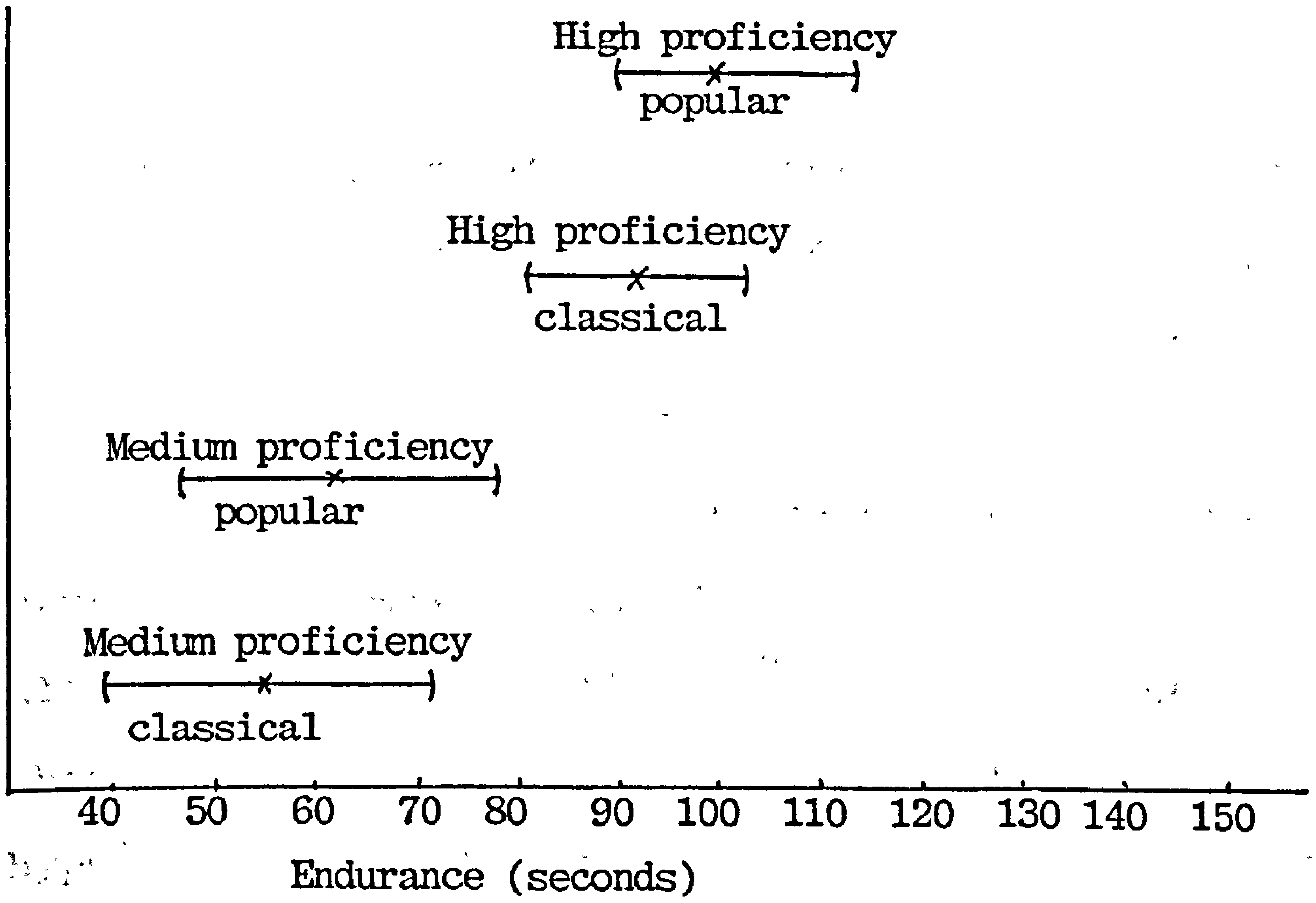
1) Endurance was measured as the length of time (seconds) that the subject was able to sustain the criterion intensity. 2) F^1: the force used to achieve the note at the beginning of the exercise. 3) F^2: the force used to achieve the note at the point where fatigue ultimately disrupted performance to criterion.

Results

Figure 3.18 shows a typical mouthpiece force trace for this test. This particular subject (A.K.) shows little variability of mouthpiece force throughout the task.

Figure 3.19 shows the means and interquartile ranges for each of the four sub-group's endurance measures. Mann Whitney U tests on the data show that the high proficiency players (n = 30) were able to sustain performance longer than the medium proficiency players (n = 30) ($z = 4.1$; $p < .001$), and also that the "popular" group were able to continue the exercise longer than the classical group ($z = 2.1$; $p < .05$).

Figure 3.19 Endurance capacity of the four performer sub-groups.



Mouthpiece Force

Firstly, a Wilcoxon matched pairs signed ranks test was carried out (n = 60) to compare the forces used at the beginning of the exercise with those applied at the end. The analysis turned out to be not significant at the .05 level. Inspection of the raw data showed that the relevant difference scores for each subject tended to be small and evenly distributed about zero. However, to check for the possibility of concealed heterogeneity within the data, the difference scores were collected into the four performer sub-groupings and became the four independent samples in a single factor analysis of variance. This yielded a non-significant result ($F3,56 = .31; p > .05$ n.s.), suggesting that no difference existed between proficiency levels or performer styles in the effect of fatigue on mouthpiece force in this test.

Also, a correlation of endurance (seconds) with the initial force level F^1 for the 60 subjects gave a Spearman rho of .102, which is well short of significance.

A further single factor analysis of variance was carried out comparing the four performer subgroups' ultimate force application (F^2). The non significance of this finding ($F = .21$ (df3,56); $p > .05$ n.s.) confirmed the hypothesis of no difference between proficiency and performance style levels in the absolute application of force in this exercise, which ties up this test with the findings of the main long note study (see study 3.I). This particular test was demanded by the fact that the endurance test occurred late on in the session, and fatigue effects may have been operating.

Discussion

The finding that there is no discernible change in mouthpiece force over a passage inducing acute fatigue is a very surprising finding which is quite contrary to the subjective report of most players. This will be discussed in more detail later, in conjunction with the stability of mouthpiece force.

Although the results show no difference in the case of instantaneous absolute levels of mouthpiece force between the referent groups, the results of the endurance test show clearly that performer variables are related to the total amounts of force used. The groups use similar force levels but demonstrate differences in the length of time they are able to withstand these forces. The correlation of endurance with mouthpiece force in this study shows quite clearly that within the limits of this experiment there appears to be no association between the instantaneous absolute level of force applied and the endurance that a player demonstrates.

V JUDGMENTAL TASKS

(i) The first test involved the active manipulation of mouthpiece force in four artificial performance conditions.

A) The subject forms the embouchure with breath support and presses the mouthpiece against the lips with the force the player judges he would have used for a semibreve G_5 at 95 dB, except that the subject does not actually play the note in this condition.

B) The subject forms the embouchure without breath support, and estimates his normal force for a semibreve G_5 at 95 dB, again without actually playing the note.

C) The subject neither forms the embouchure nor uses breath support, and simply estimates the normal force for the criterion note by pressing the mouthpiece against the relaxed lips.

D) The fourth condition which is non-judgmental involves the subject actually playing a semibreve G_5 at 95 dB.

Design

Each condition is administered twice for each subject, giving eight trials altogether. The order of presentation was determined by randomized counterbalancing. The criterion measure for each condition was the mean of the two observations of mouthpiece force.

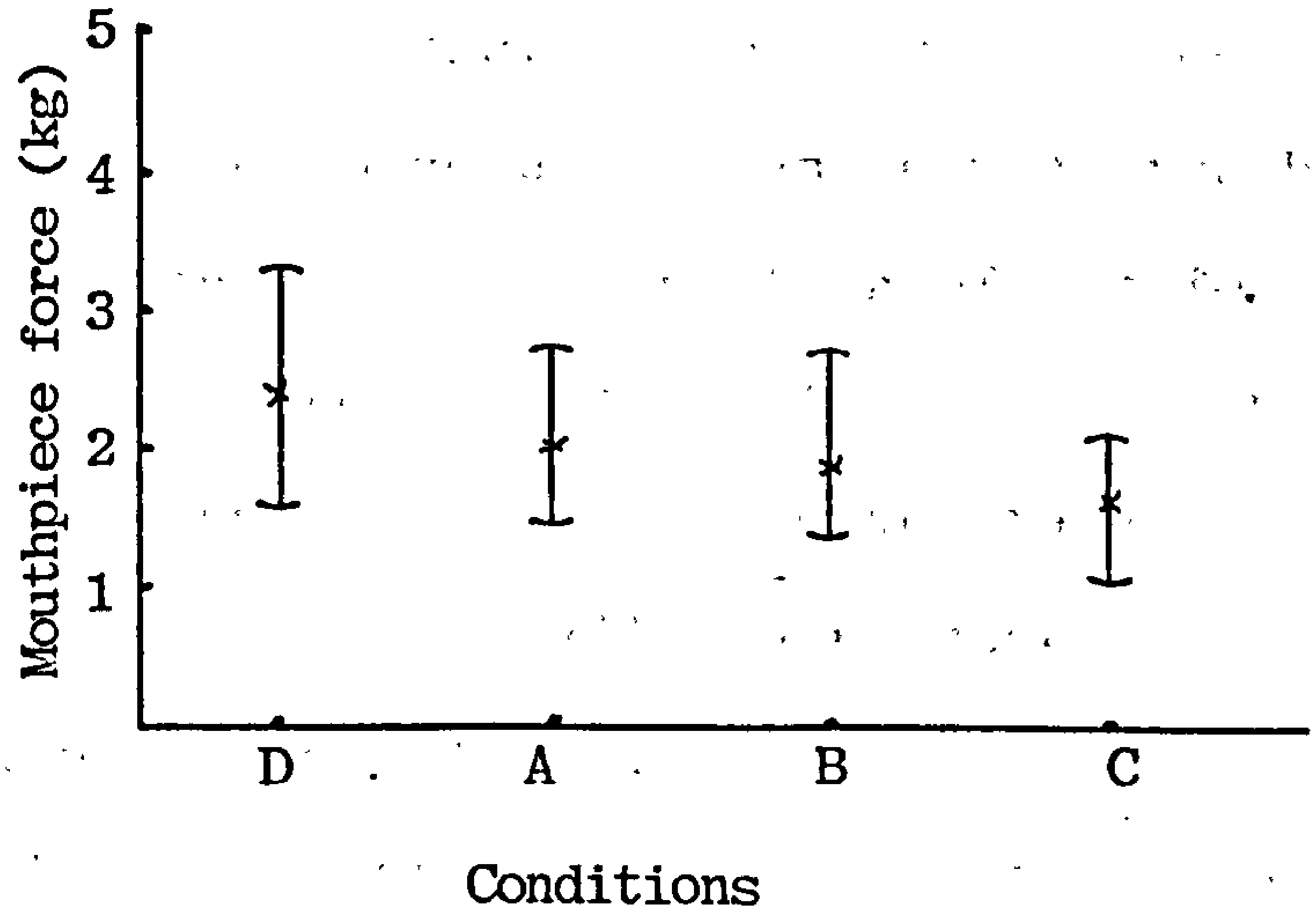
Hypothesis

Preliminary data from Weast and Hake's (1965) study suggest that the absolute judgment of force on the embouchure should produce decreasing estimates the further the judgmental conditions depart from normal trumpet performance. With the present four conditions, this would give a predicted ordering of D > A > B > C for the mean mouthpiece force recorded.

Results

Figure 3.20 gives the data plotted for the sixty subjects for the four conditions of the study.

Figure 3.20 Mouthpiece force judgments under artificial performance conditions.



As can be seen the means of the sixty subjects follow the predicted order. This composite observed order sums data from the individual subjects. To test the significance of this finding, the Jonkheree combined S test was applied, which compares every subject's individual order against the predicted order. With the normal approximation for large samples, the computed z value of 3.9 enables us to reject the null hypothesis at the .001 level of significance, and conclude that the forces applied in the quasi performance conditions are in accordance with the predicted order. The difference between the means given in figure 3.20 suggests that not necessarily every possible comparison between these means contributes to this overall result. Indeed, post hoc sign tests reveal that there is no evidence to reject the null hypothesis of no difference between conditions A and B. This is on the basis of a per comparison error rate of .017 which derives from a per experiment rate of .1 according to the reasoning of Leach (1979). However, all the other pairwise comparisons attain the chosen level of significance.

The "accuracy" of players' judgments of their normal mouthpiece force usage can be measured by calculating the deviation for each condition of the force produced from that demanded (the latter is of course given in condition D). This method of determining accuracy of sensory judgments has been used occasionally in psychophysical experiments (see for example Briggs,1968). This procedure gave three sets of difference scores representing the accuracy of the subjects for judgments under conditions A, B and C. The difference scores for the two proficiency levels were compared with Mann Whitney U tests for each condition. Table 3.19 shows that in each of the conditions A, B, and C the null hypothesis was confirmed, showing that in the circumstances of the current experiment, there was no evidence to suggest that any difference existed in the extent of the error involved in the judgment of mouthpiece force between the high and the medium proficiency players.

Table 3.19 Comparison of the accuracy of medium and high proficiency performers in the self-judgment of mouthpiece force.

	Mean "error" (kilograms)		z	p
	High Proficiency	Medium Proficiency		
A	0.48	0.32	0.824	n.s.
B	0.51	0.61	0.646	n.s.
C	0.74	0.85	0.481	n.s.

(ii) The second part of the judgment study involves the subject making a direct estimate of the force he is using. The subject is asked to play C_6 at 103 dB for four seconds. Immediately after this he is given two weights, one of one kg and one of five kg. The subjects are allowed to feel the weight of these until they are able to give a value (in kg) of the amount of force they judged they were using for the criterion note.

Results

The mean mouthpiece force actually used was 4.0 kg.

The mean of the estimates of the 60 subjects however was 1.2kg. A Wilcoxon matched pairs signed ranks showed this difference to be significant beyond the .0001 level, ($n = 60$; $z = 5.68$). The size of this judgment error was investigated using the Hodges Lehmann estimate, (Hodges & Lehmann, 1970). As this estimate uses the magnitude of the difference scores, it gives a powerful numerical indication as to the accuracy of the judgments of force. It turns out that this procedure gives the best estimate for the size of the average judgment error as 2.7kg and 2.9kg respectively for the high and the medium proficiency level groups. A Mann Whitney U test on the two sets of difference scores shows this small difference to be not significant, ($z = .62$; $p > .05$ n.s.). This shows that there is no evidence revealing any difference in the absolute accuracy of mouthpiece force self judgments between players of different proficiency levels.

In addition, the correlation between the actual forces used and the estimates given was calculated. The data from the 60 subjects gave a Spearman r_s of .12, which is well below significance.

Discussion

Study V (ii) reveals a systematic tendency of the performer to underestimate the absolute amount of mouthpiece force he is using. The errors are large and apparently unrelated to proficiency factors. It might be argued that the comparison weights chosen constituted extraneous contextual influences on the force judgments (see Poulton, 1968). It is unlikely that this was the case however, as the resulting mean judgment (1.2kg) is located at the lower extreme of the comparison range. Also, the explicit instructions to the subject attempted to minimise this effect. The correlation of judged force with actual force shows that, apart from the tendency to greatly underestimate the absolute force used, the individual judgments do not even bear a relative relation to the true applied force and reveals a startling inability of players, irrespective of proficiency level, to make meaningful judgments of their own mouthpiece force usage.

In study V (i) the three judgmental conditions involve the successive removal of selected elements from the sequence of individual components constituting trumpet performance. The analysis showed that the underestimates of actual mouthpiece force became more pronounced the further the conditions of judgment departed from the normal modality of performance. Trumpet playing is a skill like any other in that it involves the co-ordinated integration of sensory information and muscular responses. It may be that trumpet players have difficulty in isolating or abstracting the pure sensation of mouthpiece force from the interlocking complex of performance factors. This reasoning would account for the pattern of errors demonstrated in V (i), and also contribute to the tendency to underestimate the absolute force levels used in V (ii). This line of argument will be continued in the concluding discussion section of this chapter.

Stability of Mouthpiece Force

(i) Short Term Consistency

The consideration of short term changes in the application of mouthpiece force is very important for the present study. The main point of interest centres round the effects of fatigue, which is a particularly crucial aspect of brass playing. This phenomenon is therefore of interest in itself as well as being pertinent to the current analysis from the point of view of carry-over effects and experimental design. The latter methodological concern is especially important in view of the restrictions on counterbalancing strategies dictated by the relatively small number of subjects in the study.

Study IV looked at short term changes in mouthpiece force due to acute fatigue effects in a single extended endurance exercise. To investigate the change in force usage throughout the whole session, it was decided to evaluate the stability of the force associated with the occurrence of test pitch G_5 at 95 dB, as this was a test note used extensively throughout the whole study. Four measures were derived, each of which were means of the forces observed for each occurrence of the criterion note extracted from the data in studies I, III, IV and V. A fifth measure for each subject was obtained by a final performance of the criterion note at the very end of the experiment.

Results

The method to determine the stability in the use of mouthpiece force is derived from the computation of the reliability coefficient with the analysis of variance (see Kerlinger, 1973). With two proficiency levels (A1 and A2) and each subject giving five scores (factor B) for different occasions of testing, a two factor analysis of variance with repeated measures on one factor was carried out. The summary table is given at 3.20. The stability coefficient is given by the equation $r_{tt} = 1 - \frac{V_e}{V_{ind}}$ where V_e is the residual or error component and V_{ind} is the variance resulting from differences between individuals. The computed value of the stability coefficient is .98 which is extremely high.

Table 3.20 The effects of proficiency and occasions of testing on the short term stability of mouthpiece force.

Source	S.S.	d.f.	M.S.	F	p
A Proficiency levels	161.8	1	161.8	0.376	n.s.
Subjects within groups	24905.4	58	429.4	55.9	< .001
B Occasions of testing	59.8	4	14.9	1.97	n.s.
Interaction A B	40.7	4	10.2	1.32	n.s.
B x S within groups	1783.1	232	7.7		
Total	26951	299			

Discussion

The finding that the amount of force used by the individual subjects throughout the session does not demonstrably change can be seen from the noteworthy lack of a main effect associated with occasions of testing ($F_{4,232} = 1.97$; $p > .05$ n.s.) This result is not dependent on proficiency level, as can be seen from the lack of significance for this main effect ($F_{1,58} = .376$; $p > .05$ n.s.) The main source of systematic variance is that deriving from individual differences, and it is this which leads to the extremely high stability coefficient.

These findings run very contrary to subjective report. The experimental session is fairly long and most players' performance displayed some signs of fatigue. The fact that this is not reflected in mouthpiece force usage is very surprising. The implications of this will be discussed shortly.

(ii) Long Term Stability of Mouthpiece Force

A longitudinal investigation of mouthpiece force is described in this section. The period of time chosen over which the change of force usage was to be measured was fixed at six months. This type of study is very difficult to carry through owing to the problems with obtaining subjects at the required times. The potential problems with subject attrition were minimised by using only six subjects, taking care to select those players least likely to present availability problems with repeated testing. An unfortunate corollary of this was that all the subjects were from the medium proficiency level group.

The six subjects were each tested on six separate occasions with a period of exactly four weeks intervening between the repeated administrations of the test. Fortunately, only twice throughout the period of testing did the session occur more than one day away from the designated date, and these were only two and three days away respectively. The procedure for each session was an exact replication of the procedure for the long note study (see 3.I), including the warm-up time and the use of the sound level meter. From the long note data in the longitudinal sessions, the forces used for the criterion note G_5 at 95 dB were extracted for each subject.

Results

Figure 3.21 The stability of mouthpiece force over a six month period of testing.

