Mild social stress and human performance: the role of competition, evaluation and the presence of others

Weaver, Susan M.

MILD SOCIAL STRESS AND HUMAN PERFORMANCE: THE ROLE OF COMPETITION, EVALUATION AND THE PRESENCE OF OTHERS

SUSAN M. WEAVER

Thesis submitted for the degree of

Doctor of Philosophy, University of Durham

March 1978

ABSTRACT

This work reports a series of experiments designed to investigate the effects of non-*verbal* ('mere presence') social situations on the performance of routine, sensitive tasks. The particular conditions examined are coaction, competition and audience presence. Other variables of interest are individual differences (personality, sex and ability) and type of task (motor, cognitive and perceptual).

The experiments are primarily based on theories of social facilitation, particularly Zajonc's interpretation of mere presence effects and their relationship to arousal level. However, these social situations are also considered in a broader context and effort is made to theoretically integrate this area of research with the larger body of literature concerning stress and human performance. Bearing this latter concern in mind, a central and somewhat unique feature of this work is the use of 'sensitive' performance tasks and the analysis of data in terms of task strategies and shifts in attention, rather than overall measures which sometimes mask effects.

It is concluded that social comparison processes and inter-subject pacing of performance provide a better explanation of the present data than do arousal theories. The generality of these effects across different tasks and subject populations is considered, as is the role of individual differences. A model is offered which outlines the typical sequence of performance related behaviours in mere presence situations and suggestions are made, in terms of this model, for future research.

PREFACE

This work is broadly concerned with the effects of social stress on human efficiency. Social stress can be defined in a number of ways (e.g. McGrath, 1970), and experimental studies have ranged from investigations of intimate stress situations, such as interpersonal conflict, to stress encountered in a wide variety of larger social contexts. Examples of the latter type include studies of bystander apathy, conformity and anxiety in examination situations. Similarly, human efficiency can be examined at almost all levels of human functioning from basic biological processes to personal adjustment, adaptation to change or to the performance of some specific task or job. All of these disparate situations and activities could reasonably be regarded as important, perhaps essential, to the understanding of human processes and behaviour.

In addition to the specialized type of social stress research mentioned above, there is also a long tradition of work in experimental social psychology which has focused on social situations of a less overtly complex nature. This work evolved in the 1920's, largely due to the publication of Floyd Allport's (1920, 1924) studies on 'social facilitation'. Allport and others argued that the mere physical presence of other individuals was a sufficient condition for producing changes in the performance behaviour of target individuals. If this is true then reactions to other, more complex types of social stress would presumably depend on interactions with the specific situation and this fundamental response to social stimuli.

The following chapters address several questions. For example, is the mere presence of others a sufficient stimulus for producing performance changes? Also, are these changes of the type described by Allport, and later by Zajonc (1965); i.e. facilitative? Finally, provided this basic reaction can be demonstrated, what processes or mechanisms mediate between the perception of the social situation and the resulting response? This last question may involve either the examination of fundamental processes common to all individuals or the variation between individuals in their sensitivity to mild social stimulation.

Task performance was chosen as a suitable measure of human efficiency for several reasons. Firstly, it is a legitimate activity which, in one form or another, engages most people for substantial portions of each day. Secondly, changes in performance can be reliably measured, and, provided 'sensitive' tests are used, can yield information about complex mental processes. More importantly, a large body of experimental work already exists concerning the effects of environmental stressors (e.g. noise, heat) on performance. This seems particularly relevant to social stress research because of the emphasis apparent in both sets of literature on the possible mediating effects of arousal level. Considered in this context, the present work is concerned with 'stress' situations rather than 'distressing' situations. The former implies the imposition of some force or constraint which elicits attempts at adaptation or coping. Stress may conceivably take many forms (e.g. heat, noise, or others' presence) and research is needed concerning both responses to stress situations in general, and specific reactions more dependent on particular characteristics of different stress stimuli.

The situations examined in the following chapters are termed 'non-interactive', although this refers only to the absence of <u>verbal</u> interchanges between subjects. Nonverbal communication may take place even when other forms of interaction are curtailed. The adjective non-interactive is merely used to distinguish the highly restricted social situations investigated here from the more usual type of small group research which occupies a considerable part of the social psychology literature.

Chapters I and II review the literature on performance in non-interactive social situations. The latter chapter is concerned specifically with situations in which evaluation manipulations, rather than the social environment per se, are the independent variables. There are no distinct guidelines for discerning the range of social situations which subjects might construe as evaluative. Some authors have argued that all social performance situations are probably interpreted in this way. Although the literatures reviewed in these two chapters may in reality overlap to a great extent, they are methodologically different and are discussed separately in order to avoid confusion and also highlight possible areas of overlap. Since the literature on performance in social situations is voluminous, the review is necessarily selective. The aim has been to discuss a number of important issues concerning both non-interactive situations and social stress manipulations.

Chapter III considers some of the more important methodological problems in the area of social stress research and describes the rationale for the experiments presented in Chapters IV, V and VI. An overview and discussion of these results are undertaken in Chapter VII. Because of the detailed nature of the performance data, only significant findings are presented and discussed in the text. Analysis of variance summary tables, means and standard deviations for all treatments can be found in the Appendices as can the exact instructions used for various tasks and conditions.

ACKNOWLEDGEMENT

There are several people who, by their assistance or influence, have contributed in some special way to the present work. Firstly, my supervisor, Dr. G. R. J. Hockey, offered much valuable advice and criticism throughout the research period and was instrumental in the decision to give performance measures and task components a central focus in the research.

Secondly, Mrs. C. Thompson and Mr. Alan White were particularly helpful in overseeing the computer analysis of the data. The Psychology Department technical staff deserve mention for their good natured and efficient preparation of equipment, particularly Malcolm Rolling, who also produced the graphic work in the following pages. In addition, I would like to thank Mrs. Margaret Hall, who took painstaking care in typing this manuscript.

The following schools are acknowledged for their cooperation and good will during the time when experimental work was under way on the premises: Deanery Secondary School, Chester-le-Street; Framwellgate Moor Comprehensive School, Durham; Houghton-le-Spring Comprehensive School, Houghton-le-Spring; Johnston Secondary School, Durham; Thornhill School, Sunderland; and Wearside Secondary School, Durham.

Finally, I would like to extend special thanks to my parents for their constant encouragement and support during the period of this research.

TABLE OF CONTENTS

Page

LEGEND OF TABLES

LEGEND OF FIGURES

Page

CHAPTER I

THE EFFECTS OF NON-INTERACTIVE SOCIAL SITUATIONS ON PERFORMANCE

1.0 INTRODUCTION

The present chapter consists of a review and discussion of individual performance in a wide range of situations in which social interaction between individuals is limited to non-verbal cues. Typically these social situations are referred to as the 'mere presence of others' (Zajonc, 1965). Mere presence studies can be further sub-divided into those concerned with coaction, in which the individual performs a task in a group while others are working independently on the same task, or audience presence, in which the subject is observed by one or more individuals who are not engaged in task performance.

The largest part of the chapter is devoted to experimental work investigating social facilitation effects and the various theoretical models associated with this literature. However, 'social facilitation' is merely the term which has come to be loosely applied to studies involving audience and/or coaction manipulations. It is not to be assumed that such manipulations necessarily or even usually lead to facilitation of performance or performance related behaviour.

In addition to performance per se, some attention is given to the wider issues surrounding individuals' perceptions of and reactions to others in the immediate environment. Although discussion is confined to variations of coaction and audience situations, it is interpersonal rather than performance processes which are of interest. It is assumed that, even in social situations which are by definition 'minimally social', behaviour and subsequent performance is influenced to some degree by both these processes.

Also, some discussion is devoted to the methodological problems attached to investigations employing performance measures, since even very simple tasks involve quite complex processes which can interact with situational variables. Considerable attention is given to the role of arousal in mediating between environmental factors and performance outcomes. Although arousal is a topic of wide-spread experimental interest, discussion is necessarily limited to those theories which are directly relevant to the circumscribed situations of interest in this thesis. The Hull-Spence concept of drive (which later became synonomous with arousal (Duffy, 1962)) and habit strength are given particular emphasis, but it is recognized that this is only one theoretical approach to the study of arousal and the changes in behaviour associated with variations in arousal level.

1.1 SOCIAL FACILITATION - AUDIENCE AND COACTION EFFECTS

Following the publication of Allport's (1920, 1924) studies, experiments concerned with the effects of incidental social stimuli (the mere presence of others) on performance proliferated. Allport found a facilitative effect when subjects performed simple tasks (word associations, multiplication, cancellation and reversible perspective tests) in small groups as opposed to performance alone. This increment was attributed to the increase or hastening of individual responding due to the movement of others performing the same task.

Subsequent investigations agreed that the presence of others did affect performance. However, some authors reported facilitation while others found impairment. With the advent of the Second World War experimental interest was redirected toward more rigorous investigations of performance processes per se, and these inconsistencies were left unresolved until Zajonc (1965) offered a tenable interpretation of this early work based on drive theory.

Essentially Zajonc (1965) argues that the mere presence of others is a source of arousal. Based directly on Spence's (1958) restatements of Hullian drive theory, predictions are that increased arousal will facilitate performance when the dominant response (response with the greatest habit strength) is correct but will result in impairment when it is incorrect. Therefore, the presence of others would be expected to facilitate performance on tasks which are easy or well learned but to produce the opposite effects on those requiring learning (where the correct response is of initially low habit strength).

The drive theory of social facilitation has undergone some modifications in the light of subsequent research. However, the 1965 paper itself deserves some detailed attention, partly because it is often referred to and has stimulated considerable research, but also because it is based on a literature riddled with methodological flaws and tentative findings. In addition it does little to resolve the basic questions of what social facilitation is and under what conditions it occurs.

At present, definitions for the former are abundant and there is wide-spread controversy concerning the latter. The following sections consider Zajonc's theory in detail, the literature on which it is based, and the two issues stated above. They attempt to describe the foundations for recent research on this topic and establish the framework within which the experiments reported in Chapters IV, V and VI were conducted.