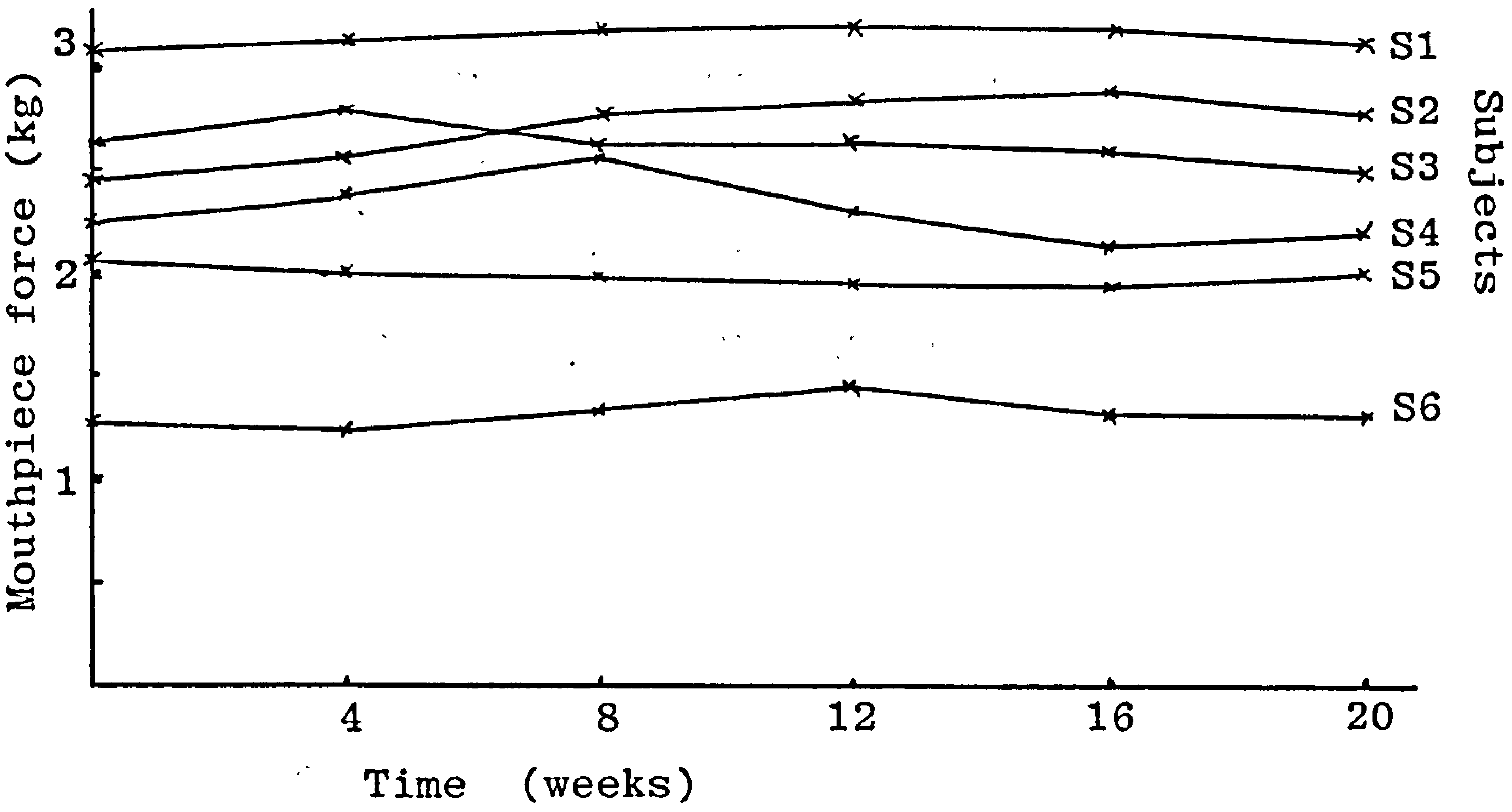


Figure 3.21 shows the data plotted for the means for each subject at the specified time intervals. A single factor analysis of variance with repeated measures was carried out with occasions of testing as the levels of the factor. Table 3.21 demonstrates the very high stability in the use of mouthpiece force over the six month period of investigation.

Table 3.21 Analysis of variance testing the stability over time of mouthpiece force.

Source	S.S.	d.f.	M.S.	F	p
Individuals	36.5	5	7.1	178	.001
Occasions of testing	0.096	5	0.02	0.5	n.s.
Residual	1.0	25	0.04		
Total	36.7	35			

The absence of a systematic effect for occasions of testing together with the larger variance due to individual differences results in a stability coefficient of .96. This is calculated according to the method in Kerlinger (1973).

Concluding Discussion

This section presents a review of the findings of this explanatory experimental study and considers some of the theoretical issues raised thereby.

The data show that mouthpiece force increases dramatically with ascending pitch, and to a marked but lesser degree with intensity. This accords with the previous studies of Henderson (1942), Weast and Hake (1965) and White (1972). The present study enhances and extends the validity of this finding because of the wider ranges of pitch and intensity employed. However, no support was found for the suggestion that proficiency is related to the observed change of mouthpiece force associated with given intensity changes (Weast and Hake), or pitch changes (Henderson). Although Weast and Hake did not report the intensity range they were actually using, which makes exact replication impossible, it is unlikely that any substantive effect was being observed. The small number of subjects they used and the lack of controls are likely to have resulted in what was no more than an experimental artifact.

The present study gives a valuable and highly detailed picture of the interactive relationship between pitch and intensity. Although White's study yielded data amenable to a highly particularised analysis, only a broad analysis of a very general nature was reported. (Notwithstanding that White's study was primarily concerned with electromyography, the omission was unfortunate).

As well as the pitch/intensity interaction, there were other interactive relationships in the data of the present study which had to be estimated indirectly by a multiplicity of tests rather than omnibus tests such as the analysis of variance. Sometimes it was possible to use the nonparametric combined S test to give a single overall demonstration of the presence of an interaction. Although it may seem unwieldy to apply a large number of single tests to this body of results, it was shown that such an exigency was dictated by the very nature of the data. There is never any statistical justification for using a particular analytical procedure purely on the basis of brevity or economy of presentation (Roscoe, 1975). In this study, omnibus tests were rejected when distributions were of questionable normality, and more especially where the variances of the sample scores were too heterogeneous.

Using many individual tests however raises the obvious argument that there will arise a number of significant results simply by chance (essentially by way of Type I errors). However, this criticism can be rebutted by observing that the very characteristics of the data which led to the rejection of the omnibus approach ensure that the alpha error effect will be of minimal importance. The reasoning is as follows. Many psychological experiments apply omnibus parametric tests without any contemplation of possible violations of their basic assumptions, because the type of datum involved ensures that such deviations will be of a minor order and hence safely ignored. For example, when comparing differences between treatment scores, the various treatment effects manifest themselves as finite increments or decrements from some actual or constructive "baseline" or "normal" score. Often this baseline takes the form of a designated control group integral to the analysis. This type of situation is one where there are unlikely to be vastly disparate variances between the treatment samples. However, the variable of mouthpiece force does not fit this model. There is no baseline score as such, and there is no conceptual or actual "control" group available. The factors of pitch and intensity combine in a direct and highly exclusive way to produce the mouthpiece force observed. This relationship between the variables had two important and interrelated consequences for the present experiment. Firstly there were large and highly consistent variance differences between the experimental pitch-intensity samples which it was decided were too large to ignore. Secondly, any random variation will only be of small effect when these are seen as being superimposed on the huge proportion of the variance taken up by the primary determinants of mouthpiece force in the context of this experiment. The possibility of type I errors are further diminished by the accuracy of the instrumentation and the replications of the experimental conditions within the design. With the post hoc analyses, the dangers of analytical error were minimised by using stringent per experiment error rates.

As far as interactive relationships are concerned, two pitch dependent effects appear most striking from the data. Firstly, it seems that equivalent increases in intensity involve greater changes in mouthpiece force the higher the overall pitch level. Secondly, fixed changes in pitch are associated with increasingly larger mouthpiece force differences the higher the general pitch at which the fixed change occurs. For example this means that to move up a whole tone, or to move from mezzo forte to forte involves progressively greater rates of change of mouthpiece force the higher in the register that these operations are performed. This means that in the upper register, a player has to contend with greater changes in mouthpiece force, as well as be subjected to the greatest overall levels which are normally applied here.

Despite the number of different analyses performed, and the differentiation between the referent performer groups in terms of proficiency and style, the long note study failed to find any consistent effects on the dependent variable of mouthpiece force. This was further demonstration of the invalidity of the "no-pressure" or "minimum pressure" (Farkas, 1962) theories of skilled performance. The value of the present study was in the uncovering of performer-proficiency and performer-style related effects as a result of the design of more complex tasks than those merely involving performance of isolated long notes. The requirements of general validity had demanded the inclusion of more complex items, and indeed it was only here where the interesting divergences in the performer attribute variables presented themselves, and the multiple test method of analysis paid off. White's study had used items of varying complexity, but then pooled all the data and led to a loss of power in many aspects of the analysis.

To begin with the present study was able to demonstrate the high degree of consistency in the application by performers of mouthpiece force across a wide variety of techniques and contexts. A note of given pitch and intensity tended to be performed at a similar mouthpiece force independent of the musical context. Although the high proficiency players tended to be <u>more</u> consistent than the medium proficiency group across different contexts, all players demonstrated a uniformity of force application which showed conclusively that mouthpiece force is not crudely applied in, and could quite properly be seen as a critical determinant of skilled trumpet performance. The detailed analysis of the flexibility exercise however revealed consistent if small departures from the simple "standard" force as given by the long note exercise. The lip flexibility exercise was a highly complex task, being one of the hardest techniques on the trumpet. In this study, there was a tendency for the players to avail themselves of greater amounts of mouthpiece force to achieve the notes of this complicated exercise than the correspondingly easier scale or arpeggio test. It was noteworthy that the higher proficiency performers, while still susceptible to this tendency, exhibited it to a lesser degree. Whether this effect is really due to complexity factors would need to be demonstrated by a further study which would employ test items specifically designed to represent different orders of complexity; the present study did not attempt any express manipulation of this variable. Similarly, the interpretation of the contextual analysis of mouthpiece force <u>within</u> the lip flexibility test would require replication and development. The post hoc nature of the interpretation of these aspects of the data was inevitable, this being an exploratory type of study, and was only intended to suggest some theories which might usefully be pursued in the future.

The tests involving the extreme use of mouthpiece force produced perhaps the most startling results, and amply justified the departure from the traditional methodology in this area. It was shown that performance in the extreme upper register was almost always associated with extremely large forces, and in as much as a player was able to play the upper notes at all, the absolute force levels were not related to the proficiency variables. However, it was seen that the maximum possible force available to the high proficiency players was greater than that which could be exerted by the medium players, suggesting that the high proficiency group have a greater tolerance for very large forces. This ties up with the endurance exercise, which showed that the better players had a greater tolerance in terms of forces exerted over time, as they were able to play the extended exercise for longer than the medium proficiency group.

We are now in a better position to describe the true nature of the high proficiency performer's use of mouthpiece force. He is someone who plays in the ordinary register with general force levels indistinguishable from his lesser skilled counterpart, but with greater consistency in its manipulation. But he has enormous reserves of mouthpiece force which he can and maybe must apply in order to play in the extreme upper register. Even in this extreme pitch region he will tend to keep some extra force in reserve, and does not indiscriminately apply the maximum possible force as a means to achieving notes in this register. (This is not of course to say he will not succumb to the expedient of applying such forces "in extremis" for example to ensure achieving the desired result in stressful or very important performance circumstances such as in the case of live broadcasts etc). Similarly, he will not react to tasks of increased complexity by resorting to an inordinate application of mouthpiece force. However, it is important to remember that for normal register performance, the striking feature is the lack of differentiation between the performer groups, the variance observed being almost exclusively attributable to the force patterns of individual performers, irrespective of proficiency or style.

The judgmental data provided some surprising findings. None of the measures in this study revealed any ability on the part of the individual performer to make accurate or meaningful judgments of the forces being applied by the performer himself, and this inability was found to be unrelated to performer proficiency or style. If this finding is correct it is of enormous importance from the point of view of teaching the trumpet, because it is difficult to see how a teacher can properly instruct a pupil in the use of mouthpiece force when he is unaware of the forces he is using himself. It is difficult to give a reason for why a player should be so unable to judge his own mouthpiece force. It is clear that trumpet performance involves profound physical changes in the metabolism during performance (see for example Faulkner and Horvarth (1967 a&b) , and it is possible that judgments of pure mouthpiece force are influenced by such factors as raised intracranial pressure, change in heart volume, blood gas content, circulatory effects etc. It may even be that a player judges his mouthpiece force from muscular fatigue in the arm! Some phenomenological aspects in the self judgment of mouthpiece force will be explored using psychophysical methods in chapter 5.

Before concluding this chapter a few words must be said in general terms about the validity of the experiment. This is necessary because the study does not use many of the strategies commonly employed in psychological investigation. To begin with the subjects were not deceived at the outset as to the nature of the experiment. Aside from the ethical consideration, the possibility of using non reactive and unobstrusive measurement was severely curtailed by the instrumentation itself, which was not amenable to concealment and could not be disguised to prevent the player from guessing its purpose. In any case, as debriefing was considered to be mandatory in fairness to the performers, it would be unrealistic to assume that players subsequently tested would be totally unaware of the general nature of the experiment. This is because the players were very likely to meet one another in the course of their musical lives. It may then be asked, what effect might such knowledge have on the outcome of the experiment? To avoid the bias with respect to the kind of "demand characteristics" that the studies of Orne (1962) revealed, the following procedures were employed; minimal communication with the subject in the original solicitation and prior to the running of the experiment, avoidance of explicit and implicit communication in the experiment proper (except in so far as to deliver the relevant instructions, which were all read verbatim from the same prepared text). The more likely source of bias would be what Rosenberg (1969) described as "evaluative apprehension". The error of motivational attitude here might be thought to produce spurious results. Although this question could only be definitively resolved by use of covert measurement techniques which are currently unavilable, it is considered to be very unlikely that the present study is impaired in this regard. This is for the very simple reason that, although there may be a tendency for players to want to "look good" by applying perhaps less mouthpiece force than they normally would, it is absurd to think that they would do this at the cost of "sounding bad"; musical performance itself is the ultimate yardstick by which any player is measured.

It might also be observed that blind scoring procedure was not used. This was considered unnecessary because the instrumentation used obviated the problems associated with subjective rating techniques. In any case, as an exploratory study, no hypotheses were directly being tested.

The above does not in any way suggest that these results would be identically replicated under different performance conditions. The generalizability of the results remains to be determined. As regards the question of population validity (Bracht and Glass 1968), the experiment is very high on this, because the study uses a large sample from an experimentally accessible population which is finite, small, and in this case not very different in actual size from the target population (at least as regards the top professional performers). The question of ecological validity is much more problematic. It is well known that performance of a musical instrument can be considerably affected by the context of that performance; for example the stress of public performance can easily enhance or destroy what has been satisfactorily accomplished in rehearsal. The next stage which might be attempted in research on mouthpiece force would hopefully be by way of a naturalistic "field study" of mouthpiece force during actual public performance. This would very likely reveal proficiency related effects of importance for teaching purposes, although the instrumentation would require some improvement and refinement to allow this to be attempted.

Certain observations can be made as a result of the present study. Although the data do not give much insight into the precise role of the skilled manipulation of mouthpiece force in the normal register, the extreme upper register is worth considering on its own. The simple question can be posed for each player; is it necessary to use extremely large forces (i.e. above 5kg) to achieve performance in this pitch region? If it is necessary then this has important consequences for the player. If he is physically unable to exert or withstand such forces, he will simply be unable to perform this register at all. Whether the ability to apply these forces is a result of a physical predisposition or whether training can give the required resistance is open to question. Briggs (1968) found evidence for a correlation between endurance and the oxyhaemoglobin saturation of the vibratory portion of the lips, and as we have seen, range and endurance on the trumpet are highly related. Research of this kind is not done very much and tends to be not very conclusive. There is in any case a reluctance to recognise research which suggests that there might be physical constraints on achievement. However, in view of the amount of effort which might be wasted by players who maybe do not possess the makeup to achieve reliable endurance or upper register performance, it would be foolish to ignore the possibility that some such physical restrictions exist. On the benefit side, it might be possible to identify those players at an early age who would be most likely to be successful and encourage their development, and allow those with less prospect of success to apply themselves to an instrument for which they do have more of an aptitude.

However, it might be pointed out that some of the players who were able to play in the altissimo register did play with force of less than 5kg. While this may give hope to players unable to exert the extreme forces we have observed, it must be appreciated that such players are exceptions to the rule and more often than not were insecure in performance in this region anyway.

Of course at this early stage it would be unwise to extend speculations as to the importance of these results <u>too</u> far, and the main purpose of the research done here is in suggesting areas in which future research might concentrate. This aspect will be covered in the concluding chapter.

CHAPTER FOUR

JUDGMENTS OF MOUTHPIECE FORCE ON THE BASIS OF VISUAL CUES

Introduction

The exploratory study (see chapter three) revealed the enormous range of general levels of mouthpiece force used by individual players. The phenomenological data which related the actual mouthpiece force used by the individual performer to the subjective self-judgments of that force revealed proficiency-independent errors of such magnitude that it was decided to devote a complete chapter to the investigation of this phenomenon (see study four, chapter five).

The third study however looks at another aspect of subjectivity, namely trumpet players' judgments of the mouthpiece force used by other players. Teachers have attached great importance to mouthpiece force, to the extent that the regulation of this variable has become an integral part of the training methods used (see for example Farkas 1962). Obviously, accuracy in the type of judgments being now considered is of crucial importance to the fulfilment of such training programs. This study therefore decided to explore the judgment of mouthpiece force in a quantitative way.

The study explores firstly the extent of judgmental agreement existing within different groups of subjects, each group representing a different level of proficiency on the trumpet. The judgment task was to rate the mouthpiece force displayed by various trumpet players in a series of photograph sets. This is a type of consensual analysis. Consensual analyses are commonly used where there is a judgmental task to be performed which involves a high degree of subjectivity. When subjective judgments have potentially important detrimental consequences, they become of urgent interest, especially to educational psychologists. In the marking of essays Valin (1961) looked at the degree of variability between judges and Akeju (1972) examined the problem of judge unreliability. Fiske (1975) has done similar work on adjudication reliability with musical judgments.

The second major approach in this study is to compare the subjective judgment of mouthpiece force with the actual mouthpiece force being used. This comparison between subjective and objective criteria is made possible by the accuracy of the transducer (see chapter two).

The third main aspect of the study seeks to account for the difference between the judments and the objective criterion. With respect to the agreement between judges, Siegel (1956) points out that if high levels of inter-judge agreement arise, this only implies that the subjects are using basically the same standards in judgment. Such consensual agreement is often used to provide a standard when there is no valid external criterion, as is often the case in the behavioural sciences. Significant concordance values may still occur however when all the judges are incorrect with respect to some external criterion and are using a "wrong" criterion consistently. This study intends to show that mouthpiece force judgments are made on the basis of a "wrong" criterion, and that this criterion is really the effort or strain which is perceived in the visual appearance of a trumpet player's performance. There seems to be a common tendency among trumpet players to believe that the more strain the performer appears to be exerting, the more mouthpiece force he must be applying. This belief is partly encouraged by the teachers of the trumpet who emphasise the necessity of relaxed performance to their pupils. This commonly found belief suggested the rationale for the present study.

Rationale

Mouthpiece force increases with ascending pitch (given a fixed sound intensity). This seems to occur with all trumpet players, irrespective of proficiency or style (see chapter three). Higher pitch levels involve greater sound energy, and it is likely that an individual player will need to use more effort the higher he plays. This increased effort is likely to be reflected in the strain which is apparent to observers. Hence, when observing an individual player playing notes of different pitches, a ranked judgment by "effort" will tend to be in agreement with the actual mouthpiece force used. Thus, a good inter-judge association index will result if this wrong criterion is consistently applied.

However, when judging mouthpiece force between different players, a problem arises. Although for each player mouthpiece force increases with pitch, this variability is overshadowed by the variability between the individual players' general levels of mouthpiece force usage. In the case of judgments between players, a player showing less strain than another may easily be using more mouthpiece force. In this case, if apparent effort is being used as the "wrong" criterion, there may be consensual agreement between the judges, but the correctness of the judgments with respect to actual mouthpiece force will be minimal unless the effort displayed happens to correlate with the players' general levels of mouthpiece force application anyway (this possibility was itself investigated).

This study concentrates exclusively on visual cues in the assessment of mouthpiece force.. To isolate the visual cues and control for other extraneous factors, the judgmental study was conducted by having groups of subjects rate photographs of trumpet players taken during performance with the transducer in position. There were sixteen sets of photographs with four photographs in each set (see photos 4.1 - 4.8). The within player condition required the subject to make comparisons, a set at a time, within each of twelve sets of photographs, each set being of the same player playing four criterion pitches; thus the subject provided judgments on twelve different players. The between player condition required the subject once more to make comparisons between four photographs at a time, only this time the four pictures were of different players playing the same note, and there were only four sets in all for this condition.

Separate indices of concordance (for the consensual factor) and correlation (for the accuracy or correctness factor) were calculated for each of the sixteen sets of four photographs. The value of using so many different players as photographic models and calculating multiple measures from these lies in the generality of the conclusions that can be drawn from the data. Fiske (1975) in comparing reliability coefficients between groups in the ratings of student trumpet performances uses this principle to good effect. We are now in a position to state our main hypotheses.

Hypotheses

I There will be more consensual agreement within each group of subjects with respect to both judgments by apparent effort and by apparent mouthpiece force for within player judgments than for between player judgments.

II The ordinal judgments by apparent effort are essentially the same as those by apparent mouthpiece force for both the within player and the between player judgments.

III The statistical association of the judgments based on apparent effort and those based on apparent mouthpiece force with the objective criterion (actual mouthpiece force) will be high for the within player judgments and negligible for the between player judgments.

We will also be interested to see if there is any relationship between subject proficiency level and the indices of consensus and accuracy. Also, individual photograph sets will be examined to see if there are any particular trends.

Photographs 4.1 - 4.8
The photographic material used in this experiment is exemplified in the four following pages (pages 184A to 186). The photographs are as follows;

BETWEEN PLAYER JUDGMENT
PHOTO SET II

page 184A

4.1

Subject T.G.
Professional
Classical
Performing E_6
Mouthpiece force 6kg

page 185

4.2

Subject D.B.
Medium proficiency
Classical
Performing E_6
Mouthpiece force 3.1 kg

4.3

Subject J.B.D.
Medium proficiency
Popular
Performing E_6
Mouthpiece force 1.8kg

4.4

Subject B.A.
Professional
Popular
Performing E_6
Mouthpiece force 5.2kg

WITHIN PLAYER JUDGMENT
PHOTO SET D
SUBJECT N.B.

page 185A

4.5

Performing C_4

Mouthpiece force .5kg

page 186

4.6

Performing G_5

Mouthpiece force 1.2kg

4.7

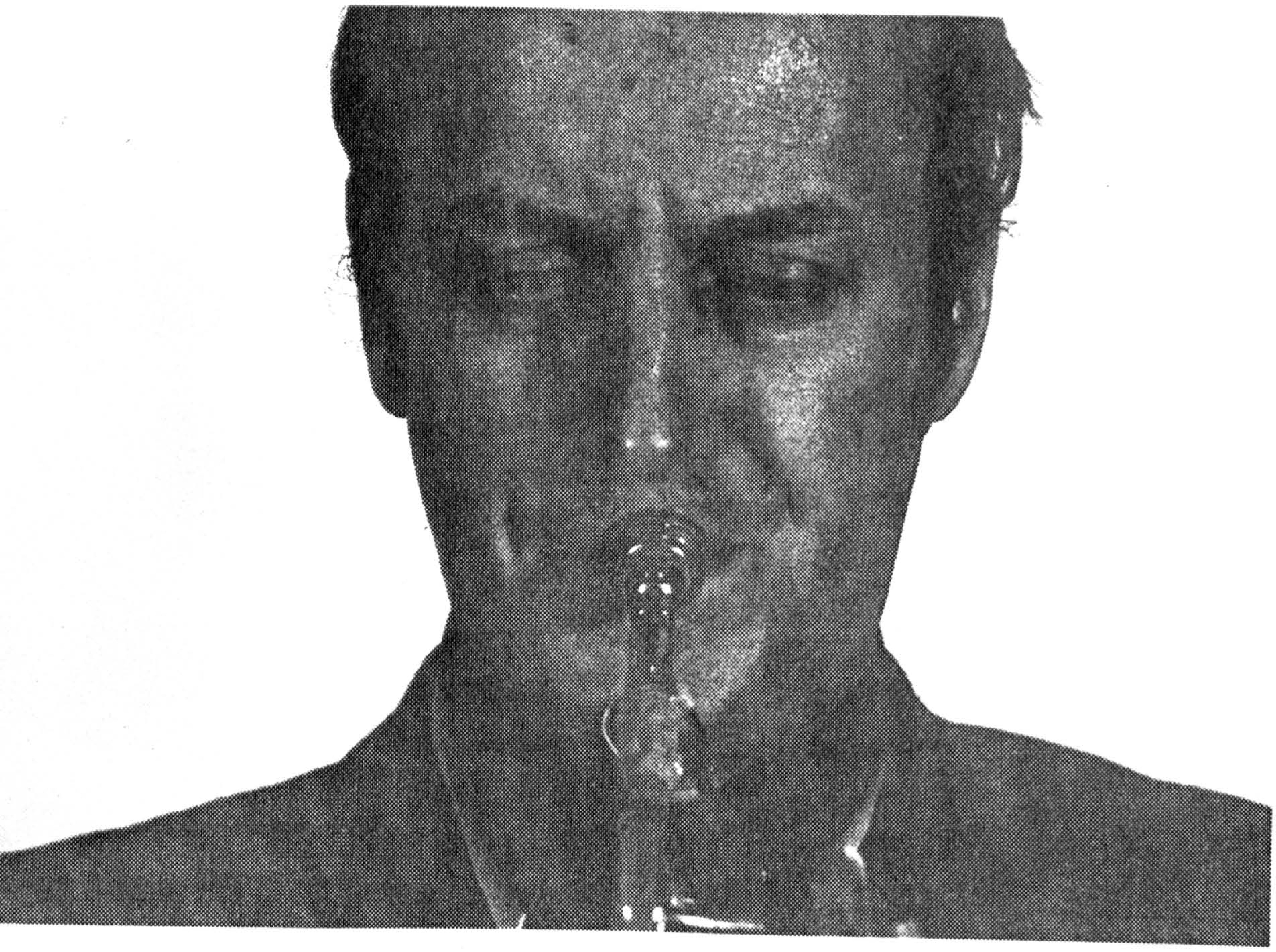
Performing G_4

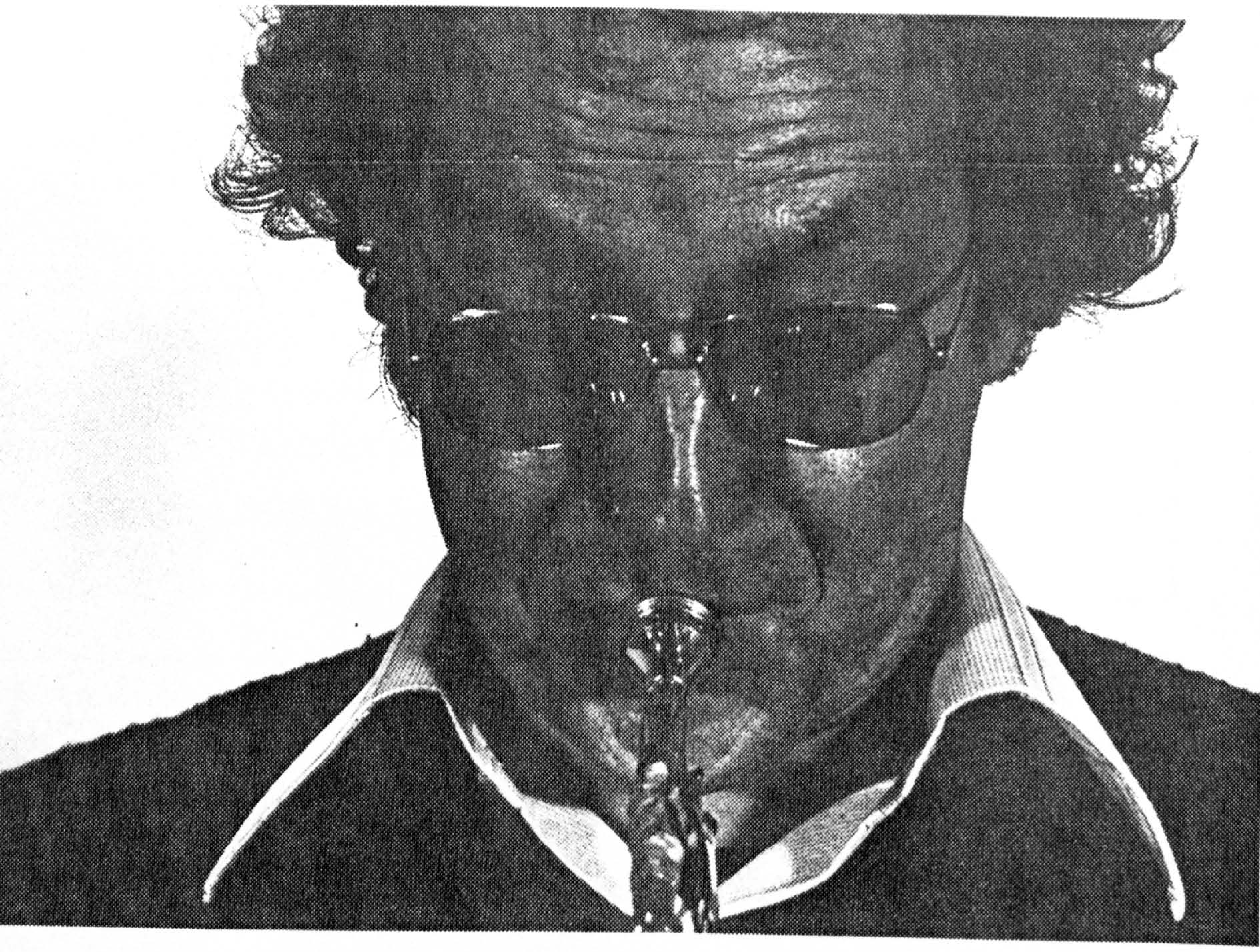
Mouthpiece force .8kg

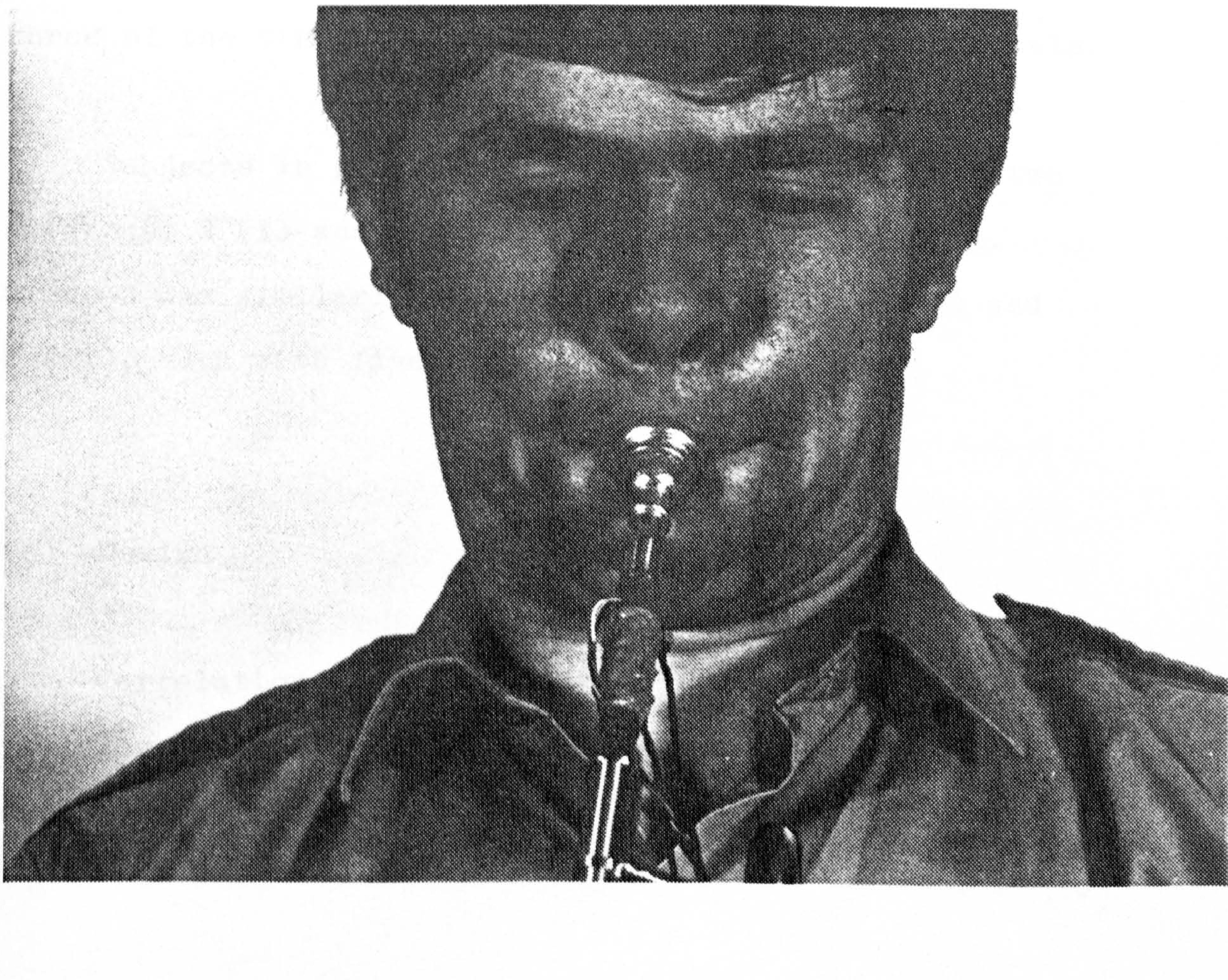
4.8

Performing E_6

Mouthpiece force 5.2kg







Method

(a) Subjects

Thirty subjects were involved in this study. These formed three main groups, each with ten subjects, representing three different levels of performance and experience with the trumpet. Group 1 were non trumpet players with little experience of practical music making. Group 2 were amateur trumpet players of intermediate standard. Group 3 were top professional performers with several years teaching experience in recognised music colleges.

No subject was tested who recognised any more than three of the trumpet players used as photographic models.

Subjects in group 2 were divided randomly into two subgroups 2 (i) and 2 (ii), each with five subjects. Group 3 was similarly divided into subgroups 3 (i) and 3 (ii), each with five subjects.

(b) Design

Correlational data form the basis of this study, yielding several different measures or indices for the experimental analysis.

Subjects in groups 1, 2(i) and 3(i) rank four photographs at a time, according to the amount of "effort" they judge as being displayed by the performer in each photograph. There are sixteen sets with four photographs in each set. There are two series of these four photograph sets; series A-N (twelve sets of within player photos) and series I-IV (four sets of between player photos). Photographs 4.1 - 4.8 show examples of the two different types of judgment conditions. Subjects in groups 2(ii) and 3(ii) perform rankings on the same photograph sets, only this time according to the amount of mouthpiece force which they judge as being displayed in the photographs. The order of presentation of series A-N and I-IV is counterbalanced in each group of judges. Within each series itself, the photo sets were randomly presented to each subject for successive ranking.

The design yields two types of index. Each individual photo set in A-N and I-IV has a measure of agreement between observers in the ranking of that set. This is the consensual index. Each particular photo set also has a measure of the agreement between the judged ranking and the objective criterion. This is a measure of the correctness of the judgment. These several individual indices of consensus and correctness then became the experimental independent variables by way of their operational definition. The dependent variables are either the proficiency levels of the subject or the two photograph categories (within player and between player) depending on the hypothesis being considered.

(c) Apparatus and Materials

Each photograph in the sixteen sets of four was a black and white print of 5" x 7" and of fixed contrast and quality (Ilford FP4, rated at 100 ASA with compensatory development in Perceptol). At the time each photograph was taken a record of the precise mouthpiece force associated with that picture was made. The within player series A-N comprised twelve different photograph sets. Each set was composed of the same trumpet player playing different notes (namely C_4, G_4, G_5 and E_6, all at the same intensity (mezzoforte). The between player series I-IV comprised four sets, each set consisting of four different trumpet players playing the same note (E_6) at the same intensity as the series A-N. Of the players pictured in the photographs, half were professional players, the rest of medium proficiency, and there was a good mix of performance styles. The four sets of between player photos were composed as follows. The sixteen photos of different players playing E_6 were divided into four lots according to the mouthpiece force associated with them. Group (i) was the four photographs with the four highest associated force readings. Group (ii) was the next four highest. Group (iii) contained the photos of the 9th to the 12th highest forces and group (iv) contained the lowest users of force at the given pitch. Photo set I was put together by drawing at random one photograph from each group, as was set II and set III, leaving the final four photographs for set IV. In this way it was ensured that there was a good variability of mouthpiece forces used in the between player series. This was necessary to avoid the possibility of biasing the selection of photographs in favour of the hypotheses.

(d) Procedure

The subjects were all tested individually. At the beginning of the session, the instructions were given. These instructions can be found in the appendix (see appendix 4.1). No ties were permitted in the rankings, and no time limit was imposed. Each photograph set was shuffled thoroughly by the experimenter before being placed on the table for ranking.

Results

Table 4.1 gives the concordance values for each group and subgroup in the ratings of the photograph series along with the significance value of each. Kendall's coefficient of concordance W expresses the degree of association among several rankings (see Kendall, 1970), and thus indicates the extent of agreement or concordance in the rank judgments. Hays (1963) points out that the W value is somewhat hard to interpret directly in terms of the tendency for the rankings to agree, although it is more comparable across different sets of data than is the average r_s (the average correlation over all possible pairs of rank orders), which is another measure of agreement often used. Also, the linearity of the statistic is another advantage. This is especially important in view of the subsequent use to which the concordance values are put in this study.

From table 4.1 it can be seen that for photo series A-N (the within player series), the concordance values are highly significant in all groups and subgroups for nearly all the photograph sets, showing that a consistently high level of agreement exists in the ranking of the series both by apparent effort and by apparent mouthpiece force. For the series I-IV (the between player judgments) the concordance values appear to be lower than the A-N series, There are fewer highly significant values (i.e. where $p < .01$)

and more $p < .05$ significance levels. Although some of the concordance values were not significant, most of those that failed were only just below the borderline. This shows that there is still a definite consistency in the judgments of the between player series, but that the subjects appear to show less unanimity than with the within player condition. This general drop in concordance values can be demonstrated with a Mann Whitney U test for independent samples. Independence in this case is not achieved by having different subjects in two samples as is usually the case. In this analysis the groups are concordance figures derived from subjects whose data contributes to both samples. The data is independent in the sense that the samples are based on discrete and independent judgment tasks, namely judgments within and judgments between players. As there is no reason to think that any judgment of series A-N would interfere with judgments of series I-IV, it was permissible therefore to use the proposed test. Table 4.1 shows that the drop in concordance values for all the groups between the respective judgment tasks is significant ($p < .05$ in each case). Thus it can be said that the indices expressing agreement between observers are higher for both judgments by apparent effort and by apparent mouthpiece force when judging within a player than between players. This gives very strong support to the first hypothesis.