1.1.1 Definitions of Social Facilitation

Before beginning an evaluation of the literature and theory concerning the study of social facilitation, it might be prudent to examine some of the commonly understood meanings ascribed to the term and identify areas of potential confusion. Firstly, social facilitation is regarded by many social psychologists as an established and fundamental product of group situations. Newcomb et al (1965) give a definition that mirrors that of Allport (1920), that "the sights and sounds resulting from group members doing the same thing serve to intensify the doing of it on the part of all of them." (p.286). Clearly this definition applies only to coaction settings. Jones and Gerard (1967) offer a somewhat different explanation, i.e. "the presence of others has an energizing effect on the subject, causing him to work with greater intensity or higher motivation." (p.602). In this case it is not clear what role the others may have. Scott (1968) declares that social facilitation has been 'clearly defined' by Crawford (1939) as "any increment of individual activity which results from the presence of another individual, and can be regarded as one of the most basic forms of social interaction." (pp.410-411). This is similar to Jones and Gerard's (1967) statement, although the phenomenon is operationally defined as 'any increment of any activity' rather than greater motivation for the task at hand.

It is this last, very general definition which seems to have been the most widely accepted, probably because it takes a diversity of behaviours into account. But its generality renders it largely useless since no guidelines are offered as to what effects different social situations should be expected to have or what behaviours might be most influenced. It will be seen in subsequent sections that there is great variation in the response of individuals depending on their relationship to the others in the environment, and there is also some question about whether social facilitation effects are in fact a basic or universal response to social stimuli or are rather trivial, incidental reactions.

In addition to inadequately defining the social situations of importance, there is some confusion in the literature as to the range of phenomena attributable to social facilitation. For example, one early experiment (Abel, 1938) is entitled "The Influence of Social Facilitation upon Motor Performance ...." Lorge et al (1958) make a similar remark, i.e. "not only can actual production be affected by social facilitation ...." (p.353). However, if the phenomenon is defined as an increase in productivity, it cannot be said to influence or affect production. To use the term in this manner makes it synonomous with social stimulation. As late as 1969 Mann wrote, "Social facilitation deals with the effects on an individual of working at a task in the presence of other individuals, but independently of them." (p.70). In this case, 'effects' could presumably be anything, not even facilitation. Zajonc (1965) avoids this problem by stressing that facilitation refers to the increased emission of dominant responses, although this is superficially confusing since the dominant response can be correct or incorrect and can result in impairment of performance as well as facilitation. If anything is clear from a scan of the literature it is that social facilitation has not been clearly defined.

Still, whatever definition is ultimately chosen, it would seem that two criteria need to be met. Firstly, some increment in some behaviour must be observed before an effect can be facilitative. Secondly, the stimulus seemingly takes the form of other people in the immediate environment. Beyond these two general observations, further definition and judgment concerning the importance of social facilitation effects are reserved until the relevant literature and theories dealing with the phenomenon have been reviewed and evaluated.

1.1.2 The Drive Theory of Social Facilitation

The cornerstone of Zajonc's (1965) drive theory is the stipulation that the mere presence of others is arousing and leads to the emission of dominant responses. From this formulation specific predictions can be made regarding performance outcomes (provided the response hierarchy is known), and performance impairment as well as facilitation can be accounted for within the same theoretical framework.

Zajonc's support for this interpretation is based on his review of the literature from the 1920 - 1940 period, discussed in terms of the effects of mere presence on learning (dominant response being incorrect) and performance (dominant response being correct) tasks. A second source of evidence comes from investigations which specifically test for changes in arousal level through the differential performance of tasks on which the habit strengths of competing responses are either known or are established through specific training.

Although Zajonc's theory has been heralded for its apparent ability to account for the discrepancies in the early studies, there is some question as to whether such acclaim is justified. Firstly, the reliability of some of the findings reported in Zajonc's review is accepted rather uncritically. Secondly, the review is highly selective, only mentioning the better known experiments, and there are some problems even with the discussion of these.

A more thorough overview of the work from the pre-Zajonc period is presented in Tables 1.1 and 1.2, which list studies concerned with coaction and audience effects respectively. In some cases reanalysis of the data was required and this was undertaken when raw data were available. When this was not possible, strong trends have been noted without regard to levels of significance. Effort has been directed toward pointing out uncontrolled variables and possible sources of error in the experimental procedures as well as alternative interpretations of the results.

In terms of audience effects, Zajonc (1965) reviews six studies which fit a drive theory interpretation (Travis, 1925; Bergum and Lehr, 1963; Dashiell, 1930; Pessin, 1933; Husband, 1931; Pessin and Husband, 1933). However, the results from two of these (Dashiell, 1930; Bergum and Lehr, 1963) can be accounted for by influences other than mere presence (see Table 1.2), Pessin and Husband reported no significant

TABLE 1.1

Summary of Studies on Social Facilitation (Coaction vs Alone)

AUTHOR(S)	SIZE OF GROUP	TASK(S)	RESULT(S)	COMMENTS
Triplett (1897)	variable	cycling	greater speed in social condition	
Burnham (1910)	classroom	classroom tasks	greater efficiency in social condition	
Allport (1920)	(3 to 5)	word association	consistently greater speeds in social condition, significant in 1 out of 5 experiments (sign test)	Subjects reported competitive feelings in group setting. E present in group condition and not in alone condition, possibly constituting an additional social stimulus
		disproving philosophical arguments	more words and ideas written in group. More superior ideas when alone	
Allport (1924)	(3 to 5)	multiplication, cancellation, reversible perspective	consistently greater speed in social condition, not significant (sign test)	same as above
Sengupta & Sinka (1926)	(5)	cancellation	significantly greater speed in group condition	Design not balanced to control for novelty of group settings after long series of alone trials. No attempt to control for competition between subjects

Continued/...

AUTHOR(S)	SIZE OF GROUP	TASK(S)	RESULT(S)	COMMENTS
Weston & English (1926)	(10)	intelligence test	higher scores in group	Different forms of test used for alone and group condition. Crude method of comparison - decile rankings. Possible confounding of results with practice effects
	(50-75)	intelligence test	no difference	
Farnsworth (1928)	(classroom)	intelligence test	no differences	
Anderson (1929)	(5)	arithmetic problems, cancellation, marble sorting	slight support for bright Ss' performance being impaired and normal Ss being facilitated in social conditions	Inappropriate statistical analysis[1]
Abel (1938)	(pairs)	maze tracing	no conclusive results	Suspicious statistical analysis.[2] No attempt to control for competition between subjects
Dashiell (1930)	(15)	multiplication, word analogies, associations	no differences	
Mukerji (1910)	(10)	cancellation, naming capitals	greater speed in social condition on both tasks	Possible novelty effect from group since Ss had practiced alone. No attempt to control for competition between subjects

Continued/...

AUTHOR(S)	SIZE OF GROUP	TASK(S)	RESULT(S)	COMMENTS
Bergum & Lehr (1962)	(2)	vigilance	non significant trend for better performance in pairs	Some comparison between individuals likely since when one performed well or badly his companion showed the same tendency
Ader & Tatum (1963)	2 subjects yoked together	avoidance learning (shock)	yoked subjects showed poorer learning	Situation was not strictly passive presence of another since target subject was responsible for his yoked mate

Notes:

1. There is some apparent error in the calculation of the standard error of the difference which forms the denominator for the critical ratio. Also this method of analysis is probably inappropriate considering that in each condition N equals only 5 subjects.

2. The critical ratios obtained are suspiciously high considering the small mean differences, possibly due to miscalculation of the probable error of the difference, based on the semi-interquartile range (Garrett, 1958). In one case the C.R. (similar to the t statistic) reaches 21.51 with a mean difference of only 2.26 seconds.

TABLE 1.2

Summary of Studies on Social Facilitation (Audience vs Alone)

AUTHOR(S)	SIZE OF AUDIENCE	TASK(S)	RESULT(S)	COMMENTS
Gates (1924)	Alone vs small (4-6) audience vs large (27-37) audience	coordination, colour naming, analogies, naming adjectives	no differences	E present in alone condition, possibly minimizing difference between conditions
Travis (1925)	(4 to 8)	pursuit rotor performance	significantly higher TOT scores in social condition (Horne, 1971)	E present in alone and social condition
Ekdahl (1929)	Experimenter	word association	fewer associations given in social condition	
Burri (1931)	alone vs attentive vs inattentive audience	learning and recalling nonsense syllables	recall superior in alone condition. More trials to learn for subjects given warning of future audience condition	E present in alone and social conditions

Continued/...

AUTHOR(S)	SIZE OF AUDIENCE	TASK(S)	RESULT(S)	COMMENTS
Pessin & Husband (1933)	alone vs unseen audience vs seen audience	learning finger maze	no significant differences	E present in all conditions. Subjects blindfolded in two conditions, perhaps hindering orientation and thus performance
Pessin (1933)	Experimenter	learn and recall nonsense syllables	impairment of learning in audience condition, facilitation of recall 3 days later in audience condition	Superior performance on third recall day in audience condition could have been due to more practice trials on first day
Dashiell (1930)	(2)	multiplication, analogies, word association	tendency for greater speed somewhat lower accuracy in audience condition	Audience consisted of probable competitors, i.e. other subjects
Bergum & Lehr (1963)	Subjects' superior officer	vigilance	performance better when subjects were monitored (social condition)	Observation was only at random intervals with audience commenting on Ss' performance. Therefore, audience was not merely present and probably was viewed as evaluative

differences between control and experimental subjects and a fourth study (Husband, 1931) does not even employ an audience manipulation.[3]

The evidence offered for coaction effects is even less satisfactory, as it consists largely of animal studies, and the results from several of these can only be accepted with qualification (Harlow, 1932 and Gates and Allee, 1933, in particular). Tolman (1968) has compiled an extensive review of the animal literature concerned with behaviour changes in the presence of other animals and cites drive theory as only one of a number of viable explanations for the variety of effects observed.

Among the experiments cited using human subjects, Allport's (1920, 1924) studies only demonstrate a tendency for facilitation in coacting groups and Dashiell (1930) found no differences between the performance of subjects working together or in separate rooms.