A closer examination of table 4.1 reveals a strong resemblance between the concordance values for judgments

Table 4.1 A comparison of between player and within player judgments for the groups and subgroups of this study. W = concordance

			Within player judgments												Between player judgments				Mann Whitney A-N vs I-IV
			A	B	C	D	E	F	G	H	J	K	M	N	I	II	III	IV	p
Group 1		W	.764	.916	.160	.904	.628	.900	.916	1.0	.904	1.0	.844	.072	.612	.444	.476	.300	< .05
		p	.01	.01	n.s.	.01	.01	.01	.01	.01	.01	.01	.01	n.s.	.01	.01	.01	.05	
Group 2	(i)	W	.808	.936	.20	.936	.648	.712	.856	1.0	.936	1.0	.792	.232	.424	.52	.456	.552	< .05
		p	.01	.01	n.s.	.01	.01	.01	.01	.01	.01	.01	.01	n.s.	n.s.	.05	n.s.	.05	
	(ii)	W	.808	.648	.024	.904	.712	.664	.856	.936	.808	1.0	.936	.20	.408	.472	.488	.472	< .05
		p	.01	.01	n.s.	.01	.01	.01	.01	.01	.01	.01	.01	n.s.	n.s.	n.s.	.05	n.s.	
Group 3	(i)	W	.728	.872	.408	.936	.648	.904	.776	1.0	.904	1.0	.744	.072	.392	.296	.472	.424	< .025
		p	.01	.01	n.s.	.01	.01	.01	.01	.01	.01	.01	.01	n.s.	n.s.	n.s.	n.s.	n.s.	
	(ii)	W	.936	.776	.40	.872	.808	.728	.792	.904	.801	1.0	.936	.392	.392	.680	.488	.408	< .025
		p	.01	.01	n.s.	.01	.01	.01	.01	.01	.01	.01	.01	n.s.	n.s.	.01	n.s.	n.s.	
2(ii)& 3(ii) combined		W	.852	.708	.092	.818	.708	.742	.808	.904	.808	1.0	.936	.268	.348	.556	.488	.40	< .05
		p	.01	.01	n.s.	.01	.01	.01	.01	.01	.01	.01	.01	.05	.05	.01	.01	.01	
2(i) &3(i) combined		W	.764	.900	.252	.926	.648	.800	.812	1.0	.916	1.0	.74	.112	.348	.372	.46	.444	< .05
		p	.01	.01	n.s.	.01	.01	.01	.01	.01	.01	.01	.01	n.s.	.05	.01	.01	.01	

Where groups 1 (n = 10), 2(i) (n = 5) and 3(i) (n = 5) rank by apparent <u>effort</u>
groups 2(ii) (n = 5) and 3(ii) (n = 5) rank by apparent <u>mouthpiece force</u>

by apparent effort and by apparent mouthpiece force. This similarity is underlined by statistical comparison of the subgroups. Comparing subgroup 2(i) with subgroup 2(ii) we have data in two samples which are related in that there are pairs of concordance values for each photograph set, one value for judgments by effort and one value for mouthpiece force judgments. This pairing yields difference scores which can be meaningfully ordered in terms of absolute magnitude (the linearity of the concordance statistic has already been mentioned), and then subjected to the Wilcoxon matched pairs signed ranks test (Siegel,1956). This analysis shows that the null hypothesis of no difference between the indices for effort and mouthpiece force cannot be rejected at the usual level of significance ($p < .05$) in a two tail test. A similar analysis between the high proficiency subgroups 3(i) and 3(ii) similarly failed to demonstrate significance.

The fact that there is no demonstrable difference between the effort and mouthpiece force concordance indices does not logically imply that the judgments by the two criteria are essentially similar. Firstly, there is the danger of a type II error; a finding of no significance does not permit the conclusion to be drawn that the samples are necessarily similar. Secondly, identical values of W can result from two separate groups ranking the same four photographs in completely different ways. W only measures the consistency within each group.

In order to demonstrate the similarity between the rankings by effort and mouthpiece force , the statistical procedure needed is one testing goodness of fit. Such tests usually illustrate the degree of association between the distribution of a set of sample values (observed scores) and some specified theoretical distribution, such as the binomial or normal distribution. In this case we are comparing two empirically observed distributions, treating apparent mouthpiece force judgments as the expected frequency category and the effort judgments as the observed frequency category, in line with the hypothesis that these two types of judgment yield the same rankings.

The only feasible way to go about using the goodness of fit test is by looking at individual photograph preference positions. The first preference position is the first rank in the judgment task, being the photograph which the subject judged to display the most mouthpiece force or effort. The second is the photograph ranking for the next highest mouthpiece force or effort and so on until the fourth preference position, which should be occupied by the photograph ranked lowest on the criterion of judgment. At each preference position, there are four categories of possible scores for both effort and mouthpiece force judgments. Each category contains the sum of a number of discrete and independent events, namely the selection of a particular photograph from four possible choices. For the within player condition the four choices are photographs of different pitches being played (C_4, G_4, G_5 and E_6), and for the between player judgments the four choices are of four different players playing the same note (E_6). The goodness of fit test is carried out at each preference position, and compares the distribution of choices of particular photographs between the effort rating group and the mouthpiece force rating group. If the degree of association between the effort and force rankings is high for each preference position with respect to the particular photographs assigned to them, this would necessarily suggest that the judgment orderings themselves are essentially similar.

There are certain restrictions on the use of the χ^2 goodness of fit tests. In particular, when there are small expected frequencies, not more than 20% of the expected frequencies may be smaller than 5, nor any observed frequency smaller than 1 (see Roscoe, 1975). For this reason, the data was collapsed across photograph series, and where necessary across categories of individual photograph choices.

Table 4.2 shows the results of the goodness of fit tests applied to the data. It can be seen from this

Table 4.2 Results of the goodness of fit tests comparing the distributions of the judgments by effort and by mouthpiece force.

Within player series A-N

Preference position	χ^2	df	Prob. under H_o that $\chi^2 \geq$ CHI SQ
1	0.18	2	.90-.95
2	0.62	3	.80-.90
3	0.81	3	.80-.90
4	1.30	3	.70-.80

Between player series I-IV

Preference position	χ^2	df	Prob. under H_o that $\chi^2 \geq$ CHI SQ
1	0.34	2	.80-.90
2	0.48	2	.70-.80
3	1.41	3	.70-.80
4	1.24	3	.70-.80

that no grounds exist for rejecting the null hypothesis, and that the rank orderings by effort and by mouthpiece force are essentially similar. Thus, the second hypothesis receives strong support.

As explained in the introduction to this chapter, the high consensual agreement implied by the general significance of the W values is not synonymous with an objective ordering. To compare these logically distinct concepts empirically, it is necessary to determine the degree of association between the rank judgments of the subjects and the ranking of the known mouthpiece forces. It is not possible to use a simple correlation such as the Spearman rank correlation coefficient r_s because this would only apply to correlations between two sets of ranks. What is required is a statistic which represents a measure of the extent to which several sets of ranks agree with one criterion ranking. This would give an indication of the extent to which the subjects' ordering of mouthpiece force is "correct" and not merely consensually consistent. It would be possible to perform a correlation on each ranking and then average the coefficient over several subjects. With only four ranks however, this would give a result which would be extremely hard to interpret, and would not be easily testable for significance. For this reason it was decided to use the combined S test (Joncheere,1954). This combines data from a number of subjects where each individual subject's data may theoretically be analysed with Kendall's S statistic (Kendall,1970). The sum of the individual S values provides the test statistic for the overall summary statement of the results. Table 4.3 shows the results of the combined S test. The value r is derived from the combined S total to convert the correlation into the standard scale from -1 to +1 (for this transformation see Siegel,1956). The significance value is for an upper tail test as our hypothesis predicts the ordering of the judgments. For the groups rating

Table 4.3 Correlations between actual mouthpiece force and the rankings of effort and mouthpiece force judgments for the different experimental groups.

		Within player judgments																Mann Whitney A-N vs I-IV
		A	B	C	D	E	F	G	H	J	K	M	N	I	II	III	IV	p
Group 1	r	.87	.90	.26	.87	.70	.93	.90	1.0	.87	1.0	.80	.2	-.13	.03	-.06	.03	< .001
	p	.001	.001	n.s.	.001	.001	.001	.001	.001	.001	.001	.001	n.s.	n.s.				
Group 2 (i)	r	.80	.93	.33	.93	.73	.73	.87	1.0	.93	1.0	.733	.26	0	.2	-.2	-.2	< .001
	p	.001	.001	.05	.001	.001	.001	.001	.001	.001	.001	.001	n.s.	n.s.				
Group 2 (ii)	r	.87	.73	.06	.80	.73	.67	.87	.93	.80	1.0	.931	.13	-.13	.33	0	.13	< .01
	p	.001	.001	n.s.	.001	.001	.001	.001	.001	.001	.001	.001	n.s.	n.s.				
Group 3 (i)	r	.73	.87	.47	.93	.73	.87	.80	1.0	.87	1.0	.80	.20	-.13	.13	.06	.2	< .01
	p	.001	.001	.005	.001	.001	.001	.001	.001	.001	.001	.001	n.s.	n.s.				
Group 3 (ii)	r	.93	.80	.33	.86	.80	.80	.80	.82	.81	1.0	.93	.26	.13	.26	0	-.13	< .01
	p	.001	.001	.005	.001	.001	.001	.001	.001	.001	.001	.001	n.s.	n.s.				
2(ii)&3(ii)	r	.90	.77	.20	.87	.76	.73	.83	.86	.80	1.0	.93	.20	0	.24	0	0	< .01
	p	.001	.001	.01	.001	.001	.001	.001	.001	.001	.001	.001	.01	n.s.				
2(i) & 3(i)	r	.76	.90	.40	.93	.73	.80	.83	1.0	.90	1.0	.77	.23	-.06	.16	-.06	0	< .001
	p	.001	.001	.005	.001	.001	.001	.001	.001	.001	.001	.001	n.s.	n.s.				

mouthpiece force (groups 2(ii), 3(ii) and 2(ii) & 3(ii) combined), the agreement with the objective criterion is very good for 10 of the 12 photographs in the series A-N, with r very highly significant. For series I-IV however the agreement with the objective criterion falls to around zero in nearly all the cells. This suggests strongly that the rankings according to mouthpiece force only agree with the objective criterion when judging within players and not when judging between players.

Further demonstration that mouthpiece force judgments on a visual basis are really judgments on an effort or strain basis is given by the association of effort judgments with the actual mouthpiece force criterion. Table 4.3 shows an identical pattern of results for the effort groups (groups 1, 2(i), 3(i) and 2(i) & 3(i) combined), with highly significant r values for series A-N dropping to negligible values for the I-IV series. The difference in all groups between the r values for the within player (A-N) and between player (I-IV) conditions were compared statistically by means of the Mann Whitney U test. Table 4.3 shows the difference in each group and subgroup to be highly significant. The third hypothesis is therefore accepted.

Table 4.4 combines the judgmental data from the subgroups to give three groups differing by proficiency, with ten subjects in each group. This enabled an analysis

Table 4.4 Judgmental data collated into the three main performer groups.

		Within player judgments A-N															
		A	B	C	D	E	F	G	H	J	K	M	N	I	II	III	IV
Group 1	r	.87	.90	.26	.87	.70	.93	.90	1.0	.87	1.0	.80	.20	-.13	.03	-.06	.03
	W	.76	.91	.16	.90	.62	.90	.91	1.0	.90	1.0	.84	.7	.61	.44	.47	.30
Group 2	r	.83	.83	.20	.86	.73	.70	.86	.96	.86	1.0	.83	.20	-.06	.26	.10	-.03
	W	.77	.78	.06	.90	.67	.66	.85	.94	.85	1.0	.85	.09	.41	.49	.46	.48
Group 3	r	.83	.83	.40	.90	.76	.83	.80	.90	.83	1.0	.86	.23	0	.4	.03	.03
	W	.83	.82	.30	.90	.71	.83	.76	.92	.82	1.0	.83	.14	.39	.41	.48	.37

of proficiency related effects to be undertaken. Friedman analyses of variance were made for series A-N and I-IV separately, for both the concordance indices and the indices of association with the objective criterion. Table 4.5 shows the results of the analyses. None of the χ^2 values was significant, so it was not possible to reject the null hypothesis of no difference between the three groups' concordance indices or correctness measures. This obviated the need to look for specific differences between any of the groups individually.

Four further Friedman analyses of variance were made to test for any differences between individual photograph sets themselves. Table 4.6 shows that for series I-IV the χ^2 value failed to reach the critical significance value for the W or the r data. For series A-N however, there was a highly significant difference in concordance ($p < .001$) and correctness ($p < .001$). Having rejected the null hypothesis, some multiple comparisons were made to locate the effects. Post hoc tests were made on two particular photo sets, C and N, which seemed to present the most difficulty to the subjects in both types of ranking tasks. The analysis compared these individual photo sets with the rest of the sets in the A-N series with respect to the W and r indices. Both photograph sets were found to have significantly lower W and r values than the other sets in the series.

Table 4.5 Friedman analyses of variance comparing consensus and correctness indices for the judgmental data across the three proficiency groups.

Photo series	Index	χ^2	d.f.	p
A-N	r	2.75	2	n.s.
I-IV	r	2.38	2	n.s.
A-N	W	0.61	2	n.s.
I-IV	W	0.52	2	n.s.

Table 4.6 Friedman analyses of variance comparing the difference between the associated concordance and correctness indices of the individual photograph sets.

Photo series	Index	χ^2	d.f.	p
A-N	r	26.3	11	$<.001$
I-IV	r	4.9	3	n.s.
A-N	W	27.4	11	$<.001$
I-IV	W	2.2	3	n.s.

This finding was on the basis of a fairly stringent post hoc significance level ($p < .001$), and prevented any other comparisons achieving significance.

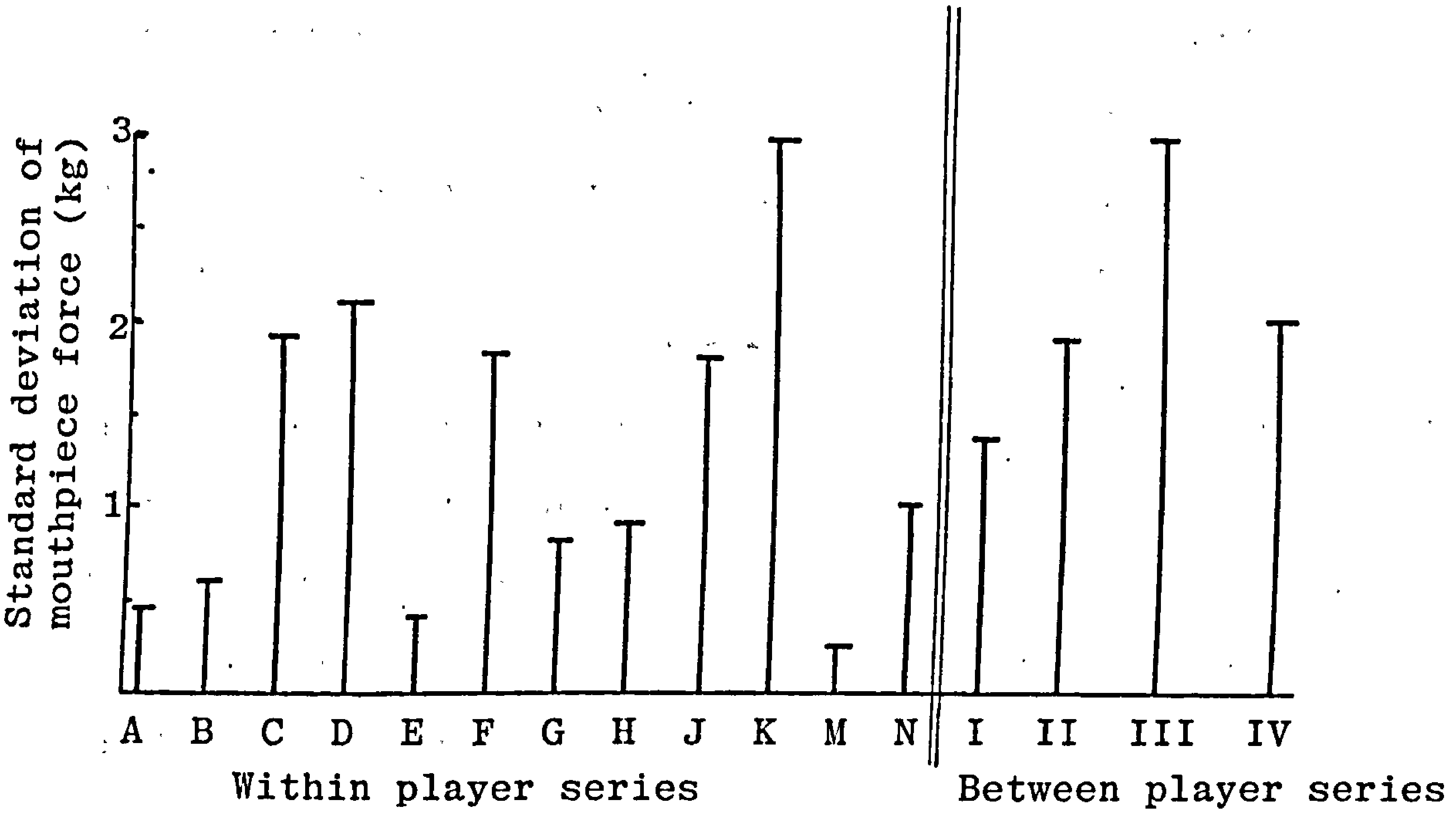
Discussion

The results, taken together provide good support for our hypotheses. However, several difficulties need to be considered.

Firstly, it could be argued that the method of selection of photographs for the between player series I-IV might have produced spuriously favourable results. In particular, if the variance of the actual mouthpiece force used by the between player photographic models happened to be significantly less than the variance of the mouthpiece force of the within player photographic models, the concordance indices could have been influenced in favour of the first hypothesis. To prevent any suggestion of this, the standard deviations of the relevant photographic mouthpiece force was calculated for each set. Figure 4.1 shows the results of the calculations. It can be seen that the variance seems to be greater for the between player series. This finding lends even stronger confirmation to the hypothesis.

Another problem to be considered is the

Figure 4.1 Variance of the mouthpiece force associated with each photograph set in series A-N and I-IV



relatively small number of subjects in the study. The strict criteria for inclusion in the highest proficiency group restricted the sample size here, especially as it was hard to find many professional players who were acquainted with less than three of the photographic models, many of whom were well known players. It is difficult to see how this particular problem can be circumvented.

Having shown that judgments of mouthpiece force tend to be made on the basis of the irrelevant and misleading criterion of apparent effort or strain, it is germane to ask why these types of errors occur, irrespective as they are of experience or skill on the instrument.

There are several possible explanations for what appears to be happening, but perhaps the best one relates to a familiar simple misconception. This is that there is a tendency to equate relaxation and apparent effortlessness with a lack of exertion. This occurs with any type of physical activity. For example in athletics, it is common for an athlete to look remarkably tranquil after a sub four minute mile, much more tranquil in fact than an average person running perhaps twenty yards for the bus! There is no doubt however who is expending more energy, and who is in the state of greater <u>actual</u> exertion. The state of their physical appearance relates to another phenomenon, namely the <u>reaction</u> to the physical activity undertaken. It is by this process that a deceptively calm exterior can mask quite extreme underlying organismic activity. The process of adaptation to such extreme physiological conditions by experienced players is only one of the possible reasons for the evident ease with which they often seem to perform. Equally important might be the pet teaching idea of the ideal of completely relaxed performance. Furthermore, many musicians are very concerned to appear to be in full control of their performance, and may suppress any outward and embarrassing evidence of strain and discomfort.

Altogether, a somewhat false picture results which has drastic implications for teaching. In chapter three we saw how almost all expert players of necessity have to withstand extremely large amounts of mouthpiece force, both from the point of view of both endurance capability and high register playing. Teaching therefore should properly be directed towards enabling the pupil to apply and withstand these necessarily large forces. Players who evince strain in performance are likely to have been encouraged by their teachers to reduce mouthpiece force usage on the erroneous assumption that they are already using too much. This research shows that it is possible that such players are in reality not able to exert and withstand enough mouthpiece force. Quite how the large forces are to be applied without detriment to performance remains to be discovered.

CHAPTER FIVE

A PSYCHOPHYSICAL INVESTIGATION OF MOUTHPIECE FORCE

Introduction

(1) Background

The previous study (see chapter 4) explored judgments by individuals of other individuals' use of mouthpiece force. This chapter is concerned with the more phenomenological question of an individual's judgment of his own use of mouthpiece force.

This again involves the determination of the nature of the relationship between two sets of variables one physical and one psychological. The correlation between psychological attributes and physical dimensions within an individual is a situation ideally suited to the application of psychophysical methods.

Psychophysical techniques are not confined to the classical problems in sensory and perceptual psychology. Effective application of the techniques has ranged widely from the estimation of the sweetness of various sugars (Moskowitz,1970), to the subjective seriousness of criminal offences, (Stevens, 1966). The reason for this wide applicability is the flexibility of the basic methods used. Corso (1967) points out that a given method is rarely used in exact accordance with its classical prototype, and that the actual method employed is determined very much by the nature of the specific sensory problem under consideration.

Although the question of mouthpiece force is well suited to such methods, there are some reservations which need to be discussed. These centre on the validity of psychophysical data in general. There are three main types of criticism which need to be overcome.

Firstly there is the radical epistemological argument which questions the very nature of the sensory scale. As early as 1879, Exner suggested that mental magnitudes were not divisible in the same way as physical magnitudes. Garner (1958) observed that the justification for scales developed from quantitative scaling procedures derives entirely from their face validity. It must be accepted that observers can actually describe their experience using the kind of arithmetic implied in the particular procedures used. Their type of argument can be used against any phenomenological data. However, such criticisms are more relevant to the use to which the data is put, and this study will avoid placing too much emphasis on the absolute nature of any derived scale.

A second criticism concerns the meaningfulness of group functions. Estes (1956) highlights the problem of drawing the conclusion that a group function necessarily reflects the nature of the individual functions from which it is derived. There is evidence that psychophysical functions are artefacts resulting from the averaging of data, which is the normal procedure in psychophysical data analysis. Pradhan and Hoffman (1963) using the arithmetic mean derived a power function for lifted weights from a group of subjects where only one individual's data yielded a power function on its own. However, several investigations have established that individual subjects can reliably yield power functions (see for example Ekman et al 1968). This study will try to establish individual psychophysical functions as well as the usual averaged group functions.

The third type of criticism is the most difficult to answer and needs to be examined closely. This is the argument that procedural variables contaminate psychophysical data to an unacceptable extent. Poulton and Simmonds (1963) go so far as to state that it is impossible to derive a measure of sensory magnitude which is not determined by the particular experimental conditions obtaining and that there can be no direct measures of sensation of any general meaning. There is no doubt that procedural variables do influence psychophysical functions. It is not the case however that such experimentation is rendered futile. Firstly there is evidence that contextual influences can be minimal with certain direct methods (Fillenbaum, 1963). Secondly, many of the criticisms are again only relevant to the use to which the data is put. Experiments can be reasonably valid if any inference refrains from claiming the discovery of absolute fundamental sensory scales. Thirdly, it is always necessary to consider the nature of the particular sensory problem under consideration. It is difficult to see how the paradigms of signal detection theory could constructively be applied to the judgment of mouthpiece force in trumpet players. This study will therefore yield to the particular sensory problem involved but maximise validity by a representative sampling of the independent variables in the light of the standard criticisms of psychophysical investigation together with the appropriate control of transfer effects.

To begin with, this study uses <u>direct</u> psychophysical methods only. Of the unidimensional scaling methods the quantitative judgment methods are least affected by procedural factors. The indirect or variability methods (for example the method of limits or the method of average error) use a rationale which is likely to result in the confusion of sensory with response criteria. The distortion by attitude factors and decision processes in the derivation of the scales make these methods least likely to provide a simple measure of sensory activity, (Swets, 1964). Even so, certain reservations must be stated in that the direct methods used in this study are to some extent based on the arbitrary acceptance of the validity of ratio judgments, that is, that the subject is able to make direct quantitative judgments in accordance with the instructions. The direct methods used in this study give judgment values directly on the subjective scale continuum. The variability is attributed to experimental error, the reduction of which it is the aim of the current methodology to achieve. We will now examine some of these critical factors of the experimental design.

(2) Choice of Independent Variables

(i) Firstly, there is the problem of the value of the absolute threshold as the natural origin of the sensory scale. In determining the nature of the psychophysical function from the various prothetic continua, Stevens (1959) noted the deviation of the functions obtained at low stimulus values. He identified the problem as that of taking the origin of the physical continuum as the physical zero point, rather than a value related to the absolute limen. He then introduced a constant S_0 into the power equation which is related (and roughly equal to) the threshold value of the subject. Corso (1963) criticises this operation as a post hoc expedient chosen to obtain the classical power function. This study avoids problems associated with the absolute threshold by using ranges of stimuli well above the minimum detectable level. All stimuli are detected by all subjects every time. This occurs by definition as the experiment requires the playing of notes on the trumpet. As seen in chapter 3, a minimum amount of force is necessary to seal the lips. At large enough stimulus values the error introduced by the lowest stimulus values should thus be negligible,(Corso ,1967). The only way a problem could arise would be if much adaptation occurred. Obviously theoretical adjustments would be necessary when the absolute threshold might vary significantly as it would in, for instance, brightness judging experiment (Ekman and Gustaffson, 1968). To avoid adaptation effects, the stimulus presentations were as brief and as well spaced as possible.

(ii) The range of stimulus values is an important consideration. Poulton (1968) gives a figure of 33% for the amount of variance in exponents for the same psychological dimension between different procedures. Smaller ranges tend to give larger exponents and vica versa (Björkman and Strangert, 1960). However, some studies show the effect is not necessarily large (Engen and Levy, 1958), if indeed it is observed at all, (Pradhan and Hoffman, 1963). Often, the studies purporting to demonstrate range effects set up an ambiguous experimental situation, where sensory input does not provide a clear basis for judgment. Garner (1954) for instance maximised the response bias of the subjects by deliberately selecting a narrow range of comparison stimuli. The present experiment uses the largest possible range of mouthpiece force and corresponds to that normally encountered in trumpet performance (0.1 - 7kg).

(iii) The position of the standard in the range of stimuli has some reported effects. Less steep functions have been found where the standard is at the extreme ends of the range of variables, (Beck and Shaw, 1961). The problem cannot be overcome by simply not using a set comparison stimulus, as the first stimulus that is presented would become to all intents and purposes (Poulton, 1968) a standard for the subjects. The best procedure, as suggested by Marks (1974), is to present the first stimulus at different intensities to different subjects although not using any extreme values, and in the case of magnitude estimation, to allow the subject to freely assign a number to represent the subjective magnitude of the first stimulus.

(iv) To control for the distance between the standard (or first stimulus) and the second stimulus, the order of presentation of the second stimulus needs to be systematically varied for the subjects, again however, avoiding initial extreme values.

(v) Poulton (1968) claims there is an influence on the psychophysical functions due to the difference between fractional estimates (giving a finite number set limited by the zero at the far end) and multiple estimates (giving an infinite number set). He reports transfer effects here as being important. However this effect will be small, if evident at all, if it is pointed out to the subjects that they are free to use fractions as estimates, which would likewise give an infinite number set towards the zero.

(vi) The choice of modulus has an effect on the psychophysical function obtained (Poulton and Simmonds, 1963). However such effects are often found in studies using numbers not commonly employed by the subjects. There seems no reason to make the task even more difficult for the subject than it already is. In this study a modulus is only used where it is impossible to use free scaling methods.

(vii) Another general problem to be resolved is that of stimulus spacing. This is a problem which relates to the stimulus range discussed earlier. Beck and Shaw, (1965) note a localized steepening of psychophysical functions when stimuli are bunched together. Obviously as this study is interested in employing as large a range as possible of the subjective mouthpiece force, there is no reason to space the stimuli irregularly. However it is difficult to decide a priori how to space the stimuli as nothing is known about the sensation spacing for this dimension.

(viii) The final general problem concerns the question of naive versus practised subjects. Poulton (1968) outlines the dangers of unduly influencing subjects' choice of numbers in the non iterative type of practice procedures usually employed. Also Stevens and Tulving (1957) show that it is perfectly possible to produce good results with unpractised subjects. Furthermore, the subjects, being trumpet players, will be aware of the range of stimuli involved from their experience on the trumpet. Hence it was not considered necessary to use subjects practised in psychophysical scaling techniques.

(3) The Problem of Passive as Against Active Application of Mouthpiece Force.

Ideally, this study would have involved passive judgment of mouthpiece force. This would simplify the analysis of the data and interpretation. However, it would not be at all easy to do this. Chapter 3 shows the very high levels of mouthpiece force used in the upper register by many subjects. Application of these same forces to a passive subject would be extremely hazardous. The use of such large forces on the lips is part of a complex co-ordinated sequence of behaviour involving formation of the embouchure, breath support and a host of oropharyngeal adjustments. If the subject himself were not in complete control of the mouthpiece force, the lack of integration of the above performance characteristics could result in physical damage to the lips. It is extremely unlikely that any professional performers would take part in such a study.

(4) Aims of the Experiment

(i) The first aim is to try to obtain psychophysical functions for mouthpiece force in trumpet performance and to relate any discovered functions to other psychophysical data from the literature.

(ii) The second aim of the study is to compare different groups of subjects with respect to (a) their psychophysical judgments of mouthpiece force and (b) the accuracy of the subjective judgments. The groups used represent different levels of proficiency on the instrument.

D'Amato (1970) stresses the point that in order to avoid bias by using a single scaling method, it is wise to attempt to use several different methods. . The rationale for each method will be discussed in turn.

As regards the comparison of subjects at different levels of skill, Briggs (1968) studied the psychophysical functions of labial force judgments, also with a transducer. In this case a sustained isotonic contraction (that is, a clenching of the lips together) for a period of seven seconds was measured. He compared trumpet players and non trumpet players using as the dependent variable the deviation of the force demanded from that produced in a bisection production experiment. He found no difference with respect to accuracy of subjective force judgments between the groups. In this study, the method of the proficiency analysis will depend on the nature of any psychophysical function obtained.

A) RATIO PRODUCTION

This study employs the fractionation procedure of Stevens and Volkmann (1940) where the subject indicates the magnitude of a sense ratio (the ratio between two subjective magnitudes on a psychological continuum). Because the subject directly produces the proportionality judgment, there is no "set of comparison stimuli". This reduces the variability of the estimates and diminishes the effects of context. No "doubling" is used to balance out the "halving" on which the scale is based. This is because of an imbalance in the data that might arise. It is possible to use halving over the entire range of commonly used mouthpiece force. This is not the case with doubling, in which the comparison stimuli may only reach half the maximum force used. Particularly in view of the lack of information as to what type of function is to be expected, it would be too difficult to design a balanced study involving both halving and doubling.

Although often criticised as being vulnerable to procedural bias, the method is a strong one from the point of view of the minimal amount of number manipulation necessary.

B) MAGNITUDE PRODUCTION

This is one of the last invented scaling methods. A standard and modulus are necessary in this case to prevent the subject from placing his first adjustment too close to the extreme ends of his available stimulus range.

The ratio and magnitude production experiments require the subjects to depart from normal playing procedure. They are asked to play a note in the normal way and then adjust the mouthpiece force while still playing the same note. In the circumstances of the experiment this is the only way to avoid the problem of artificial reliability described by

Underwood (1966). According to this, if a subject is likely to recognise a specific stimulus as having been presented before he may make particular judgments merely in order to be internally consistent rather than respond purely to the apparent sensory magnitude, thus providing the stability he presumes that the experimenter expects of him. This problem would arise if different notes were used to provide mouthpiece force stimuli. As the basis for the reliability of the data is to be found in taking enough repeated measurements at the specific stimulus levels, it is necessary to alter the mode of playing for these studies.

C) MAGNITUDE ESTIMATION

The problem of artificial reliability is especially relevant in the magnitude estimation study. This is because in this case the only practicable way to deliver specific stimulus mouthpiece force is to get the subject to play specific notes. To prevent spurious reliability in the results it was decided not to use multiple presentations at the same stimulus level. This unfortunately eliminates the possibility of discovering individual psychophysical functions for this method, and increases the number of subjects necessary to provide enough data for a group analysis. However, as the method of magnitude estimation is generally accepted as being the most resistant to extraneous factors, it was decided to include it in the study. "Free scaling" (where there is no designated standard or modulus) was used to minimise the contextual effects.

Method

Apparatus

In this study the transducer was fitted to the subject's own trumpet, or to a standard model in the case of non trumpet players. The chart recorder monitoring mouthpiece force was run continuously throughout the experiment. The judgments were made in a sound proof room. For details of the apparatus, see chapters 2 and 3.

Subjects

(i) Individual Study

In view of the judgments necessary to establish individual psychophysical functions, only six subjects were used; two non-trumpet players, two medium proficiency performers and two professional players with extensive teaching experience in recognised music colleges.

(ii) Group Study

For this part of the study there were three groups of subjects. Group one consisted of twelve non trumpet players. Group two consisted of intermediate trumpet players numbering thirty two in all, while group three contained thirty two professional trumpet players. Twenty subjects from group two and twenty subjects from group three were randomly selected and allocated to the magnitude estimation experiment, and the remaining twelve in each group randomly allocated to the ratio production and the magnitude production experiments, giving six in these two studies for each of groups two and three. For this arrangement see table 5.1. To avoid transfer effects it was decided that each subject should only provide data for one technique, which limited the number of the available subjects for each group. The twelve non trumpet players were randomly allocated to the ratio production and magnitude production experiments.

Procedure

Each psychophysical method involved getting the player to produce the comparison stimulus mouthpiece force. In the case of the trumpet players this was done as follows.

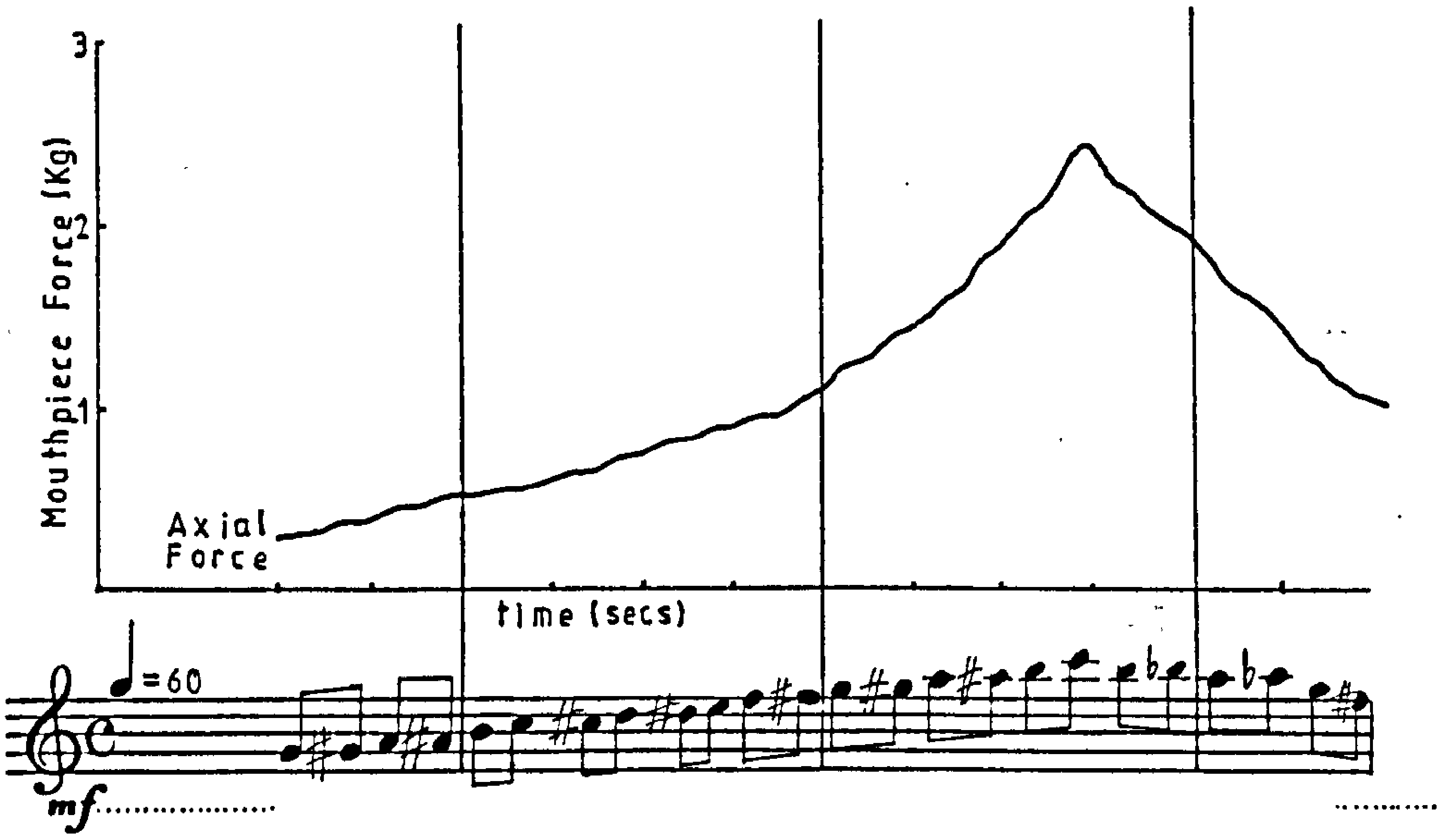
The subject performs a standard chromatic exercise to reveal his particular mouthpiece force profile, (see figure 5.1).

As the players' use of force tends to be very consistent, (see chapter 3), the desired level of mouthpiece force can be simply achieved by instructing the subject to play the note associated with a given level on the chart recorder. For example in figure 5.1 the subject will produce 2kg by playing F$\sharp_5$. If the subject produces a force more than .1kg away from the target level, he is asked to play gradually more quietly or loudly until the desired level is reached (see chapter 3 for the effect of sound intensity on mouthpiece force).

Table 5.1 Numbers of subjects allocated to each psychophysical investigation technique.

	Non trumpet players	Intermediate players	Professional players
Ratio Production	6	6	6
Magnitude Production	6	6	6
Magnitude Estimation	0	20	20
Total	12	32	32

Figure 5.1 Example of a mouthpiece force trace for the chromatic exercise. The axial force only is plotted, after correction for sagittal cross-sensitivity.



The subject is given a few trials before the start of the experiment proper to get used to any initial adjustments. Fortunately the correction procedure was simply and swiftly carried out and rarely interrupted the flow of the experiment.