Even if the animal studies could be construed as offering evidence for the facilitation of dominant responses in group settings, generalization between these data and that of human performance is of questionable validity. Most of the behaviours measured have been survival related, i.e. eating (Bayer, 1929) or nest building (Chen, 1937) and have been ones which normally occur in a social context. In fact a persistent observation in the animal literature is the apparent distress and behaviour disruption shown by many species when placed in isolated conditions (Scott, 1968; Rajecki et al, 1975; Stamm, 1961; Clayton, 1976), implying that the facilitation sometimes demonstrated in social conditions may be the product of reduced fear or stress rather than increased arousal. It is also clear that imitation of the other animal's behaviour may be a significant factor (Zentall and Hogan, 1976). In

3. Pessin and Husband (1933) merely alluded to the possible disruptive effects of spectators as part of the rationale for their later investigation which did vary audience presence.

general, the animal literature is concerned with situations essentially different from that of the human subject performing a task in response to specific instructions and in a setting which is rigidly structured and quite often novel. Performance in this latter circumstance is not spontaneously occurring behaviour and is likely to be mediated by different processes.

Leaving Zajonc's review for the moment and considering the larger body of experimental work presented in Tables 1.1 and 1.2, it can be seen that few experiments employ exactly the same social situation with the same tasks so that generalizations must be based on findings obtained under widely differing conditions. Little evidence is present to support the notion that coaction situations, in general, affect individual performance in any predictable way. In cases where effects do occur, the results are just as likely to be due to competition as to mere presence, and this is even noted in the animal literature (Harlow, 1932). Only in the Triplett (1897) study is competition more or less held constant over alone and group conditions so that the effects of the presence of another can be inferred.

Findings from experiments dealing with audience effects are equally confusing. It is not the case that well-learned, routine tasks are facilitated by audience presence, since recalling nonsense syllables, giving word associations and coordination tests have shown in some cases impairment or no differences (Gates, 1924; Ekdahl, 1929). There is also confusion over what constitutes an audience or, for that matter, a coaction situation. Examples are given where the presence of the experimenter has been associated with either facilitation or impairment of performance. Yet in other studies the experimenter is present in all conditions and 'audience' is defined as one or more additional spectators. Furthermore, audiences have been composed of probable competitors

(Dashiell, 1930) and group and audience size has varied from one additional person to 50 or more. Similarly, coaction studies have been confounded by the presence of the experimenter in one or both conditions (Allport, 1920, 1924) and have consistently reported competitive influences.

It seems fairly clear that these social situations can affect individual performance. However, they do not always do so and the effects are often different between individuals. The diversity of findings is not explainable solely on the basis of differences in the tasks used, as most have measured the speed of execution of either easily mastered or well-learned activities. In some cases learning tasks were employed, but results have been inconsistent. Few of these tasks have permitted a detailed analysis of performance components, and, although intra-subject variability was occasionally measured, no very strong trends have been reported.

In 1935 Dashiell reviewed the literature on the effects of social situations and behaviour and concluded that the findings attributed to coacting groups were equivocal and that little had been added since Allport's studies, apart from strong evidence that 'pure' alone, coacting or even spectator conditions were remarkably difficult to create experimentally. His conclusion about audience effects were equally noncommittal because of the range of situations and tasks used. Oddly, Hollingworth (1935) concluded the opposite i.e. that the effects of being observed by a passive audience were about the same as those of working in a coacting group; thoughtfulness and originality were likely to suffer, although activity may increase. This conclusion seems unwarranted, and the merging of audience and coaction effects is misleading. For example, although Allport (1920) found a slight increase in the speed of word association in coacting groups, Ekdahl

(1929) found the opposite effect in an audience situation.

Fortunately, there is now a sizeable literature which specifically addresses a drive theory interpretation of mere presence effects, so the inadequacies of Zajonc's review are less crucial than would otherwise be the case.

Typically these experiments employ the paradigm devised by Spence et al (1956) in their work on drive and the learning of competitional and noncompetitional paired-associates. Cottrell et al (1967) found that the performance of observed subjects was impaired on highly competitional lists. Similarly, on pseudo-recognition tasks, in which habit strength was established via training, the presence of an audience has resulted in an increase in the frequency at which subjects report words which have been practised more (Zajonc and Sales, 1966; Henchy and Glass, 1968). It has also been shown that subjects give word associations of greater commonality (according to existing norms) when tested with an observer present (Matlin and Zajonc, 1968) and are more conservative (dominant response) in risk-taking in both audience and coaction situations (Zajonc et al., 1970).

Using learning versus performance manipulations of habit strength, an audience has been found to impair <u>learning</u> of a complex motor task but facilitate <u>performance</u> of the same task (Martens, 1969a; Haas and Roberts, 1975), and coaction has yielded similar results for maze learning (Hunt and Hillery, 1973). Also, an audience has been found to impair short term recall in paired-associate learning but to facilitate long term recall (Geen, 1971, 1973; Deffenbacher et al, 1974).[4] When varying task difficulty, facilitation has resulted from the mere

4. These studies are based on the observations that high arousal impairs short-term recall, which occurs during the perseverative process, but facilitates long-term recall because of enhanced neural consolidation (e.g. Walker and Tarte, 1963).

presence of others for simple digit symbol, figure drawing and maze learning but has been associated with impairment on complex versions of these same tasks (Shaver and Liebling, 1976; Kiesler, 1966; Hunt and Hillery, 1973).

All but the last two of these studies employed an audience manipulation. However, Cottrell (1972) has reviewed several studies in which coaction has been associated with increased frequency of dominant responses, where the responses were the subjects predetermined preference for colours or stimulus position (left/right).

All of these reports indicate that the presence of others can facilitate the emission of dominant responses in the ways predicted by arousal theory.

1.1.3 Others as a Source of Arousal - Mere Presence or Potential Evaluation?

Zajonc (1965) has argued that the mere presence of others is arousing. Presumably such arousal was held to be innate, and an 'instinct' interpretation fits with the author's concentration on animal studies (e.g. Zajonc et al, 1969). However, several researchers working within a drive theory framework have failed to obtain results consistent with the theory from manipulations of mere presence (Thayer and Moore, 1972; Martens and Landers, 1972; Klinger, 1969; Innes, 1972; Hardesty et al, 1963; Paulus and Murdock, 1971; Henchy and Glass, 1968; Cottrell et al, 1968; Pederson, 1970; Wankel, 1975; Paulus et al, 1972; Singer, 1970; Freishlag, 1974; Shrauger, 1972; Carment and Latchford, 1970; Carron and Bennett, 1976; Borden et al, 1976).

Other studies have reported some support for the theory but with qualifications. For example, Innes and Sambrooks (1969) found a complex interaction between coaction, sex and birth order for paired-associate learning, and in another study audience presence was associated with poorer learning of the competitional list but produced no differences on the non-competitional one (Berkey and Hoppe, 1972). In an interesting study by Good (1973) an audience resulted in higher commonality scores on word association only if subjects were led to believe that they would do well on the task and that the audience was evaluating their performance, suggesting that different task strategies were adopted according to the situation. In addition, Geen (1976) found no main effect due to audience presence on the range of cue utilisation,[5] but obtained an interaction between audience and subjects' anxiety level.

Cottrell (1968) has argued that subjects only demonstrate the predicted performance outcomes when others are seen as sources of positive or negative evaluation. In his original investigation (Cottrell et al, 1968) no differences were found between the performance of subjects tested on a pseudo-recognition test when either alone or in the presence of two blindfolded subjects. The emission of dominant responses was enhanced, however, when the observers were not blindfolded and could see the subject. The latter constituted a source of potential evaluation in that the observers were able to informally judge the subjects' competence, even though evaluation was not specifically mentioned.

Other studies have directly manipulated evaluative cues via instructions emphasizing either the expertise of the audience members or their interest in the subjects' performance level and have shown the predicted effects when evaluation was implicated but yielded only tentative support for mere presence (Henchy and Glass, 1968; Paulus and Murdock, 1971; Smith and Crabbe, 1976; Musante and Anker, 1972). Similar results have been obtained in coaction studies using reaction

---

5. Cue utilisation is expected to be restricted under conditions of high arousal (Easterbrook, 1959).

time, vigilance, paired-associate learning and motor performance (Innes, 1972; Klinger, 1969; Thayer and Moore, 1972; Martens and Landers, 1972). Furthermore, Carment and Latchford (1970) found that simple motor responding was unaffected by the presence of coactors but was facilitated by the presence of the experimenter (in both alone and coaction conditions).

The fact that mere presence sometimes but not always affects performance casts doubt on its status as an innate source of drive. In a thorough review of the literature following Zajonc's (1965) paper, Cottrell (1972) concludes that arousal due to the presence of others is a learned source of drive based on the individual's previous experience in performance situations. Weiss and Miller (1971) support this interpretation and liken the drive induced by audiencesto aversive states such as frustration or anxiety.

Assuming that the presence of others, in the context considered here, is a learned source of arousal, then the problem for future research becomes one of identifying what features of the social environment act as evaluative stimuli for subjects and how these vary in both the degree and manner in which they influence performance.

1.1.4 Some Comments about Audience and Coaction Studies

The process of identifying the salient features in social situations raises many problems, and recent literature offers little illumination. Much of the difficulty arises from basic differences in definitions of social facilitation, what behaviours this term should be restricted to and the failure to differentiate between the effects of different social environments. In addition, there is still a lack of attention to clearly defining 'audience', 'coaction' and 'evaluation', even in recent literature. This is partly due to an abundance of single, isolated studies, usually designed to demonstrate social facilitation on some specific task and under some specific condition (e.g. Rittle and Bernard, 1977). Apart from crude similarities such as the presence or absence of others and in some cases the type of evaluative manipulation used, these studies are as disparate as those listed in Tables 1.1 and 1.2: classification and summary on any finer basis is next to impossible.

Although evaluation potential as a source of arousal has been the most popular explanation in the literature to account for the effects of others' presence, there are important differences between coaction and audience situations, both in the number and type of evaluative cues possible. Foot (1973) points out that coactors can be used as sources for both performance and emotional feedback and comparison; such information is perhaps available from audiences but would be more difficult to assess. In addition, although subjects may fear evaluation from observers or coactors, competitive behaviour is possible in coaction settings; this would hardly be a feature of performance in the presence of an audience. For example, Innes (1972) found no effects on reaction time due to the mere presence of a coactor, but facilitation occurred when feedback was provided as to which of the pair had responded fastest; no further improvement occurred when evaluative instructions were added. Similar results have been noted for vigilance performance (Klinger, 1969), and impairment on a motor learning task resulted when coactors could make direct observations of each others' performance (Martens and Landers, 1972).