For the non trumpet players a certain amount of training on the instrument was required. A simple chart of chromatic valve positions was given to enable the subjects to produce a chromatic series of mouthpiece forces as per figure 5.1. Although certain difficulties were initially encountered, the non trumpet players managed eventually to produce the required stimulus levels almost as fast as the skilled performers. Unfortunately however the non trumpet players were not able to do the magnitude estimation experiment owing to their lack of facility on the instrument. We now turn to the aspects of procedure specific to each method used.

Ratio Production

The subject is given a warmup period of two minutes to adjust to the use of the transducer. After performing the standard chromatic exercise, he is given the following instructions (these are given here in full owing to their importance as procedural factors):

> "You will be asked to play a given note mezzoforte. Hold this note for two seconds. I may ask you to play the note a little louder or softer. Then I will ask you to reduce the mouthpiece force until it feels to you to be half as much as the force you were using at the moment I asked you to make the reduction. Don't worry if you lose the note, just keep

the normal playing position and adjust the mouthpiece pressure. Hold the level judged to be one half for at least two seconds.

Remember that we are not concerned with the actual physical difference between the pressure used, but only what feels half as much to you. You will not be required to use any more or les mouthpiece pressure than you are accustomed to applying".

After a little practice at yielding the particular stimulus values, the experiment began.

(i) Individual Study

Each subject performs 15 judgments at each of six stimulus values ranging from 0.5 - 5kg. Some subjects were able to perform additional judgments at higher stimulus values.

(ii) Group Study

Each subject makes 5 judgments at each of six stimulus levels, ranging from 0.5 - 5kg. Some subjects, if they were able, made additional judgments at higher stimulus levels.

For both studies the order of presentation of the stimuli was completely randomized except for the first stimulus, which was of medium intensity, (1 or 2kg). There was at least a 10 second gap between each trial to prevent adaptation from occurring, and there was a one minute break between every 10 estimates, as the task required much concentration.

Magnitude Production

The procedure for eliciting the stimulus level required is identical to that in the ratio production experiment. In the case of magnitude production however, a standard of one kilogram was used, and a modulus of 10 assigned to this. The instructions given to the subjects were similar in the following respects. After achieving the initial standard stimulus by the normal method, the subject was told:

> "Let the number 10 stand for the apparent value of the mouthpiece pressure you are exerting. You will be asked to exert pressures in proportion to other numbers while still playing the same note. For example, if given number 25, you should try and use a force 2½ times as great as that we called 10. If you are given the number 2, exert a force $\frac{1}{5}$ of that called 10".

Each subject performed 5 judgments at the six stimulus levels required, (5, 7.5, 20, 40, 80, and 100). It was not considered safe to attempt to elicit forces above 4kg as this seems to be the maximum value that all players can obtain without strain. The order of the judgments after the standard was completely randomized.

Magnitude Estimation

The subject is required to produce a note corresponding to a mouthpiece force of 1kg. He then freely assigns a number to this. He is then asked to produce different notes which correspond to specific different force values. The method of eliciting stimuli in this technique precludes the use of a group of non trumpet players, as they are unable to achieve the right notes quickly enough for a valid series of judgments. The order of stimuli was completely randomized apart from the first of 1kg. For each subject only two judgments were made at at least 6 different levels of mouthpiece force from 0.5 to 5kg. Some subjects were able to supply data above 5kg depending on their customary overall levels of force.

After eliciting the required note and mouthpiece force according to the above procedure, the subject was instructed:

> "Once you have achieved the desired note, try to give a number appropriate to the mouthpiece pressure you are using. Remember that there are infinite numbers upwards and fractions downwards which can be used. If the amount of one mouthpiece pressure seems twice as much as another then you should respond by assigning a number double the other estimate. Any type of number, whole number or fraction might be used".

Results

(A) Ratio Production

(i) Individual Data

The results for individual subjects are plotted for the halving function, as is usual in fractionation experiments (Harper and Stevens, 1948). Figure 5.2 shows the six graphs for the function relating the stimulus judged half (which is

Figure 5.2 The halving functions for six individual subjects in the ratio production experiment, together with the psychophysical functions derived from these. The horizontal axis gives the actual mouthpiece force (kg) while the vertical axis gives the force which was judged to be half that actual force.

Non trumpet player 1

Non trumpet player 2

Intermediate player 1

Figure 5.2 (continued)

the median of the 15 judgments at each stimulus level), and the physical stimulus for the 6 subjects.

A plot of these medians on log-log co-ordinates gives a near linear relationship in most cases. The straight line in each graph is determined using the method of averages.

In order to construct a sensory scale it is necessary to perform a computation on the halving function data. Firstly a unit must be defined to measure the psychological dimension. Harper and Stevens (1948) used the "veg" to represent the unit of apparent heaviness. This experiment uses one unit of mouthpiece force as the apparent force produced by 1kg of actual mouthpiece force. Figure 5.2 shows the final psychophysical functions which result. The method by which these were derived is described in the appendix (see appendix 5.1). Mathematical procedures were preferred to graphic interpolation for reasons of accuracy. Table 5.2 gives the value of the exponents calculated for the functions of each subject. No constants are given as these were arbitrarily determined by the unit of subjective magnitude chosen. It is the size of exponent which is of primary importance.

(ii) Group Data

Figure 5.3 demonstrates the same procedure applied to the geometric means of the group data, each point being the halving function representing 30 judgments (except at the

Table 5.2 The following are the tabulated exponents of the psychophysical functions from the different methods applied in this experiment.

INDIVIDUAL FUNCTIONS

All the following are from the ratio production experiment.

	subject	exponent
Non trumpet players	1	1.98
	2	1.50
Intermediate players	1	1.38
	2	1.82
Professional players	1	1.52
	2	1.75

Mean exponent 1.66
Standard deviation 0.21

GROUP FUNCTIONS

A) Ratio production

Non trumpet players	exponent 1.71
Intermediate players	exponent 1.54
Professional players	exponent 1.60

Mean exponent 1.6

B) Magnitude production

Non trumpet players	exponent 1.60
Intermediate players	exponent 1.54
Professional players	exponent 1.49

Mean exponent 1.54

C) Magnitude estimation

Non trumpet players	no data
Intermediate players	exponent 1.28
Professional players	exponent 1.25

Mean exponent 1.26

Figure 5.3 The halving functions for the three groups of subjects in the ratio production experiment, together with the psychophysical scales derived from these.

highest stimulus levels which fewer subjects could provide data for). The exponents are given in table 5.2 for the three groups. The geometric mean is preferred for averaging the group data to utilize all the results without being too distorted by extreme estimates, of which there were a few.

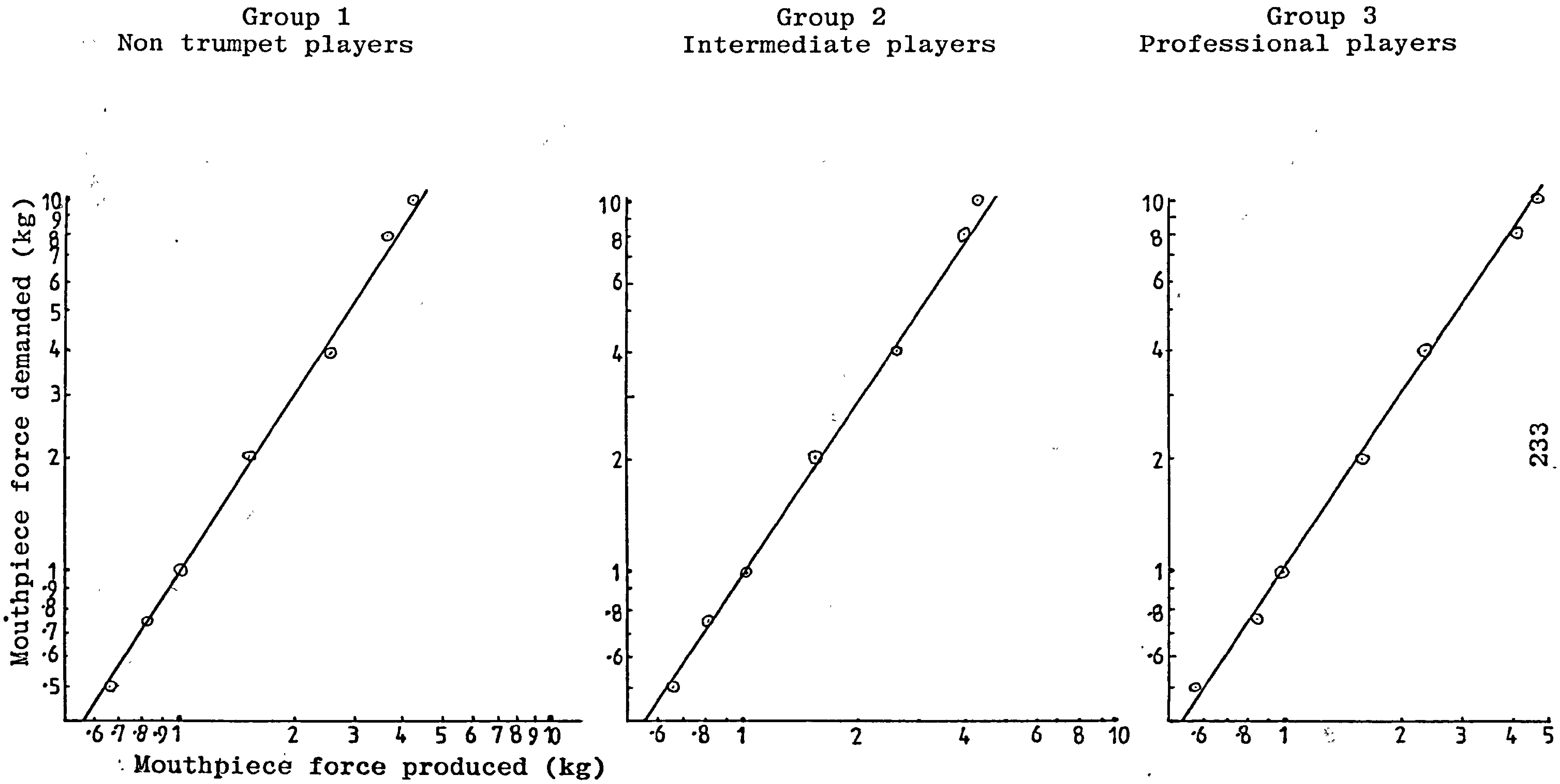
(B) Magnitude Production

As this technique has the observer respond in psychological scale units in the first place, there needs to be no more calculation than the central tendency of the group judgments. The geometric mean was again used. Figure 5.4 shows the psychophysical function plotted by the method of averages for the three groups, with the exponents given in table 5.2.

(C) Magnitude Estimation

This technique again involves the direct use of scale units. However, as free scaling was used it was necessary to bring all the estimates to a common unit. In keeping with the unit of mouthpiece force as 1kg of subjective force, the estimates were all adjusted to give the value of 1 as the judged amount of mouthpiece force resulting from an actual force of 1kg. Figure 5.5 gives the plot for the magnitude estimation function from the geometric means of the group data. As explained earlier, it was only feasible to use the procedure on moderately experienced trumpet players.

Figure 5.4 Psychophysical functions for the three proficiency groups in the magnitude production experiment. The standard was 1kg with the modulus 10. The values on the ordinate have been converted to kilograms.

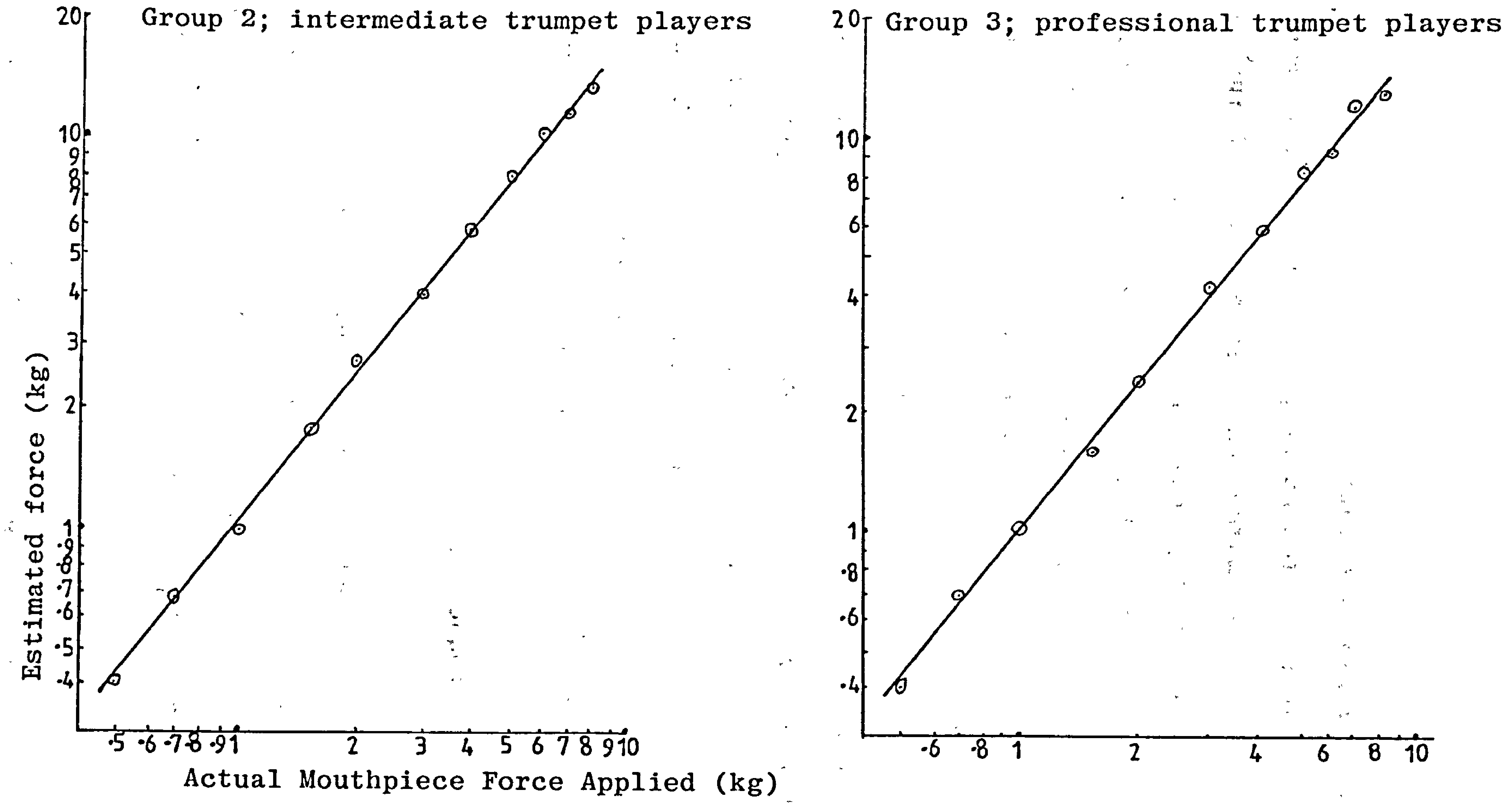


Group 1
Non trumpet players

Group 2
Intermediate players

Group 3
Professional players

Figure 5.5 Psychophysical functions for the two proficiency groups in the magnitude estimation study. The apparent magnitudes given by the ordinate have been adjusted such that 1kg = the apparent force resulting from 1kg actual force.



Comparison of Proficiency Groups

There would be many ways of comparing the judgments in the three groups of subjects, but the two most obvious approaches were as follows.

(i) Comparison of groups by the accuracy of judgment with respect to the objective criterion. To satisfy the conditions of independence of data in the statistical tests used, the groups were compared at each stimulus level.

(a) Ratio Production

For each subject a calculation was made of the average deviation from the objective criterion of the halving judgments made. This gave six independent measures of the error in judgment for each proficiency level at each stimulus level 0.5 to 5kg. At higher stimulus levels there were fewer subjects providing data but the results were still analysable. Using the normal approximation to the Jonckheere test (Leach,1979) at only one of the ten stimulus levels was there any statistical support for the hypothesis that the accuracy of judgment increases with the level of subject proficiency on the instrument. At this 5kg stimulus level the data supports this prediction at the .05 confidence level. However, with tests at ten separate stimulus levels this is hardly indicative of a consistent trend for proficiency level, and is more likely to be merely a type I error.

(b) Magnitude Production

The response variable in this case was the average deviation (the arithmetic mean of 5 deviations at each force level) for each subject of the force produced from that demanded. There are six separate stimulus levels, each giving indices of "error" for each of the three proficiency level groups.

In this case, using the normal approximation in the Jonckheere test, no support is found at any stimulus level for the hypothesis that accuracy in mouthpiece force judgment improves with proficiency level.

(c) Magnitude Estimation

The response variable again is the average deviation of the force exerted from that estimated. With only two groups this time, a Mann Whitney test shows no significant difference between the groups at any of the stimulus levels.

(ii) Another mode of comparison involves taking into account the psychophysical scale which has been developed relating the psychophysical and physical dimensions. This calculation is made by applying the identical statistical tests, as before, only to the variability of the individual's judgments at each stimulus level. The response variables become the standard deviation which is calculated for each subject's estimates at the respective stimulus levels. For the magnitude estimation study, the range was used as there were only two estimates for each subject. At all the stimulus levels tested, there was no support for the hypothesis that proficiency level was related to the variability of the estimates.

Discussion

The results clearly show that a psychophysical function of an exponential nature can be obtained for the stimulus situation which the study entailed. Any deviations observed from the function appear to be random rather than systematic, except for a general tendency for variability to increase at upper extreme stimulus levels.

A good exponential fit was obtained for individual subjects with the ratio production method. As the basis for the reliability of psychophysical data is in taking enough repeated measurements,the individual function experimentation is very time consuming and it was thus not feasible to demonstrate individual functions using the other techniques.

The group functions (see figures 5.3, 5.4 and 5.5) show that measures of group central tendency yield more stable functions. Clearly though they have little value in predicting what any single subject will do in an experiment, as although the medians for the single subjects could be fitted well by straight lines in log - log co-ordinates, the slope varied from one subject to another (see table 5.2).

The data reveal a marked difference between the production experiment and the estimation study. The exponent for the magnitude estimation is rather less (average 1.26) than the ratio and magnitude production (averaging 1.6 and 1.54 respectively). This type of difference has been observed before when investigating the same stimulus situation by different methods. Stevens and Greenbaum (1966) referred to this phenomenon as the "regression effect in psychophysical judgment". They relate the magnitude of what they call "recruitment" to the "difficulty" of the task although the latter is obviously hard to define. As some kind of response restriction is likely to occur, it is often held to be necessary to use both types of technique wherever possible and use a combination of the results.

This experiment has therefore met the two main requirements for a valid determination of a subjective scale as given by Ekman (1961) namely; (i) Scalability of the stimulus by physical procedures, enabling the assignment of a numerical value to the magnitude of a physical attribute. This is ensured with great accuracy by the design and calibration of the transducer, (see chapter 2). (ii) The variables selected for judgment should encompass the entire range of the psychological stimulus values appropriate to the experiment. This study used the whole range of mouthpiece force encountered in normal performance (0.5 to 9kg).

There is some difficulty in interpreting the functions found in this study. Garner (1954) makes the point that we cannot assume the validity of a scale simply because the subject has been required to make numerical judgments. It may be the case that the subject is not capable of faithfully reporting the magnitude of his sensations. The scales produced by psychophysical experiments have often been accused of telling us more about the number behaviour of subjects than of any purported sensory dimension (Teghtsoonian,1971). Warren and Warren (1963) propose a physical correlate theory whereby the subject is quantifying the stimulus rather than the sensation magnitude. Ross and Di Lollo (1968) suggest a multidimensional model for the determination of the scale produced. It seems unavoidable that any sensory scale will be affected by response criteria and various central processes. However, while this means that there may be no case for an absolute sensory function, there is a case for sensory functions developed from standardised conditions which would then be comparable with similarly developed scales, especially if several techniques are used for the same stimulus situation.

Given then that the scales produced in this particular study have some validity, and that the scales developed do represent sensory magnitudes, it is not exactly clear which sensory functions are involved. The subject might judge mouthpiece force from any number of cues. These could include simple force on the lip, kinaesthetic information from the muscular force exerted by the arms, the force on the hands holding the instrument, or any other usable source of information. Table 5.3 shows 12 exponents culled from the

Table 5.3 Some psychophysical functions from force judgments in different stimulus situations.

Psychophysical Function	Stimulus	Scaling Methods	Function	Experimenter
Tactile intensity	Single pulse (mechanical)	Magnitude estimation	cS^{1}	Greenspan (1980)
Force of handgrip	Dynamometer	Magnitude est. Magnitude prod.& Ratio production	$cS^{1.6}$ $cS^{1.9}$ $cS^{1.9}$	Stevens and Mack (1959)
Force applied to the skin	Pushrod on palm	Magnitude estimation	$cS^{1.1}$	Stevens and Mack (1959)
Muscular effort	Lifting weights	Magnitude estimation	$cS^{1.38}$	Bernyer (1962)
Weights		Constant sum Ratio estimation	$cS^{1.4}$	Baker and Dudek (1955)
Weights		Fractionation	$cS^{2.2}$	Harper and Stevens (1948)
Weights		Fractionation	$cS^{1.8}$	Warren and Warren (1956)
Weights		Fractionation & Constant sum	$cS^{1.3}$	Guildford and Dingman (1954)
Weights		Magnitude estimation	$cS^{1.2}$ (free est.) $cS^{1.6}$ (modulus)	Stevens and Galanter (1957)
Muscular effort	Rotational pressure	Fractionation	$cS^{1.7}$	Girden (reported in Stevens and Mack, 1959)
Muscular effort	Lever pushing	Fractionation	$cS^{1.7}$	Bernyer (reported by Stevens and Mack, 1959)

past psychophysical literature,which represent a good cross section of the procedures used for investigating the psychophysical functions of force judgments of one type or another. Ten of the studies investigate the active application of a force, most of them by lifting weights. They all give exponential functions with exponents reliably greater than 1 (i.e. concave upwards).

Only a few studies however report psychophysical functions from the passive force applied to the skin. Stevens and Mack (1959) found an exponent of 1.1 for a push-rod applied to the palm of the hand. Greenspan (1980) reports finding exponential functions in a series of experiments on tactile intensity. He found differing values of the exponents (around 1) depending on the individual tested or the locus of the stimulation. It seems that spatial factors may well affect judgments of tactile intensity. Greenspan relates the size of the exponent discovered to the factor of skin compressibility,which is a spatial variable. Stevens and Marks (1971) in a study on cutaneous stimuli note that differences in the local density of neural innervation may have an influence on judgments of sensation magnitudes, perhaps by altering the effective area of the stimulus.

Given the size of the exponents found in the present study on judgments of mouthpiece force, it may tentatively be suggested that the psychophysical functions more resemble those normally found in judgments of the active application of force rather than the depth of indentation of the skin. A firmer conclusion could be made were there more reports in the literature on local differences in pressure stimulation, but this appears to be a neglected area.

If it were the case that judgments of mouthpiece force are made by kinaesthetic force information rather than lip indentation, it might explain why no differences were found in the psychophysical functions between the different proficiency level groups, the reasons being as follows. It is widely evident that sensory acuity can be increased by practice and experience (witness the developed ear of piano tuners and the remarkable skill of wine tasters). If in the case of mouthpiece force it were the lip force alone which provided the basis for judgment, one would expect to find more acuity in more practised or skilled players, as this mode of application of force is not one which is encountered normally by non trumpet players. However, the lifting of weights and the exerting of pulling forces is a more familiar stimulus situation to all the groups of subjects, and might lead to similar resulting psychophysical functions in all the groups. The question could only be fully answered by devising a study which attempted to apply forces to the lips of a passive subject. As mentioned before, this would be difficult to manage safely, and could probably only use moderate stimulus values.

To recapitulate, this experiment has successfully demonstrated both individual and group functions in the psychophysical judgment of mouthpiece force. This has been achieved largely by three important features of the method: firstly there was close control of the experimental conditions and variables; secondly the stimuli presented were extremely fast and accurately provided, with the responses rendered simply and swiftly (the design of the transducer made this possible); thirdly, enough judgments were made at each stimulus level to give good reliability (although this necessitated long and arduous sessions for some of the subjects).

Although there is some argument about the fundamental nature of the sensory scales discovered, the main use to be made of the data was in comparing the effect of proficiency level on the resulting functions. The failure to demonstrate any difference between the groups with respect to subjective judgment of mouthpiece force is a further demonstration of the extent to which skilled performers and experienced teachers are unaware of some of the most fundamental aspects of their own performance. Together with the evidence from the last chapter on people's judgments of <u>other</u> individuals use of mouthpiece force, there are serious implications for trumpet teaching in general.

CHAPTER SIX

A CONCLUDING NOTE

It is reasonable to presume that in attempting to explain the mechanical principles which underlie motor acts, these acts should thereby benefit (Seymour, 1966). However, the situation is not that simple, and empirical studies into practical skills often raise more questions than they answer. In musical performance investigators using universally accepted scientific methods are still in dispute as to the nature of some of the most basic phenomena, and the positions taken seem sometimes no less polarised than the normative controversies which inspired the objective research in the first place.

It would be gratifying to be able to attribute the ever increasing standards of brass performance to the growing body of scientific research, but probably other factors have contributed materially to this improvement, for example the greater opportunities for children taking up musical instruments, more leisure time, more chances to observe and learn from the great performers on the different instruments.

Despite this, we must put our faith in the efficacy of the scientific method. Not only does it offer the only intellectually honest approach to the problems at hand, it also can have positively beneficial effects on improving performance standards. However, to begin with it must be accepted that there will be few simple answers. Frohrip (1972) suggests that there will be no single optimum pedagogical method which is appropriate for any particular instrument. As the description of skilled performance requires involved models with several interrelated variables, so will any body of recommendations for improvement of these skills be necessarily complex. Amstutz (1970) for example advances a model purporting to explain the production of notes of different pitch which entails the interaction of the factors of tongue arch, instrument pivot, mouthpiece force and teeth aperture in a co-ordinated pattern, whereby there are several different ways of achieving the same result.

As regards mouthpiece force, some fairly simple yet profound observations can be made as a result of this study. Chief amongst these is the finding that most players will have to cope with large forces applied to the embouchure, both in terms of absolute levels (for example in the high register) as well as in the sense of force over time (endurance). There seems little doubt that mouthpiece force does have detrimental effects on performance with many players. This is on the a priori basis of general medical knowledge (see chapter one) as well as the specific problems which have been encountered amongst experienced players (see for example the dental studies of Porter (1967) and Herman (1981)). This is a dilemma which must be resolved. Future research might look for instance at metabolic factors. It has been suggested (Briggs, 1968) that these factors, rather than strength per se may be significant in the offering of resistance to the external forces of the trumpet mouthpiece. Alternatively, a player might be able to resist such forces by muscular development. It would be most interesting to tie up electromyographic work on the embouchure musculature with the use of mouthpiece force by individual players to see if there is any definite relationship. Another avenue which might be explored is the question of non-axial forces. Although this study did not explore lateral force application statistically, there was evidence to suggest that the high proficiency performers used relatively less sagittal force in the extreme upper register than the medium proficiency group. Speculatively, the even distribution of the force on the lips might be another way of lessening the deleterious effects of the large forces at these pitches.

Although much work remains to be done, it is hoped that this study has shed some light on some of the processes which contribute to skilled trumpet performance.

Teachers of the trumpet are at present not in a position to be able to evaluate physiological and metabolic differences amongst their pupils. An ideal teaching method would take such factors into account. Nevertheless some concrete teaching guidelines emerge from this research.

To begin with, attempts to teach pupils to use general mouthpiece force levels sufficient only to form an airtight seal (the "no pressure" or "minimum pressure" methods) should be abandoned. Pupils should be taught to utilize mouthpiece force as an integral aid to the achievement to pitch and intensity.

However, there is no <u>normal</u> level of mouthpiece force for any given note, and pupils should be allowed to use their own individual range of forces. Emphasis instead should be placed on the attaining of consistency and regularity in the forces applied for given notes across different contexts (as far as this is possible depending on the complexity of the musical material). It is very important for players to resist the tendency to respond to difficult musical tasks with an inordinate increase in mouthpiece force application. This is particularly essential in extreme upper register performance, where it is noticeable that the best performers, although using large forces, still have spare capacity. A teacher should only interfere with a pupil's overall force levels where he has very good grounds for believing that a particular individual's use of force is inappropriate for particular reasons. If a teacher is worried about the effects of large mouthpiece forces, he would do better to find a way for his pupil to learn to withstand and apply such forces rather than artificially try to reduce them.

Equally crucial is the question of the problems and dangers attendant on the teachers' subjective assessments of the mouthpiece force usage of their pupils. This study is evidence for the proposition that mouthpiece force is a non judiciable aspect of trumpet performance. The stark fact seems to be that it is not possible to judge the absolute amount of mouthpiece force that a player is using from his visual appearance, and that professional players are no more accurate than non trumpet players in this respect. The situation is even worse than this however, because it seems that trumpet players use the erroneous criterion of apparent effort as a basis for mouthpiece force judgments, leading to systematic and consistent misjudgment. It is vital for teachers to understand the nature and extent of these types of errors, in order to properly implement the recommendations suggested by this study.

Finally, teachers should be aware of the dangers of training pupils in accordance with introspective interpretations of their own mouthpiece force usage. This study shows that errors of self judgment are as extensive in professional players as they are in intermediate players.

While it is appreciated that certain potentially important areas have not been covered in this study of mouthpiece force application in trumpet performance, it is hoped that some light has been shed on several of the principal issues.

REFERENCES

Akeju, S.A. (1972). The reliability of General Certificate of Education examination English composition papers in West Africa.. Journal of Educational Measurement, 9, 175-180.

Amstutz, A.K. (1970). A videofluorographic study of the teeth aperture, instrument pivot and tongue arch and their influence on trumpet performance. D.Mus.Ed. Thesis, University of Oklahoma.

Anfinson, R.E. (1965). A cinefluorographic investigation of supralaryngeal adjustments in selected clarinet playing techniques. Ph.D. Thesis, State University of Iowa.

Arban, J.J.B. (1936). Arban's complete conservatory method for trumpet (cornet). New York: Carl Fischer, Inc.

Aurand, W.O. (n.d.). Survey on brass instrument playing. Unpublished manuscript, University of Michigan, Ann Abor. 11pp.

Bach, V. (1925). The art of trumpet playing. New York: Vincent Bach Corporation.

Badia, P. and Runyon, R.P. (1982). Fundamentals of behavioural research. Addison-Wesley Publishing Corporation.

Bahrick, H.P. (1957). An analysis of the stimulus variables influencing the proprioceptive control of movements. Psychological Review, 64, 324-8.

Baker, K.E.and Dudek, J.E. (1955). Weight scales from ratio judgments and comparisons of existent weight scales. Journal of Experimental Psychology, 50, 293-308.

Barbenel, J., Kenny,P. and Davies, J.B. (1985). Mouthpiece forces produced while playing the trumpet. Medical and Biological Engineering and Computing, (in press).

Barrett, H.C. and Barker, J.R. (1973). Cognitive pattern perception and musical performance. Perceptual and Motor Skills, 36, 1187-93.

Bartlett, M.S. (1937). Some examples of statistical methods of research in agriculture and applied biology. Journal of the Royal Statistical Society Supplement, 4, 137-70.

Basmajian, J.V. (1967). Muscles alive; their functions revealed by electromyography,(second edition). Baltimore: Williams and Wilkins.

Basmajian, J.V. and Stecko, G. (1962). A new bipolar indwelling electrode for electromyography. Journal of Applied Physiology, 17, p849.

Beck, J. and Shaw, W.A. (1961). The scaling of pitch by the method of magnitude estimation. American Journal of Psychology, 74, 242-251.

Beck, J. and Shaw, W.A. (1965). Magnitude of standard, numerical value of stimulus and stimulus spacing in the estimation of loudness. Perceptual and Motor Skills, 21, 151-156.

Bellamah, J.L. (1960). Brass facts. A survey of teaching and playing methods of leading brass authorities,(third edition). San Antonio: Southern Music Company.

Bernyer, G. (1962). Ētude sur la validitē d'une echelle de sensation d'effort musculaire. Annee Psychologique, 62, 1-15.

Björkman , M. and Strangert, B. (1960). The relation between ratio estimates and stimulus dispersion. Report from the Psychological Laboratory, University of Stockholm, 81.

Blanton, P.L., Briggs, N.L. and Perkins, R.C. (1970). Electromyographic analysis of the buccinator muscle. Journal of Dental Research, 49(2), 389-394.

Box, G.E.P. (1953). Non-normality and tests on variance. Biometrika, 40, 318-35.

Bracht, G.H. and Glass, G.V. (1968). The external validity of experiments. American Educational Research Journal,5,437-447.

Briggs, G.V. (1968). Electrophysiological examination of labial function in college-age trumpet performers. D.Mus.Ed. Thesis, University of Oklahoma.

Brown, G.L. and Burns, B.B. (1949). Fatigue and neuromuscular block in mammallian skeletal muscle. Proceedings of the Royal Society, B, 136, 182-195.

Campbell, D.T. (1968). Quasi-experimental design. In Gills, D.L. (Ed.), International Encyclopedia of the Social Sciences, 5. New York: Macmillan and Free Press.

Chen, W.W. (1976). The lasting effects of electromyography biofeedback relaxation training. Ph.D. Thesis, University of Maryland.

Christenson, L.B. (1977). Experimental methodology. Boston, Mass.: Allyn and Bacon, Inc.

Clarke, H.L. (1934). Characteristic studies for the cornet. New York: Carl Fischer, Inc.

Cochran, W.G. and Cox, G.M. (1957). Experimental designs. New York: Wiley.

Colin, C. (1948). Vital brass notes. New York: Charles Colin.

Cook, N.H. and Rabinowicz, E. (1963). Physical analysis and measurement. Reading, Mass.: Addison-Wesley Publishing Inc.

Corso, J.F. (1963). A theoretico-historical review of the threshold concept. Psychological Bulletin, 60, 356-370

Corso, J.F. (1967). The experimental psychology of sensory behaviour. Holt, Rinehart and Winston Inc.

Cox, D.R. (1958). Planning of experiments. New York: Wiley.

D'Amato, M.R. (1970). Experimental psychology. McGraw Hill Book Company.

De Sousa, O.M. and Vitti, M. (1965). Estudio electromiografico do m. buccinator. O Hospital, 68(3), 105-117.

Deye, H.W. (1947). Use of the tongue in brass instruments. The Instrumentalist, 1(4), 16-17.

De Young. D.D. (1975). A videofluourographic analysis of the pharyngeal opening during performance of selected exercises for trombone. Ph.D. Thesis, University of Minnesota.

Eby, W.M. (1933). Eby's scientific method for cornet and trumpet. New York: Walter Jacobs, Inc.

Ekman, G.A. (1961). A simple method for fitting psychophysical power functions. Journal of Psychology, 51, 343-350.

Ekman, G.A. and Gustaffson, V. (1968). Threshold values and the psychophysical function in brightness vision. Vision Research, 8, 747-758.

Ekman, G.A., Hosman, B., Lindman, R., Ljungberg, L. and Akesson, C.A. (1968). Interindividual differences in scaling performance. Perceptual and Motor Skills, 26, 815-823.

Engen, T. and Levy, N. (1958). The influence of context on constant-sum loudness judgments. American Journal of Psychology, 71, 731-736.

Estes, W.K. (1956). The problem of inference from curves based on group data. Psychological Bulletin, 53, 134-140.

Exner, S. (1879). Hermann's "Handbuch der Psychologie", II, (ii), 242.

Farkas, P. (1962). The art of brass playing. Bloomington, Indiana: Wind Music Inc.

Faulkner, M.E. and Horvath, S. (1967a). The function of the lung in playing a brass instrument: blood gas measurement. The Brass World,3(1), 205-217.

Faulkner, M.E. and Horvath, S. (1967b). Measurement of heart volume and the results of an electrocardiogram during blowing of a trumpet mouthpiece. The Brass World, 3(2), 245-252.

Ferguson, G.A. (1971). Statistical analysis in psychology and education. McGraw Hill Book Company.

Fillenbaum, S. (1963). Contextual effects in judgment as a function of restrictions in response language. American Journal of Psychology, 76 103-109.

Fiske, H.E. (1975). Judge group differences in the rating of secondary school trumpet performance. Journal of Research in Music Education, 23, 186-196.

Fiske, H.E. (1977). Relationship of selected factors in trumpet performance adjudication reliability. Journal of Research in Music Education, 25, 252-263.

Fitts, P.M. and Posner, M.I. (1973). Human performance. London: Prentice Hall International.

Frohrip, K.R. (1972). A videofluorographic analysis of certain facial factors involved in performance of selected exercises for trombone. Ph.D. Thesis, University of Minnesota.

Gagné, R.M. (1962). Military training and principles of learning. American Psychologist, 17, 83-91.

Garner, W.R. (1954). Context effects and the validity of loudness scales, Journal of Experimental Psychology, 48, 218-224.

Garner, W.R. (1958). Half loudness judgments without prior stimulus context. Journal of Experimental Psychology, 55, 482-485.

Garrett, H.E. (1966). Statistics in psychology and education. David McKay Company, Inc.

Gibson, D.R. (1972). A photographic study of twelve professional trumpet embouchures while playing from the low to the extreme upper register. Ph.D. Thesis, University of Minnesota.

Greenspan, J.D. (1980). The effects of skin compressibility upon psychophysical functions of tactile intensity: A comparison of force and depth of skin indentation as stimulus dimensions. Ph.D. Thesis, Florida State University.

Guildford and Dingman, (1954). A validation study of ratio-judgment methods. American Journal of Psychology, 67, 395-410.

Hadden, J.K. (1969). Use of ad hoc definitions. In Borgatta (Ed.), Sociological methodology. Jossey-Bass Inc.

Hall, J.C. (1954). A radiographic, spectrographic and photographic study of the non labial physical changes which occur in the transition from middle to low and middle to high registers during trumpet performance. Ph.D. Thesis, Indiana University.

Harper, R.S. and Stevens, S.S. (1948). A psychological scale of weight and a formula for its derivation. American Journal of Psychology, 61, 343-351.

Haynie, J.J. (1967). A videofluorographic presentation of the physiological phenomena influencing trumpet performance. Booklet. Denton:School of Music, Texas State University (North).

Hays, W.L. (1963). Statistics. Holt, Rinehart and Winston.

Heffner, R.M.S. (1950). General phonetics. Madison, Wisconsin: The University of Wisconsin Press.

Henderson, H.W. (1942). An experimental study of trumpet embouchure. Journal of the Acoustical Society of America, 13, 58-64.

Henderson, W.A. (1979). E.M.G. potentials of the sternocleidomastoid muscle during trumpet performance. International Trumpet Guild Journal, 4, 30-35.

Herman, E. (1981). Influence of musical instruments on tooth position. American Journal of Orthodontics, 80, 145-155.

Hiigel, L.E. (1967). The relationship of syllables to pitch and tonguing in brass instrument playing. Ph.D. Thesis, University of California.

Hodges, J.L. and Lehmann, E.L. (1970). Basic concepts of probability and statistics,(second edition). San Fransisco: Holden-Day.

Holbrook, R.T. and Carmody, F.J. (1941). X-Ray studies of speech articulations. University of California Publications Modern Philology, 20, 187-237.