Additional evidence of performance comparison comes from two studies in which subjects were allowed access to their own and coactor's scores. In the coaction situations subjects not only requested the other subjects' scores, they increased self audits as well, suggesting a comparison of scores (Hake et al, 1973). Investigations of competitive behaviour point out that competition is greatest when opponents' skills are roughly equal (Caillois, 1961). Consistent with this view, a second experiment demonstrated that subjects requested both self and coactor audits more frequently when the two scores were about the same (Vukelich and Hake, 1974). These two studies indicate that subjects are at least interested in the performance level of coactors and will utilize sources of feedback if they are available.

1.2 ALTERNATIVE EXPLANATIONS FOR MERE PRESENCE EFFECTS

1.2.1 Theories of Attention

Although the drive theory of social facilitation is most often referred to for explaining performance outcomes in non-interactive social situations, other interpretations have been suggested.

Sarason (1972) and Wine (1971) have argued that evaluation stress impairs the performance of individuals who are fearful in testing situations because interfering responses, which are personal rather than task oriented, are elicited. Such interference takes the form of thoughts about failure or personal inadequacy, suggesting maximum performance impairment on tasks involving higher order cognitive processes. This is not inconsistent with drive theory if response hierarchies could be considered in terms of the individual's whole behavioural repertoire. In other words, test anxious subjects may emit personal, task-irrelevant responses because in evaluative situations these have greater habit strength than those which are task related.

Duval and Wickland (1972) have formulated an attentional theory similar to Sarason's. According to these authors, individuals tend to focus attention inwardly if they are confronted with a photograph, mirror or any stimulus which alerts them to their status as an object in the environment (objective self-awareness). For example, latencies to self-relevant rather than neutral words (on the Stroop test) have been shorter under conditions designed to increase objective self-awareness (Geller and Shaver, 1976). Presumably the presence of others in a performance situation need not elicit evaluation fear but may cause subjects to attend more closely to their object state i.e. appearance, level of performance. Increased objective self-awareness is expected to result in conscious increases in effort and striving and in most cases to performance facilitation. However, the authors point out that on very difficult tasks subjects may engage in excessive self-examination, because of perceived failure, and this may result in lower performance levels due to reduced concentration on the task.

Experiments have shown improved performance on simple tasks when subjects were confronted with a mirror (Wickland and Duval, 1971), but less improvement under the same conditions when evaluative instructions were added (Liebling and Shaver, 1973a). Interestingly, when <u>accuracy</u> on a star tracing task was stressed, the presence of an audience was associated with slower performance (Innes and Young, 1975), implying increased attention to the quality of performance: arousal theory would predict faster but perhaps less accurate performance.

The significance of theories of objective self-awareness would seem to be in the identification of self-evaluation as a factor in performance situations involving audience presence.

1.2.2 Dearousing Effects

In direct opposition to the drive theory of social facilitation are studies which suggest that the presence of companions decreases arousal. Most of this evidence comes from animal experiments demonstrating a reduction in fear responses to noxious stimuli when animals are in pairs (Rasmussen, 1939; Hake and Laws, 1967; Baum, 1969; Davitz and Mason, 1955). At least one such study has found similar results with humans (Seidman et al, 1957). Zajonc (1965) argues that such findings are due to an increase in the emission of dominant responses e.g. non-fearful responses. However, a few studies have found evidence of decreases in arousal based on subjects' self-reported feelings of anxiety (Wrightsman, 1960; Schachter, 1959; Buck and Parke, 1972), although some of these effects were restricted to first-born children. One study has found an increase in interaction (presumably anxiety reducing) under stressful conditions (Morris et al, 1976).

However, in a review of the literature, Epley (1974) concludes that, although the presence of others is often associated with reduced feelings of anxiety and fear reactions, it is not likely to be sufficient for such changes. In most cases where calming effects have been noted, the other persons present were confederates or were naive as to the stressful nature of the experiment. Hence, the presence of a calm companion may have calming effects on subjects who are experiencing stress.

Dearousing effects due to the presence of others are not likely to be of central concern in this discussion, as most of the situations described would not be regarded as very stressful, at least not as stressful as those mentioned above which involved threat of physical pain. However, there are individual differences in sensitivity to stressful stimuli as well as considerable inaccuracy in estimating the stressfulness of various experimental treatments. Therefore, calming effects cannot entirely be excluded from consideration.

1.3 THEORETICAL AND METHODOLOGICAL CONSIDERATIONS IN PERFORMANCE SITUATIONS

As mentioned previously, with the beginning of the Second World War interest in performance in the presence of others abated and was replaced by a much more rigorous approach to the measurement of performance;

war-time industry demanded a more precise and detailed understanding of the laws governing the learning and performance of skilled operations.

Of major impact on the performance literature has been Spence's (1956, 1958) work on drive level and performance. This is particularly relevant in the present discussion since Zajonc's (1965) theory of social facilitation is directly extrapolated from Spence's formulations. However, later modifications of drive theory have incorporated some rather more complex assumptions. For example, it is now more or less accepted that a relationship exists between level of arousal and performance which fits an inverted U-shaped curve. The original Spence theory implied that as drive increased so should performance, provided there were no competing responses. But the data indicate that there is an optimal range of drive associated with optimal performance and that levels above or below this are accompanied by less efficiency (Broadhurst, 1959). Broen and Storms (1961) offer a theoretical explanation of the inverted U-shaped relationship based on the hypothesis that all responses have an effective ceiling; once responses of strong habit strength have reached this ceiling subordinate responses will tend to be enhanced as well, resulting in the performance decrement associated with super-optimal arousal.

The presence of such a relationship complicates investigation considerably, since at present there is no reliable way of measuring the level of drive operative in most experimental situations. A situation construed to be highly stressful by an experimenter may not be experienced as such by the subjects or vice versa. Similarly, the range of situations employed in any one experiment may be too narrow to evoke low, optimal and high levels of arousal. As a result, the inverted U-shaped curve is typically employed to explain results but is seldom successful in predicting them.

Although physiological measures have been employed with some success as indicators of arousal in social situations, particularly those believed to be stressful (Malmo et al, 1957; Ulrich, 1957; Frankenhaeuser and Patkai, 1964; Luria, 1932; Leiderman and Shapiro, 1964; Thackray and Pearson, 1968; Ulrich and Burke, 1957; Mason, 1968; Malmo, 1965; Kiritz and Moos, 1974), they have only infrequently been incorporated in studies of mere presence. Chapman (1973a, 1974a) has reported higher EMG levels from subjects when the experimenter was present and Martens (1969b) noted indications of higher arousal in audience conditions based on palmer sweat prints. However, two studies found no evidence of changes in autonomic functioning due to an audience (Borden et al, 1976, Henchy and Glass, 1968). The independent assessment of arousal level via physiological recordings is problematic because of difficulties with the collection and interpretation of such data and individual differences in the patterning of physiological reactivity (Lacey et al, 1963).

In addition to the basic problem of measuring arousal levels, there is a further complication resulting from the relationship between task difficulty and arousal (Yerkes and Dodson, 1908). Specifically, optimal performance of a difficult task would be predicted to occur at a lower level of arousal than would be the case for a simple task. Clearly some tasks can, with reasonable certainty, be classified as simple or difficult, but these represent the extremes. To date, there is no classification system for tasks which is universally accepted, despite numerous discussions on the topic (Fleishman, 1975; McGrath, 1970; Hackman, 1970).

Another confounding factor is the possible interaction between habit strength and arousal over time (Hicks, 1975). The performance of tasks on which the correct response is of initially low habit strength would be expected to suffer under conditions of high drive. However, as learning proceeds, correct responses are reinforced and assume a higher position on the response hierarchy. Therefore, although high drive may be initially detrimental, it may facilitate performance in later trials and camouflage any overall effects. Such an interaction would of course only be expected for tasks on which improvement could occur and mastery was not accomplished in just a few trials.

In addition to task difficulty, there are other task variables which are problematic. For example, tasks vary in their degree of intrinsic stressfulness. What may be an essentially easy task to learn and execute may be experienced as stressful because of time pressure or monotony or other unpleasant features such as noise (Fitts and Posner, 1967; London et al, 1972), resulting in higher levels of arousal.

In summary, any discussion which offers predictions about performance based on subjects' level of drive or arousal must also take into consideration task variables and temporal factors as well as the possibility of a non-linear relationship between drive level and performance.

Although the Spence theory of drive and performance has been given most attention both in this discussion and in the social facilitation literature, there are other theories concerned with arousal and performance which have utility for the situations of interest here. Easterbrook (1959) has argued that increases in arousal lead to restricted cue utilization i.e. a narrowing of attention or focusing on central aspects of the task. Such an effect has been found in studies employing manipulations of Manifest Anxiety or endogenous drive (Zaffy and Bruning, 1966), threat of electric shock (Kohn, 1954), as well as in audience situations (Geen, 1976).

Effectively, in terms of performance in social situations, the implications of Easterbrook's (1959) theory may not be much different from those based on habit strength i.e. the performance of complex tasks, where a number of cues must be simultaneously attended to, would be expected to suffer, whereas performance of simple tasks would be facilitated due to increased focusing of attention. However, there are many tasks, such as vigilance or digit span, for which there are no easily perceivable response hierarchies. For tasks which are attention/perception based and for which learning is not required or is impossible, cue utilization may be a more appropriate model for making predictions based on inferred arousal level.

In terms of experimental methodology, the introduction of analysis of variance has considerably altered the style of psychological research from that described in Tables 1.1 and 1.2. The advantages of being able to estimate the combined effects of variables are obvious. On the other hand, the use of factorial designs and large numbers of subjects has meant the abandonment of the very rich, detailed analysis of individual performance which is the major strength of Allport (1920, 1924) and Dashiell's (1930) work. One of the outstanding findings in this early literature was the great variation in individual susceptibility to influence from the presence of others, often resulting in performance changes in opposite directions. Individual differences tend to be overlooked in factorial designs, unless specifically manipulated, and hence a possible source of explanation and hypothesis generation has been lost.

Another methodological consideration in more recent performance literature is the emphasis on employing "sensitive" tasks. Lazarus et al (1952) and Underwood (1976) have argued that concentrating on overall performance measures often obscures important differences in the strategies that subjects use for executing tasks and offers little illumination of the actual processes affected by the experimental manipulations. They suggest analysing performance in terms of task components. Similarly, Fitts and Posner (1967) have argued that experimental tasks should provide measures of output, error and variability, as not only do gross measures provide little insight into performance processes, but they may simply not reflect changes in performance which a finer analysis might highlight.

Clearly it is difficult to bear in mind all of the above considerations in any one experiment, and the literature on performance in social situations can be justly criticized for its lack of proper controls, post hoc explanations of results and inadequately detailed data analysis. This has resulted in the near replication of much experimental work, which is both costly and confusing. However, equipped with relatively crude methods for controlling and measuring stress stimuli, it is unlikely that great improvements in experimental design and methodology in this area of research will emerge in the near future. Certain problems have been highlighted and these should be given attention. Also, the results from single experiments, despite the elegance of their design, should be regarded with caution until there is a better understanding of which effects are general and which are related to specific tasks, procedures and situations.

1.4 GROUP PROCESSES AND INDIVIDUAL PERFORMANCE

It has been suggested that non-interacting groups vary in the degree to which members provide performance and/or emotional information to other group members (Foot, 1973). Also, it seems likely that tasks which enable the assessment of performance by others will alert subjects to evaluative features in the situation. However, there is little information available concerning the ways in which performance feedback from others affects actual performance.

Festinger (1954) has outlined a theory of social comparison which maintains that subjects in groups try to evaluate their performance level by comparing it with those of other members of the group. If the individual perceives his performance to be discrepant, efforts will be made to reduce the discrepancy. That subjects seek comparison with others on various personal dimensions is supported by several studies; performance (Wheeler, 1966; Schwartz and Smith, 1976), emotional state (Gerard, 1963), opinions (Pettigrew, 1967) and personality test scores (Gruder et al, 1975). But, whether or not this comparative information is then used as the basis for altering the attribute in question has not often been directly tested.

Most texts concerned with group processes discuss the pressures toward uniformity that are exerted on individual group members (Davis, 1969; Kelley and Thibaut, 1969; Secord and Backman, 1974; Kiesler and Kiesler, 1969), and it is from these sources that Festinger's second proposal receives some indirect support. Considering the centrality of this notion in the social psychology literature it is curious that theories concerning mere presence effects have not encompassed the possibility of individual compliance to group standards. Allport (1924) reported a series of experiments which suggested that individual judgments of odours and weights were less extreme in group settings, and there is ample evidence from the literature on conformity (patterned on Asch's (1952) work) that discrepant feedback from others can influence an individual's judgment of lengths of lines (Asch, 1952), the autokinetic effect (Sherif, 1936), opinions, attitudes and answers to logic problems (Crutchfield, 1955). Kelley and Thibaut (1969) noted tendencies toward uniformity in group problem solving situations even when unanimity was not emphasized, and point out that poorer performance in 'brainstorming' groups is at least partly due to the group's tendency to pursue a single line of thought for an unduly long time.

There are various explanations given to account for conformity behaviour, but most focus on the individual's fear of rejection should performance be perceived to be discrepant from the rest of the group (Elms, 1972; Festinger, 1954; Kiesler and Kiesler, 1969). Singer (1966) argues that the comparison process is more general than Festinger maintains, often involving several relevant attributes and mainly concerned with protecting or enhancing self-esteem; such motivation has appeared to be central in at least two experiments examining social comparison (Thornton and Arrowood, 1966; Hakemiller, 1966).

That such pressures exist within group performance situations seems all the more reason to consider audience and coaction conditions separately. Although it is known that knowledge about personal attributes of an anticipated audience can affect the content of subjects' recall (Grace, 1951; Schramm and Danielson, 1958), sources of performance evaluation and comparison are usually not available in audience settings. Alternatively, the conformity studies mentioned above involve public performance and usually verbal feedback, which is not a feature of mere presence studies. Allport's (1924) subjects wrote their judgments privately. Even so, although his results could be explained in terms of dominant responses or cue utilization, it is possible that the group situation encouraged the production of standard, normative responses. Most experimental tasks provide some degree of feedback about the performance of other group members e.g. time to finish, pages completed, and Allport's subjects would have had access to coactors' facial expressions and possibly other reactive cues.

The degree to which individuals respond to uniformity pressures depends on personal and situational factors. Crutchfield (1955), on the basis of personality tests, found that those who conformed were less effective intellectually, less mature in social relationships, less confident and higher in need for social approval. Uniformity is also related to the ambiguity of the situation. For example, Asch (1951) found conformity to be positively related to task difficulty, and Nordholm (1975) refers to the "classic" effect of stimulus ambiguity on conformity. Similarly, Singer (1966) points out that, although dissimilar others are usually not chosen for comparison, when the anchor points of a distribution are unknown, discrepant others may provide useful information about the appropriate range of performance. Usually in laboratory experiments, performance norms are unknown. In such situations it is likely that social comparison would be elicited and that individuals might attend especially to the range of group members' performance in order to establish a subjective group norm. In turn, awareness of others' performance level may to some degree modify the individual's efforts or strategy in executing the task.

Theories dealing with audience and group effects have concentrated largely on performance outcomes (output, errors), overlooking the possible influence that basic group processes may have on these outcomes.

1.5. OVERVIEW

In conclusion, it would seem that the presence of others can increase individual performers' level of arousal, this being reflected in the facilitation of dominant responses. The reliance on habit strength as a predictor of performance, however, presents some problems. As has been mentioned, except for a limited number of tasks and under controlled conditions, it is usually not possible to determine the individual's response hierarchy on any particular activity. This of course does not invalidate the theory, but does restrict its applicability.

On the basis of the experiments reviewed, there is strong evidence that performance in non-interactive social situations is dependent on learned responses to evaluative/competitive stimuli. However, this does not imply that mere presence has no effect on behaviour. For example, it has been shown that children laugh more in the presence of other children when listening to humorous stories. This facilitative effect was found when the others were present as co-listeners or as an audience and when the material listened to was different between coactors (Chapman, 1973, 1974b; Chapman and Chapman, 1974). However, if some fundamental social response is activated due to the physical presence of others, this response is likely to be masked or minimised in performance situations where there are well established learned patterns for appraising and coping with potentially threatening stimuli. It was suggested that emphasis be placed on identifying which features of the environment are responsible for differential performance outcomes and which processes (social and performance) are affected.

It has also been stressed that an audience and a coacting group are essentially different social situations. They vary in their degree of evaluation potential, feedback availability and possibilities for competition or comparison. Each of these two general categories can in turn be subdivided on several dimensions. Foot (1973) has presented a diagrammatic outline of audience and coaction situations which stresses several of the points mentioned in the foregoing discussion, particularly the motivational and informational properties of coactors. However, his scheme identifies only the more overt, physical aspects of audience situations (visible/invisible, active/passive), and it is not clear how these features are related to the motivational (evaluative) potentialities of observers. Although such an outline provides a general breakdown of these two situations, clearly characteristics of the task, the others present and the context of the activity must be given more detailed attention. There may even be a difference between conditions in which evaluation is certain (anticipated) or merely possible (potential), although these terms are used interchangeably in the literature.

It is also worth noting that, although the effects of mere presence are of theoretical interest, the range of real-life situations which meet this specification is quite restricted. Pure coaction may only be an academic contrivance, and it is questionable whether a group situation in which no cues were exchanged between individuals could even be defined as 'social' (Foot, 1973). There is some logical contradiction in experimental efforts which attempt to create 'pure' non-interactive situations as a means for studying the effects of social stimuli. Unfortunately, although non-verbal cues are generally regarded as an important source of interpersonal communication, their influence has been overlooked in investigations of performance. Even more obvious factors such as the publicness of the task and the opportunities available for performance comparison have been virtually ignored.

In addition, although drive theory postulates an increase in general arousal level, little attention has been given to the effects of such changes on behaviours which are not directly related to performance. Those which are most strongly affected might be irrelevant to the task (may be affiliation or escape oriented (Lazarus, 1966)). If so, then it is a mistake to discuss performance outcomes as though they were directly related to social-environmental manipulations. Reactions to evaluative stimuli are basically social in nature, as are tendencies to uniformity. Hence changes in performance which occur in association with different social situations may be largely the indirect product of changes in social behaviour.

Performance measures are often used as an indirect means of assessing the effects of environmental stressors on internal processes (information processing, arousal). Social situations are perhaps not very much different from other types of environmental stressors, except that an additional mediating process is involved i.e. the subjects' appraisal of the situation (Lazarus, 1967). Usually in experimental psychology the connection between environmental stress and performance is conceived to be:

stress →	change in arousal →	change in
(noise, heat)	level	performance

However, when social stimuli are the stressors the paradigm must incorporate an additional stage:

stress →	appraisal of →	change in →	change in
(other people)	situation	arousal level	performance

Clearly, the effect of the social environment on performance is by no means direct.

Early in this chapter attention was given to usages of the term 'social facilitation', and thus far no statement has been made as to a preferred definition. It is likely that the phenomenon applies more to behaviour in general than to performance per se, and, although changes in arousal may mediate between the perception of environmental stimuli and performance outcomes, a number of other processes are also likely to intervene. In some ways it may be sensible to direct less experimental effort toward proving that social facilitation exists and more toward a general understanding of the effects of social situations on performance, facilitative or otherwise. Concentration on social facilitation has perhaps hampered such understanding, diverting attention from the fundamental differences between tasks, individuals and social situations. Therefore, little advantage would seem to be gained by defining this term until the specific situations leading to facilitation have been better elucidated.

CHAPTER II

EVALUATIVE CUES, INDIVIDUAL DIFFERENCES AND PERFORMANCE

2.0 INTRODUCTION

Considerable attention was given in the previous chapter to the role of evaluative cues in enhancing mere presence effects. Although the rationale for the interaction between these cues and the social environment seems straightforward, 'potential evaluation' is not a clearly defined condition. Judging from the diversity of manipulations used to suggest evaluation, these cues, it would seem, can be extremely subtle. Knowing that potential evaluation can lead to differences in performance in social situations is not particularly useful unless there is some means for discerning what situations are likely to be perceived as evaluative.