Isley, C.L. (1971). Facial muscles in brass embouchure. Unpublished manuscript, Appalachian State University.

Isley, C.L. and Basmajian, J.V. (1973). Electromyography of the human cheeks and lips. Anatomical Record, 176(2).

Joncheere, A.P. (1954). A test of significance for the relation between m rankings and k ranked categories. British Journal of Statistical Psychology, 7, 93-100.

Kaplan, L. (1961). The relationship between certain personality characteristics and achievement in instrumental music. Ph.D. Thesis, New York University.

Kelsey, C.A., Mirifie, F.D. and Hixon, J.T.(1969).Application of Ultrasound in speech research. Journal of Speech Research, 12, 564-75.

Kemp, A. (1981a). The personality structure of the musician. Psychology of Music, 9(1), 3-14.

Kemp, A. (1981b). Personality differences between the players of string, woodwind, brass and keyboard instruments, and singers. Council for Research in Music Education Bulletin No. 66-67, 33-39.

Kendall, M.G. (1970). Rank correlation methods,(fourth edition). London: Griffin.

Kerlinger, F.N. (1973). Foundations of behavioural research. Holt, Rinehart and Winston.

Kirkpatrick, G.S., Day, E.E.and Lehmann, J.F. (1969). Investigation of the performance of an ischial load bearing leg brace. *Experimental Mechanics*, *9*, 31-35.

Lacey, R.B. (1969). The teaching of the trumpet using special material. Ed. D.,Boston University.

Lamp, C.J. and Keys, N. (1935). Can aptitude for specific musical instruments be predicted? *American Journal of Educational Psychology*, *26*, 587-96.

Leach, C. (1979). *Introduction to statistics. A non-parametric approach for the social sciences*. J. Wiley and Sons Ltd.

Leanderson, R., Öhman, S. and Persson, A. (1967). Electromyographic studies of facial muscle co-ordination during speech. *Acta Otolaryng. Suppl*., *224*, 307-310.

Legge, D and Barber, P.J. (1967). *Information and skill*. Methuen.

Leno, H.L. (1970). Lip vibration characteristics of selected trombone performers. Ph.D. Thesis, University of Arizona.

Levee, J.R., Cohen, M.J. and Rickles, W.H. (1976). Electromyographic biofeedback for the relief of tension in the facial and throat muscles of a woodwind musician. *Biofeedback and self regulation*, *1(1)*, 113-120.

Lewis, D.G. (1968). *Experimental design in education*. University of London Press.

Lindquist, E. (1953). *Design and analysis of experiments*. Boston: Houghton Mifflin.

Lippold, O.C.J. (1952). Relation between integrated action potentials in a human muscle and its isometric tension. *Journal of Physiology*, *117*, 492-9.

Malek, V. (1954). A study of embouchure and trumpet-cornet mouthpiece measurements. Ph.D. Thesis, Northwestern University.

Marks, L.E. (1974). *Sensory processes. The new psychophysics*. New York and London: Academic Press.

Martin, D.W. (1941). A physical investigation of the performance of brass musical wind instruments. Ph.D. Thesis, University of Illinois.

McCutcheon, M.J., Fletcher, S.G., Akira, H. and Hasegawa, A. (1977). A videoscanning system for the measurement of lip and jaw motion. *Journal of the Acoustical Society of America* , *61*, 1051-5.

Meidt, J.A. (1967). A cinefluorographic investigation of oral adjustments for various aspects of brass instrument performance. Ph.D. Thesis. University of Iowa.

Merton, P.A. (1954). Voluntary strength and fatigue. Journal of Physiology, 123,553-564.

Minium, E.W. (1970). Statistical reasoning in psychology and education. John Wiley and Sons, Inc.

Moskowitz, H.R. (1970). Ratio scales of sugar sweetness. Perception and Psychophysics, 7,315-320.

Mursell, J.L. (1937). The psychology of music. New York: Norton.

Myers, L.S. and Grossen N.E. Behavioural research. Theory, procedure and design. San Fransisco: W.H. Freeman and Co.

Neff, C.H. (1974). A study of trumpet pedagogy and repertoire in the U.S. and Scandinavia. D.M.A., University of Oregon.

Nichols, R.L., Hanson, F., Daehling, W.A. and Hoffman, H. (1971). The improvement of brass instrument teaching through the use of a profile of the physical aspects involved. Washington D.C.:U.S. Government Printing Office.

Orne, M.T. (1962). On the social psychology of the psychological experiment, with particular reference to demand characteristics and their influence. American Psychologist, 17,776-782.

Paillard, J. (1960). The patterning of skilled movements. In Handbook of Physiology, Vol.III. Washington D.C.: American Physiological Society. 1679-1708.

Perkell, J.S. (1969). Physiology of speech production: Results and implications of a quantitative cineradiographic study. Research Monographs, No. 53. Cambridge, Mass.: M.I.T. Press.

Pflieger, F.C.(1971). An experimental study to determine the relationship of various independent psychomotor abilities to performance success on selected musical instruments. Ph.D. Thesis, West Virginia University.

Porter, M.M. (1967). Dental problems in wind instrument playing. British Dental Journal, 123, 393-396.

Poulton, E.C. (1968). The new psychophysics. Six models for magnitude estimation. Psychological Bulletin, 69, 1-19.

Poulton, E.C. and Simmonds, D.U.C. (1963). Value of standard and very first variable in judgments of the reflectance of grays with various ranges of available numbers. Journal of Experimental Psychology, 65 297-304.

Pradhan, P.L. and Hoffman, P.J. (1963). Effect of spacing and range of stimuli on magnitude estimation. Journal of Experimantal Psychology, 66, 533-541.

Ramsey, G.H.S., Watson, J.S., Tristan, T.A., Weinberg, S. and Cornwell, W.S.,(Eds). (1960). Cinefluorography. Springfield, Illinois: Charles C. Thomas.

Reid, C. (1928). The mechanism of voluntary muscular fatigue. Quarterly Journal of Experimental Physiology, 19,17-42.

Richtermeyer, L.C. (1966). A definite analysis of brass embouchure abnormalities including recommended remedial techniques. Ph.D. Thesis, Michigan State University.

Roscoe, J.T. (1975). Fundamental research statistics for the behavioural sciences, (second edition). Holt, Rinehart and Winston, Inc.

Rosenberg, M.J. (1969). In Rosenthal, R. and Rosnov, R.L. (Eds.), Artifact in behavioural research. New York: Academic Press.

Ross, J. and Di Lollo, V. (1968). A vector model for psychophysical judgment. Journal of Experimental Psychology Monograph Supplement, 77 (3,part2).

Russell, G.O. (1928). The vowel; its physiological mechanism as shown by X-ray. Columbia, Ohio: Ohio State University Press.

St. Jacome,C. (1894). Grand method for trumpet of cornet. New York: Carl Fischer, Inc.

Scheffe, H.A. (1953). A method for judging all possible contrasts in the analysis of variance. Biometrika, 40, 87-104.

Schlaefer, C.M. (1975). A comparative study of an experimental method for increasing range and endurance on brass instruments (cornet-trumpet). D.M.A., The Catholic University of America.

Seashore, C.E. (1906). The tonoscope and its use in the training of the voice. The Musician,11,331-332.

Seashore, C.E. (1919,1939). Measures of musical talents, (original version 1919, revised edition 1939, latest edition 1960). New York: The Psychological Corporation.

Seymour, W.D. (1966). Industrial skills. London:Pitman.

Shannon, C.E. and Weaver, W. (1949). The mathematical theory of communication. University of Illinois Press.

Siegel, S. (1956). Nonparametric statistics for the behavioural sciences. McGraw-Hill Book Company.

Sime, W.E. (1975). Psychophysiology of muscle tension/ relaxation: A comparative study of E.M.G. biofeedback and autosensory feedback relaxation training. Ph.D. Thesis, University of Pittsburgh.

Sloboda, J. (1985). The musical mind. Oxford University Press.

Stammers,R. and Patrick, J. (1975). The psychology of training. Methuen.

Stetson, R.H., Hudgins, C.V. and Moses, E.R. (1940). Palatograms change with rates of articulation. Archives Néerlandaises de phonetique experimentale, 16, 52-61.

Stevens, J.C. and Marks, L.E. (1971). Spatial summation and the dynamics of warmth sensation. Perception and Psychophysics, 9, 291-298.

Stevens, J.C. and Tulving, E. (1957). Estimation of loudness by a group of untrained observers. American Journal of Psychology, 70, 600-605.

Stevens, R. (1971). Embouchure self analysis and the Stevens-Costello triple C embouchure technique. New York: The Stevens-Costello Embouchure Clinic.

Stevens, S.S. (1959). Tactile vibration: Dynamics of sensory intensity. Journal of Experimental Psychology, 57, 210-218.

Stevens, S.S. (1966). A metric for the social consensus. Science,151,530-541.

Stevens, S.S. and Galanter, E.H. (1957). Ratio scales and category scales for a dozen perceptual continua. Journal of Experimental Psychology, 54, 377-411.

Stevens, S.S. and Greenbaum, H.B. (1966). Regression effect in psychophysical judgment. Perception and psychophysics, 1, 439-441.

Stevens, S.S. and Mack,J(1959). Scales of apparent force. Journal of Experimental Psychology, 58,405-413.

Stevens, S.S. and Volkmann, J. (1940). Relationship of pitch to frequency: A revised scale. American Journal of Psychology,53, 329-353.

Sweeney, L. (1948). Double and triple tonguing on brass instruments. The Instrumentalist, vol II (5), 20-21.

Swets, J.A. (1964). Signal detection and recognition by human observers. Contemporary readings. New York: Wiley.

Tanner, P. (1976). About trombone teaching: A national consensus, parts I and II. In Brass Anthology.Evanston, Illinois: The Instrumentalist Company.

Teghtsoonian, M. and Teghtsoonian, R. (1971). How repeatable are Stevens' power law exponents for individual subjects? Perception and Psychophysics, 10, 147-149.

Testa, R.F. (1972). The effect of jaw thrust instruction on four selected aspects of trumpet performance and overjet of young trumpet players. Ph.D. University of Miami.

Trosper, O.W. (1962). The principles and practice of producing a vibrato in brass instrument performance. Ph.D. University of Columbia.

Underwood, B.J. (1966).Experimental psychology. Meredith Publishing Company.

Valin, E.(1961). La valeur des examens: Etude docimologique realisee au Liban. Etude et Documents d'Education, 40. Paris:.U.N.E.S.C.O.

Wagner, C., Piontek, E. and Teckhaus, L. (1973). Piano learning and programmed instruction. Journal of Research in Music Education, 21,106-122.

Warren, R.M. and Warren, R.P. (1956). Effect of the relative volume of standard and comparison object on half heaviness judgments. American Journal of Psychology, 69, 640-643.

Warren, R.M. and Warren, R.P. (1963). A critique of S.S. Stevens' "New-Psychophysics". Perceptual and Motor Skills, 16, 797-810.

Weast, R.D. (1963). A stroboscopic analysis of lip function. The Instrumentalist.

Weast, R.D. (1961). Brass performance: An analytical text of the physical processes, problems and techniques of brass. New York: McGuiness and Marx.

Weast, R.D. and Hake, A. (1965). A definitive analysis of mouthpiece pressure, The Brass World, 1(2), 38-43.

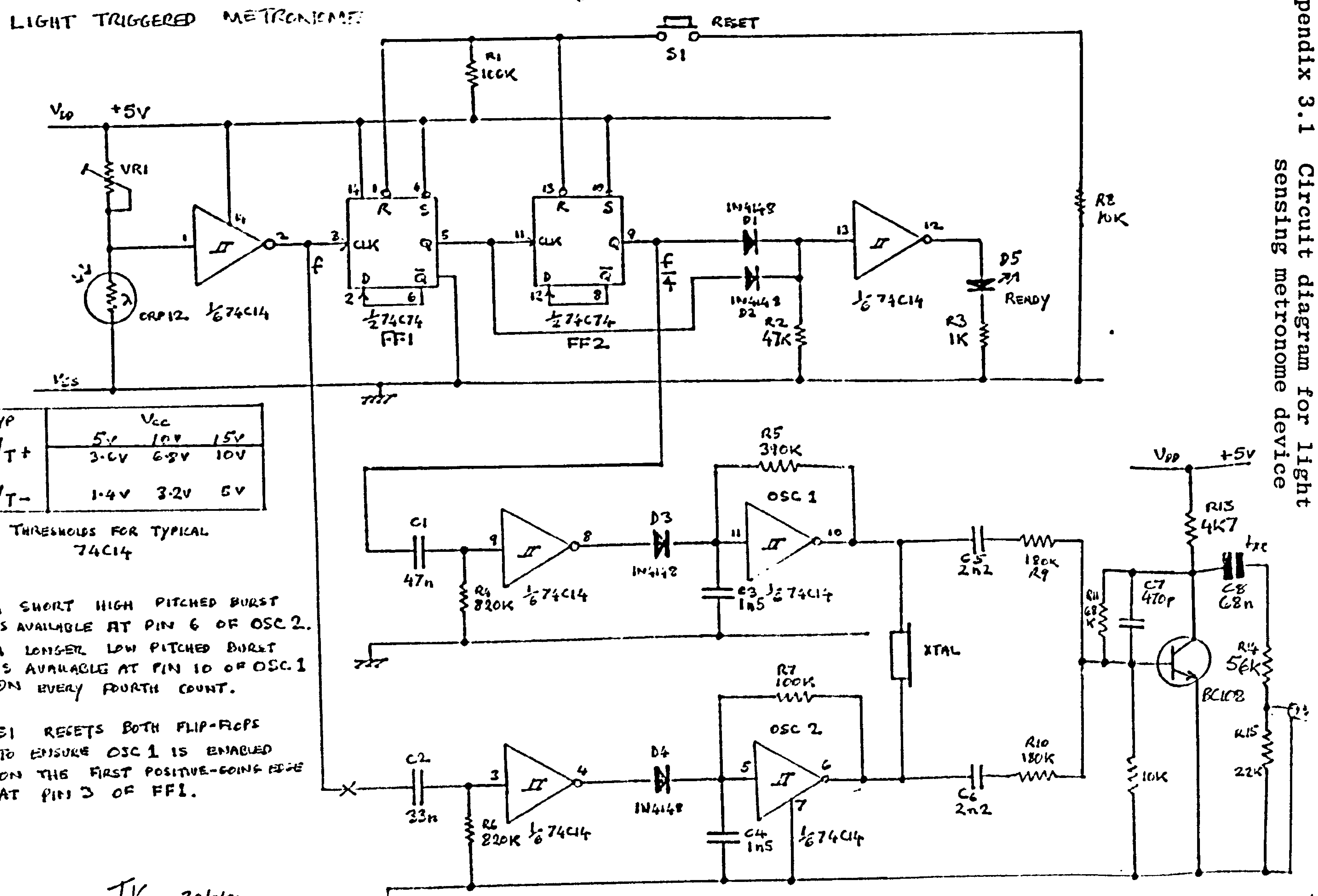
Welford, A.T. (1958). Ageing and human skill. Oxford University Press.

White, E.R. (1972). Electromyographic potentials of selected facial muscles and labial mouthpiece pressure measurements in the embouchure of trumpet players. Ph.D. Thesis, Columbia University.

Whitworth, J.R. (1961). A cinefluorographic investigation of the supralaryngeal adjustments in the male voice accompanying the singing of low and high tones. Ph.D., Iowa State University.

Winer, B.J. (1962). Statistical principles in experimental design. McGraw-Hill Book Company.

Appendix 3.1 Circuit diagram for light sensing metronome device



T/P	Vcc		
	5V	10V	15V
V_{T+}	3.6V	6.8V	10V
V_{T-}	1.4V	3.2V	5V

THRESHOLDS FOR TYPICAL 74C14

A SHORT HIGH PITCHED BURST IS AVAILABLE AT PIN 6 OF OSC 2.

A LONGER LOW PITCHED BURST IS AVAILABLE AT PIN 10 OF OSC 1 ON EVERY FOURTH COUNT.

S1 RESETS BOTH FLIP-FLOPS TO ENSURE OSC 1 IS ENABLED ON THE FIRST POSITIVE-GOING EDGE AT PIN 3 OF FF1.

J.Kerr 30/7/80

MODIFICATION JK 30/10/80 C2 NOW LINKED TO FF1 Q OUT PIN 5

Appendix 4.1

The following gives the format for the instructions which were given to the subjects in the judgmental study (chapter four).

> " You will be given four photographs of a trumpet player (trumpet players). Spread them out on the table in front of you and look at them closely. Then, put them in order, from the least to the most effort (mouthpiece force) which you feel that the player is using in each photograph. Once you have done this, I will give you another set of photographs, and you will then be required to perform the same operation with this, and so on until we have completed several sets of judgments.

The above gives the instructions for the effort judgment of the photograph series A-N. The alternative instructions (in brackets) give the relevant instructions for the mouthpiece force judgments and for the between player series I-IV.

Appendix 5.1

The following calculations show the method for determining the halving functions shown in figures 5.2 and 5.3, together with the method for deriving the associated psychophysical functions. For further details of this technique see D'Amato (1970)

The example which will be worked through is the data from a single subject (see figure 5.2, fifth graph).

Stimulus S_t(kg)	Stimulus judged half $S_{\frac{1}{2}J}$(kg)	Log S_t	Log $S_{\frac{1}{2}J}$
8	5.3	.903	.724
5	3.4	.699	.483
4	2.5	.602	.401
3	1.8	.477	.274
2	1.35	.301	.130
1	0.66	.000	$\bar{1}.819$
0.5	0.32	$\bar{1}.699$	$\bar{1}.505$

The straight line is fitted to the points on figure 5.2 by the <u>method of least squares</u>.

$$\log S_{\frac{1}{2}J} = c \log S_t + k$$

$$\begin{array}{ll}
.724 = c\,(.903) + k & .274 = c\,(.477) + k \\
.483 = c\,(.699) + k & .130 = c\,(.301) + k \\
.401 = c\,(.602) + k & \bar{1}.819 = c\,(.000) + k \\
.247 = c\,(.477) + k & \bar{1}.505 = c\,(\bar{1}.699) + k
\end{array}$$

$$\boxed{\frac{1.882}{4} = c\,\frac{(2.68)}{4} + k} - \boxed{\frac{-.272}{4} = c\,\frac{(.477)}{4} + k} \rightarrow \frac{2.154}{4} = c\,\frac{2.204}{4} \qquad c = .9773$$

$$\boxed{\frac{1.882}{4} = .977\cdot\frac{(2.68)}{4} + k} + \boxed{\frac{-.272}{4} = .977\frac{(.477)}{4} + k} \rightarrow 1.6 = 3.086 + 8k$$

$$\rightarrow k = -.1845$$

or $\bar{1}.8155$

Appendix 5.1 (continued)

The determination of the pschphysical function:

Log $S_{\frac{1}{2}J}$ = .9773 log S_t - .1845

1mf. = the unit of the psychological dimension of mouthpiece force. 1mf = 1kg.

mf	S_t(kg)	Computation
0.5	.653	log $S_{\frac{1}{2}J}$=.9773x0(log 1kg)-.1845 $S_{\frac{1}{2}J}$=.6538
1.0	1.0	
2.0	1.54	log $S_t = \frac{0.00(\text{i.e log 1kg})+.1845}{.9773}$ =.188 S_t=1.54
4.0	2.41	log $S_t = \frac{.188+.1845}{.9773}$ = .381 S_t=2.40
8.0	3.79	sim
16	6.05	sim
32	9.74	sim

To fit the final psychophysical function, the <u>method of least squares</u> is again used.

This yields the values of 1.52 and .108 for the constants c and k, which are effectively the exponent and the constant of the power function.

This finally yields the power funcion

$$mf = .108\ S_t^{1.52}$$ for the subject in figure 5.2