Although manipulations of evaluation stress are a relatively recent development in the literature dealing with mere presence, the subject has received detailed attention in other areas of psychological research. For the most part this research has employed instructional sets which are designed either to be neutral or to emphasize competition and/or evaluation. There are several reasons for looking at these findings in detail. Firstly, the social facilitation literature has relied heavily on audience manipulations and little is known about the effects of potential evaluation on coacting groups. Competition between individuals is one possible outcome of evaluation stress in group situations, and this may elicit dynamic social processes which would not be possible in audience situations. Secondly, the relative importance of mere presence and evaluation stress has not been delineated. In the Cottrell et al (1968) study for example, social facilitation resulted when the audience could see the subject performing but not when observation was eliminated. In this case the observant audience 'cued' evaluation and presumably led to higher arousal. However, the actual importance of the audience's presence is unclear since no comparable evaluative condition existed in which subjects performed in isolation. Once the individual had been cued that his competence might be important, the actual physical presence of an observer may have been redundant. Finally, there are a number of studies which look specifically at subtle variations between types of evaluative stimuli e.g. size and type of audience. This literature consists mainly of single, diverse studies. However, it offers some guidelines as to what variables might be important in suggesting evaluation.

The first section of this chapter therefore deals with instructional manipulations and the second with these secondary environmental features. The third section comprises a review of the literature concerning personality factors in relation to performance in stressful situations. Since there appears to be considerable variation between individuals in both the direction and intensity of their reactions to social-evaluative situations, a comprehensive discussion would seem to require at least a brief overview of those individual factors which are known to be related to performance measures and stress manipulations; these being anxiety, extraversion/neuroticism and need for achievement. There are of course many personality dimensions which could be included in this discussion, but those listed above have generated the most research and a summary of the findings can be made with more confidence.

Sarason's (1972, 1975) work on test anxiety is given detailed attention partly because of its direct relevance to the situations of interest, but also because it suggests a different theoretical framework for conceptualizing reactions to evaluation stress. More specifically, Cottrell's (1968) notion of a more or less direct relationship between evaluation stress and arousal is by some accounts theoretically unsound. Zajonc (1965) argued that the mere presence of others was arousing, offering no explanation of the source of arousal apart from an implied innate basis. Cottrell's argument, however, relies on learning factors i.e. that evaluative cues are associated with potential failure and anticipated negative outcomes. This is essentially Sarason's position, although the latter's 'association theory' makes rather different predictions and does not cite arousal as an intervening variable. According to the association theory not only do subjects generalize between experimental settings and past evaluative situations, but the experimental situation acts as a stimulus for specific (coping) responses learned in previously stressful situations.

Farber (1955) discussed the theoretical confusion between drive and associative factors:

> "The associative function of a .... variable is identified in terms of its tendency to elicit or facilitate a limited class of responses only. The drive function .... is demonstrated if a) its presence energizes or intensifies indiscriminately all reaction tendencies existing in a given situation ...." (pp. 311-12)

Sarason is mainly concerned with the effects of anxiety and the tendency for highly anxious subjects to emit task-irrelevant responses which interfere with efficiency. However, considered more generally his analysis of the role of situational cues may be more credible than Cottrell's. If evaluation stress is an associative variable rather than one more directly related to arousal level, it may be expected to elicit markedly different responses between individuals. It is known for example that persons exposed to extreme stress (potential loss of life) exhibit diverse reactions including psychological escape or denial and fatalistic rationalization (Speisman et al, 1964; Friedman et al, 1963; Katz et al, 1970). Therefore it is probably simplistic to think of evaluation stress solely in terms of its potentially energizing properties.

This chapter therefore covers a wide range of experimental material.

Although there are several individual themes which could be dealt with in depth, a wider approach to the subject was preferred in order to illustrate that, although evaluation stress is a commonly employed experimental device, the conditions which elicit it and the effects attributed to it are neither simple nor well understood. The following discussion attempts to review what is known about the more commonly used evaluative manipulations and their effects on performance.

2.1 INSTRUCTIONAL MANIPULATIONS

2.1.1 Some Comments about Instructional Manipulations

It has been fairly well established that subjects' reactions to various experimental situations can be modified by the initial information they are given, creating a predisposition or 'set' to perceive subsequent events in a particular way (Speisman et al, 1964; Schachter and Singer, 1962). In addition, it is likely that reactions to performance situations will differ depending on how the task is presented, i.e. as a test of ability or a normative study. For example, Mandler (1972) suggests that, in terms of information processing, instructions are one means by which the individual selects programs and determines what strategy to apply to any given task and situation: there is some direct evidence from social-performance situations supporting this notion (Good, 1973).

If the task is presented without any explanation of its meaning or the purpose for which it is being given, subjects are likely to interpret the situation in the light of their previous experience in related settings. This tendency has been labelled 'cognitive appraisal' (Lazarus, 1966) or more generally 'perceptual style' (Heider, 1958), and a similar idea labelled the 'social construction of reality' can be found in the sociological literature (Ball, 1972). The effects of this idiosyncratic set are likely to be a major source of variation in many experimental designs (Lazarus et al, 1952). Of course it is not possible to stop subjects from trying to interpret the situation they are in; for most individuals an experiment is novel and probably somewhat stressful (Spence and Spence, 1966; Paivio and Lambert, 1959). However, the ambiguity of the performance situation can possibly be reduced by giving specific instructions as to the meaning and purpose of the task.

The effectiveness of instructions in establishing a set is based on certain assumptions about the relationship between the experimenter and the subject. Presumably, the former is in the position of an expert or authority figure while the latter is there to do a job or favour for the experimenter. In any event, the experimenter is in the controlling position and the subject adopts an assigned role. It is also assumed that the experimenter is able to evaluate the subjects' performance, and therefore that subjects will listen to the instructions and take them seriously. There are of course problems with these assumptions which will be discussed in Chapter III.

There is at present a substantial literature dealing with the effects of different instructional sets on performance. Traditionally these studies have fallen into two rather gross categories; (1) experiments primarily concerned with the effects of competition and (2) those interested in the effects of ego-threat. The latter manipulation may involve instructions which emphasize potential failure or which give subjects direct feedback indicating inferior performance. Usually instructions in the ego-threat category stress the relationship between the task and some valued ability and highlight competence e.g. 'this is a test of intelligence, and we want to see how favourably you compare with others from your class'. However, competitive instructions focus on winning rather than failing and often the activity is given the flavour of a game or contest e.g. 'see if you can do better than the others'.

Although the difference in orientation between these two sets is subtle, in some respects it is analogous to the distinction between need for achievement (positive motivation) and fear of failure (negative motivation) (Atkinson and Litwin, 1960). Persons who are highly achievement motivated are believed to be characterized by behaviour which is oriented toward attaining high performance levels. Alternatively, individuals who report high fear of failure display behaviour which is directed toward avoiding failure; the latter need not take the form of trying to attain high standards but may be expressed in defensive or escape responses which are irrelevant to the task. Whether or not instructions are actually powerful enough to establish either a general orientation toward achievement or, alternatively, failure avoidance has never been directly examined. However, some differences might be expected in the way subjects approach and carry out a given task under these different orientations.

Besides emphasizing different performance outcomes, studies utilizing competitive instructions almost of necessity involve the physical presence of other competitors, and typically the instructions point out the fact that others are present. This is not a feature of experiments using ego-threat, or rather, the social environment is seldom considered. In these latter studies details of the testing procedure are often omitted and when they are included it is fairly clear that the choice of testing conditions has been largely determined by features of the task and administrative convenience. For these reasons the two types of instructions are being discussed separately in the following review. Obviously there is a fair amount of overlap in the work on instructional manipulations and that concerned with social facilitation. Unfortunately, since neither areas of investigation have systematically varied both the social environment and evaluative instructions it is difficult to combine the two literatures.

Studies using failure feedback have not been included in this discussion. Although failure manipulations are considered to be a more extreme variety of ego-threatening instruction (Costello, 1964), they are essentially different. Feedback which indicates inferior performance is not 'potential evaluation', it is unequivocal failure. Furthermore, knowledge of results (whether valid or not) can have informational properties which may affect performance by leading subjects to change their strategy for performing the task (Leamon, 1974). For example, a failure report may result in performance decrement because the subject has abandoned a successful strategy for a less efficient one (independently of any changes in arousal or motivation).

In the following sections the term 'instructional manipulation' is sometimes used interchangeably with stress. Spence and Spence (1966) have discussed the dynamics of instructional manipulations in the context of experimental situations, placing emphasis on their potentially stressful features. In addition, McGrath (1970d) includes this device among a variety of environmental stressors under the general heading of 'experimentally induced stress'. Presumably the social situations described in Chapter I could be viewed in the same way.

A likely goal of the present discussion would be to identify differences (or similarities) between situations and the effects of differing combinations of situational variables on performance. The suggestion has been made that stress manipulations may be additive in their effects (Weiss and Miller, 1971). However, the problem is complex and little literature exists on which predictions can be made concerning these combined variables. For example, ego-threatening instructions might be expected to facilitate performance on a given task for most subjects but may result in decrement when combined with observation by the experimenter. Studies which have attempted to increase stress along one dimension such as the number of people present, have indicated an additive effect (McCullagh and Landers, 1976; Hillery and Fugita, 1975). However, other reports imply that combinations of different stressors may be interactive rather than additive in terms of performance outcomes (Wilkinson, 1969; Good, 1973; Berkey and Hoppe, 1972).

2.1.2 Competitive Situations and Performance

Most studies employing competitive instructions have been confined to tasks measuring gross motor activities, possibly because competition has traditionally been associated with athletics and tests of physical prowess.

The earliest experimental investigation in the area was that of Triplett (1897) who found that on a simple motor task most subjects (children) improved performance under competition but about half either suffered a decrement or showed no change. In subsequent studies employing simple motor or reaction time tasks facilitation of overall performance has usually been noted (Whittemore, 1924; Sorokin et al, 1930; Wilmore, 1968; Church, 1962; Wankel, 1972; Carment, 1970; McManis, 1965). However, in some cases no differences have been found (Whiting and English, 1925; Kozar, 1973).

Similarly, superior performance has usually been found in competitive situations involving simple cognitive tasks such as mathematical operations, word association and symbol substitution

(Hurlock, 1927; Sims, 1928; Dashiell, 1930; Rudow and Hautaluoma, 1975). In one case the learning of a complex relay assembly task was facilitated with no reduction in accuracy (Williams, 1956). However, the experiment employed 'team' rather than 'individual' rivalry. Rudow and Hautaluoma (1975) noted reduced accuracy on two cognitive tasks when competition was individualized.

A few studies have investigated individual differences. High anxiety combined with low need for achievement has been associated with inferior performance on a pegboard task under conditions of competition but with facilitation of performance in a neutral setting (Ryan and Lakie, 1965). Similarly, under competition high anxiety has been related to poorer performance on simple motor steadiness (Vaught and Newman, 1966), and there is some evidence that highly anxious females and subjects of relatively high ability may be most detrimentally affected (Martens et al, 1976; Clifford, 1971).

In spite of the diversity between the instructions used and the procedures followed, most of these experiments report facilitation of performance under competition or in some cases no differences. The only qualification would seem to apply to highly anxious subjects, who have sometimes been detrimentally affected. Group rivalry would seem to be almost as effective as competition between subjects, especially if individual recognition is involved (Hurlock, 1927; Williams, 1956). Furthermore, in only one study was accuracy impaired by the instructional manipulations, although accuracy was not measured in most cases.

The effects of the physical presence of other competitors is equivocal. Wankel (1972) found that instructions alone were sufficient to facilitate simple motor performance and that others' presence made no difference, while Carment (1970) obtained facilitation only when instructions were paired with others' presence. However, there were several procedural differences between these two experiments (particularly in the way feedback was controlled) which may account for the discrepant results.

The experimental literature concerned with competition and performance is inadequate in many ways. The tasks used have in the main provided indices of output for fairly routine operations. Only one learning task is cited here (relay assembly) and this does not involve higher order processes such as reasoning or decision making. Cases in which more complex cognitive activities have been examined have almost always involved comparisons of competitive and cooperative rather than neutral and competitive situations. Usually these have been concerned with discovering the most efficient and least stressful methods for learning in collective settings, particularly in the classroom.

Johnson and Johnson (1975) have reviewed the literature on competition and cooperation and conclude that competition is only associated with facilitation on simple drill or mechanical tasks where speed and quantity of work are important, and even then may only stimulate those who believe they have a reasonable chance of winning. They discuss as myth the belief that competition is necessary for drive and ambition and point out that it is actually associated with decreases in the quality of work and with rigidity in problem solving. This conclusion is virtually identical with that of Allport (1920, 1924) regarding mere presence effects. However, Okun and DiVesta (1975) have shown that the effects of cooperation and competition on performance are complex, depending on group goal structure and reward conditions.

Another area of research on competition has centred on the developmental origins of competitive behaviour and the various functions it may serve for the individual. An overview of this literature suggests that implicit competition is probably a feature of almost all performance situations. It will be recalled that competitive influences were noted as contaminating variables in the control conditions of many studies concerned with social facilitation (Dashiell, 1930; Harlow, 1932; Allport, 1920). This has also been true of experiments directly manipulating competition (Janssens and Nuttin, 1976). For example, Whittemore's (1925) subjects reported feelings of competitiveness during non-competitive trials as well as suspicions that other group members were surreptitiously competing with them. Whiting and English (1925) noted what appeared to be a 'test-taking attitude' in neutral as well as competitive trials and cited this as the probable reason for their failure to find performance differences between the two conditions.

Studies on the developmental origins of competition are important in that they demonstrate its pervasiveness and the different sources from which competition can be expected to arise.

There is evidence that children and adolescents prefer working for themselves rather than for a group (Thompson, 1962; McClintock and Nuttin, 1969) and that competitive responses are well established in children by the time they begin formal education (Leuba, 1933; and Greenberg, 1932). Caillois (1961) observes that long before the emergence of regulated competitions, children often challenge others to endurance contests such as staring at the sun or enduring tickling.

Several different motivational bases have been proposed to account for competitive behaviour (Pepitone, 1967), the most basic of these being the imposition of structural constraints i.e. setting a goal which can only be attained by one or a few group members. However, competition also occurs in situations where no such constraints are evident, and

Pepitone (1967) outlines three likely sources of self-based competitive behaviour; self-evaluation, self-enhancement and self-validation.

The need for self-evaluation (Festinger, 1954), already discussed in Section 1.4, can lead individuals to compare their performance with others. However, the manner in which self-evaluation is related to competition is not quite clear. It is logical that feedback may elicit greater effort on the part of poor performers, but comparative information should provide little incentive for individuals of relatively high ability. Self-enhancement is a rather more straightforward motive i.e. the individual competes in order to raise his self-esteem or status in the eyes of others. Self-validation is similar in some respects to the resolution of cognitive dissonance (Festinger, 1957). Specifically, individuals strive to obtain rewards proportional to their perceived level of proficiency, thereby confirming their self-concept. However, Pepitone (1967) only discusses increases in motivation due to self-confirmation in high ability individuals. Self-validation can also lead to self-perpetuating failure (Gurin and Gurin, 1972). Cratty (1967) suggests that individuals weigh their performance and compare it with norms judged to be representative of their reference group and from this comparison develop a personal level of aspiration. The relationship between this and performance measures can be reasonably direct (Locke, 1966a, 1966b; Locke et al, 1968). Singer (1968) also discusses competition in the forms of competing against established norms, against one's own record or against other individuals.

Considering the various sources from which competitive behaviour can arise, it seems unlikely that any performance situation could be completely free from such influences and that what is being compared in neutral and competitive group situations is really performance under implicit or explicit competition.

Early research assumed that all individuals were stimulated by competitive situations, and when impairment occurred it was due to overstimulation (e.g. Triplett, 1897). However, there is evidence that individuals who are high in self-reported fear of failure tend to avoid competitive situations (Birney et al, 1969). Therefore, the general observation that individuals perform better in contests and competitions may be based on a self-selected group of people who tend to be favourably affected by these conditions. The laboratory situation is a forced competition, in some ways more analogous to an academic test than to a game or contest, and is therefore not identical with spontaneous competition. In the former situation, as subjects are not self-selected, individual differences may contribute more to the variability of performance outcomes. The above discussion suggests some of the sources of variation between individuals in their reactions to competitive stress.

2.1.3 Ego-Threatening Instructions

In addition to the observation that competitive manipulations are usually achievement oriented whereas those involving ego-threat emphasize failure, there are other differences. Although the obvious intention underlying both types of set is motivational, it is not clear why ego-threatening instructions should be very effective, since performance takes place in the absence of others, thereby minimizing the consequences of both success and failure.

It has been argued that the relationship between the subject and the experimenter is important in this respect; that subjects will work hard in order to comply with the experimenter's requests and to avoid his negative evaluation (Orne, 1969). The emphasis often placed on the future comparison of subjects' scores with those of others may also be significant, since others who are not physically present can represent reference groups with whom comparisons are made (Sherif and Sherif, 1953). Even instructions which state 'this is a test of intelligence' imply comparison with others, since intelligence scores are only meaningful in a relative sense. Hence, the subject may work both to fulfil the requirements specified by the experimenter and in order to compare favourably with his own reference group. It would be reasonable therefore to consider ego-threatening instructions as a variety of social stress. The point of interest is whether or not, and in what ways, the effects of ego-threat per se differ from those of potential evaluation in others' presence.

The literature concerned with ego-threatening instructions and performance is vast but largely unsystematic, consisting mostly of single, isolated studies. Only Sarason has tested different combinations of variables in a fairly comprehensive fashion, and therefore his work is given more detailed attention. A more extensive coverage of this literature can be found in Costello (1964).

Although the effects of ego-threat and especially failure stress were popular research topics before 1950, the early literature is too diverse to offer more than very tentative statements concerning these manipulations. This is the conclusion reached by Lazarus et al (1952), and following their review the topic was approached with somewhat more sophistication. Mandler and Sarason (1952), in the context of their research on test anxiety, questioned whether specific instructions were necessary to elicit anxiety responses in predisposed individuals, implying that the experimental situation per se might be sufficiently test-like. Hence more recent studies have tended to employ situations which are superficially similar to testing situations while giving careful attention to the number and type of evaluative cues in the actual task instructions.

Unlike the experimental work on competitive manipulations, research on ego-threat has favoured the use of cognitive rather than motor tasks. This is due largely to the emphasis placed on intelligence which typifies these instructions and is also a by-product of efforts to simulate test-like conditions. Generally, the tasks used fall into three categories; cognitive tasks with little or no learning component, problem solving or creativity tasks and serial or paired-associate learning.

Simple cognitive tasks have either shown facilitation when performed under ego-threatening conditions or have been relatively impervious to this type of stress. For example, digit symbol and card sorting performance have both been enhanced by ego-threatening instructions (Sarason and Palola, 1960; Angelini, 1973), although no differences emerged between easy and difficult versions of either test. Curiously, subjects asked to give word associations under conditions of ego-threat yielded responses of lower commonality (compared to norms) (Sarason, 1959a), an effect exactly opposite to that which would be predicted by arousal theory and interpreted by the author as indicative of interference. In one of the few studies to use a perceptual motor task Fleishman (1958) found that airmen of high ability improved their performance on a rudder control task after receiving ego-threatening instructions while the performance of low ability subjects was unaffected, suggesting that only those who anticipated success were influenced by the instructions.

Experiments utilizing learning tasks have yielded conflicting results. Serial learning under ego-threat has been associated with performance facilitation (Sarason, 1957) or no differences (Sarason, 1956) and impairment has been noted under failure stress (Sarason, 1956).

One experiment noted greater variability but no differences in trials to criterion (Russell, 1952). However, consistent with other findings on immediate and delayed recall (Walker and Tarte, 1963), retention scores in serial learning have been shown to be higher when the original learning took place under conditions of ego-threat (Alper, 1948, Sarason, 1957). Although serial learning has by far been the most widely used task, impairment of performance under ego-threat has been noted for paired-associates learning (Long and Bessemer, 1971) and digit span (Moldawsky and Moldawsky, 1952); the latter more than likely indicates restricted cue utilization rather than learning in the usual sense.

Problem solving tests have tended to be insensitive to manipulations of ego-threat. No differences attributable to the experimental treatments have been found on difficult anagram solution (Sarason, 1961), identifying rules to number series (Sarason and Palola, 1960) or on the Vocabulary subtest of the Wechsler Bellevue Scale (Moldawsky and Moldawsky, 1952). In addition, originality scores on the Torrance Test of Creative Thinking were shown to be unaffected by ego-threat, although fluency scores were lower than those of controls (Belcher, 1975).

One reason for these inconclusive findings is no doubt that implicit evaluation in the neutral situations has reduced the differences between control and stress conditions. The power of these implicit features in cueing evaluation has been illustrated in an interesting experiment by Sarason et al (1972). Prior to a typical serial learning experiment subjects were given pre-test interviews, the content of which centered around the subjects' reactions to either testing situations or university life in general. Following the interview,task-oriented (neutral) or ego-threatening instructions were read. The overall results showed superior performance for those subjects receiving the campus interview, with no additional effects attributable to the post-interview instructions. Sarason argues that simply talking about testing situations may cue anxiety responses which may persist after the interview has ended. Clinical work employing desensitization procedures is of course based on a similar rationale (Wolpe, 1958). It is also true that the campus interview may have had a relaxing effect on the subjects which neutralized the impact of the subsequent ego-threatening instructions.

In spite of the inconsistencies mentioned, some summary of the findings is possible. Firstly, it would seem that fairly routine cognitive tasks are generally facilitated by ego-threatening manipulations, while problem solving tests are relatively insensitive to these treatments. However, the latter have a tendency to generate intrinsic motivation, rendering them impervious to many types of environmental stress (Kahneman, 1973). Learning tasks, on the other hand, have more often than not been associated with impairment under conditions of ego-threat.

Costello (1964) reaches essentially the same conclusion based on his review of the literature, although he is more tentative regarding the results on learning tasks. Some of the difficulty in interpreting the data may lie in the classification of 'routine' and 'learning' tasks. Most experimental tasks are novel and require some degree of learning for efficient execution. A categorization of tasks based on difficulty might be more useful in this context. For example, serial learning is the most widely used task and has produced the most inconsistent findings. Yet this would seem to be an easier task than paired-associate learning because of the advantage of having a constant order of presentation. Tasks which are neither clearly difficult nor simple might be more likely to show either disruption or facilitation depending on subtle variations in experimental procedures.

Considered in total, the results from experiments manipulating either the presence of others, competition or ego-threat have been remarkably similar: all of these conditions would seem to be associated with the facilitation of easy tasks and the impairment of difficult, complex tasks, particularly those involving the learning of new material. This raises again the question of the relative strengths of these differing social variables. Cottrell (1968, 1972) argues that potential evaluation is a necessary condition for the occurrence of social facilitation; studies utilizing ego-threatening instructions suggest this is often a sufficient social stimulus.

However, the range of evaluative cues is potentially much greater when performance takes place in the presence of others. Both the personal characteristics of the other individuals as well as specific information available from them may either reinforce or neutralize the idea of evaluation suggested by the instructions or the testing situation. A performance situation would seemingly always be one of 'potential' evaluation. Whether or not that potential becomes translated into actual 'anticipated' evaluation would appear to depend on other more subtle features of the testing environment. Some of these are discussed in the following section.

2.2 SECONDARY ENVIRONMENTAL FEATURES AND POTENTIAL EVALUATION

If a social situation is judged to be evaluative, two categories of environmental factors would seem to be of central importance; (1) characteristics of the others present and (2) features of the physical setting which suggest or enhance evaluation. Variables such as the number of others present, the context of the subject's activity and information given about the others may or may not be relevant to the subject's assessment of and performance in any given situation.

2.2.1 Characteristics of Others

Although others in a performance situation can be viewed as sources of potential evaluation, it is unlikely that subjects will attribute evaluator status to just any individual who is physically present. Webster and Sobieszcek (1974) take the view that only 'significant others' are regarded as sources for evaluation and are thereby given the power to influence subjects' decisions and judgments. A significant other must be more capable than the subject is of evaluating performance, either because he is believed to possess greater ability or has access to objective standards. If the other's ability is unknown, as is often the case, information about status may be equated with actual ability information i.e. high status evaluators are believed to have high ability. This implies that, at least in the typical experimental situation, lecturers, peers and younger students are probably progressively less salient as evaluators.

Supporting evidence comes from many areas of the psychological literature. Two experiments indicate that subjects find being observed by friends or classmates more stressful than observation by strangers or younger people (Brown and Garland, 1971; Levin and Baldwin, 1958), and latencies on hidden words problems have been shown to be greater when subjects performed before an audience of postgraduates and lecturers as opposed to other undergraduates (Cohen and Davis, 1973). Even testing by an 'attractive' rather than an 'unattractive' female experimenter produced relative impairment on both a verbal and a motor skill learning task (Hartnett et al, 1976).

When audience expertise has varied, pursuit rotor learning has been poorer with both a non-expert and expert audience (as opposed to performance alone), but only the subjects in the expert audience condition failed to improve relative to the no audience condition in later trials (Lombardo and Catalano, 1975). Furthermore, the recall of an array of objects has been shown to be selective when subjects are given information about the sex or attitudes of an anticipated audience (Schramm and Danielson, 1958; Simmerman and Bauer, 1956). In an experiment by Winkel and Sarason (1964), serial learning was found to be more sensitive to the test anxiety of the experimenter than to the experimental instructions. In this instance subjects showed no performance differences when instructions were either ego-threatening, neutral or reassuring, but those who were tested by an experimenter of low test anxiety showed superior performance under ego-threat. It would seem that experimenters who are themselves sensitive to testing situations might transmit subtle cues to subjects which effectively reduce the differences between conditions.

2.2.2 Evaluative Cues in the Physical Environment

Although the relationship between the status or personal characteristics of others and evaluation potential seems fairly straightforward, the effects of the physical environment are less clear.

Several studies suggest that an invisible audience (one-way mirror or TV camera) can be at least as potent an arousal device as the physical presence of others (Geen, 1973) or even more disruptive (Laughlin and Wong-McCarthy, 1975; Wapner and Alper, 1952; Criddle, 1971; Ganzer, 1968). Presumably being observed by people who are unobservable in return and about which subjects can not estimate status or number is more unsettling than a live audience which is available for unambiguous assessment.

Studies examining the effects of audience size on performance have produced equivocal results. Levin et al (1960) found that children told shorter stories to an audience of six adults than to only the experimenter, whereas students elaborated and enlarged when retelling a story to an audience of 10-12 (Hanawalt and Ruttinger, 1944). However, these results are compatible if the tasks are viewed respectively as one of creativity and one of recall. Presumably, all things being equal, a large audience would have more evaluation potential than a small one, and some studies have shown small but linear relationships between performance and audience size (McCullagh and Landers, 1976; Hillery and Fugita, 1975). However, McCullagh and Landers (1976) found that, although self-reported activation increased in proportion to audience size, performance was unaffected. Gates (1924), using several tasks, failed to obtain any differences between performance before only the experimenter, a small or a large audience, and concept attainment has demonstrated no sensitivity to observation by either one or two people (Laughlin and Wong-McCarthy, 1975). Generally it would seem that, although individuals may report different preferences or feelings of anxiety regarding audience size, these are not strongly related to performance outcomes.

However, the actual size of the audience may be less important than its ability to view and evaluate the subjects' performance. Subjects' ratings of 'feeling observed' have been shown to be positively related to the amount of time a confederate spent looking at the subject (Argyle and Williams, 1969). Geen (1973) found a similar trend when subjects were told they were being observed by the experimenter rather than just having him in the room. Performance has been affected in a manner consistent with arousal theory when observers could see the subjects' performance rather than being present but blindfolded (Cottrell et al, 1968). Obviously the task often dictates the degree to which a performance situation is public. Paper and pencil tests, where no evidence of even output is apparent, would to some extent protect the individual from evaluation by others: even the experimenter, who has access to norms, can only give an evaluation after the subject has left the situation.

Another seemingly important, although virtually unexplored situational variable, is the number of individuals in a coacting group and their proximity to each other. The available literature suggests little difference in performance between subjects tested either alone or in pairs, while groups of 3 or 4 have led to greater muscular endurance (Martens and Landers, 1969), poorer motor coordination (Martens and Landers, 1972) and poorer acquisition of a motor skill (Burwitz and Newell, 1972). However, there is some evidence that performing in pairs rather than alone may reduce the impact of observation by an audience (Laughlin and Jaccard, 1975), suggesting the dearousing effect of companions in stressful situations discussed by Epley (1974).

In terms of proximity, Emiley (1975) found no direct relationship between density and task performance. Similarly, using several different tasks, Freedman et al (1971) noted no effects on performance with varying densities and group size (5 or 9) and with two different populations (housewives and students). However, Chapman (1975) found that children seated close together engaged in more eye contact and laughed more in response to humorous stimuli than those who were seated farther apart. There is also some evidence of sex differences, as males have shown increases in competitive type behaviour when in confined spaces with other males while females have shown the opposite tendency in single sex groups (Freedman et al, 1972).

The relationship between density and task performance is apparently complex. Competitive instructions have produced superior performance in small groups (2) and when individuals were seated far apart, while larger groups (4) seated close together performed better under a cooperative set (Seta et al, 1976). Curiously, evaluative instructions have also been associated with superior performance when half of the group members have been working on different tasks (Seta et al, 1977), perhaps suggesting a tendency toward group uniformity in the single task group. A rather more predictable finding by Schachter et al (1951) is that group members who are led to believe they are individually compatible (based on personality test scores) display greater productivity and respond more favourably to directions from other group members than individuals who believe they are incompatible, although this was an examination of perceived similarity rather than physical proximity.

Obviously, it would not be possible to assess the effects and interactions of all these variables within any one experiment. However, since these physical dimensions are probable sources of error variance, it would seem important to hold as many of them constant as possible when assessing the effects of any one or two discrete environmental features. This has not been the case in much of the experimental work on performance in social situations, since either these variables have not been controlled for in the experimental procedures or studies have been one-off investigations which, because of situational variations, do not enable direct comparison with other studies.

2.3 INDIVIDUAL DIFFERENCES IN RESPONSE TO SOCIAL/EVALUATIVE SITUATIONS

2.3.1 The Problem of Individual Difference

Lazarus et al (1952) in their early review of the literature on psychological stress and performance criticized the theories of that period for their failure to account for individual differences. Although these factors are generally regarded to be a major source of variation in the whole area of social stress research (Eysenck, 1966), recent work has offered few additional insights.

Individual differences of course cover a variety of attributes ranging from ability, race and age to differences in social class and personality. Physiologically based differences, such as age and race, are thought to be fairly stable in their effects on performance. However, personality variables are often defined on the basis of the individual's reactions to social situations, and logically these would be expected to interact with situational variables. Argyle (1972) has argued that performance on working sites is determined largely by personal traits of the individual in social situations.

Because of the difficulties intrinsic to the definition, measurement and control of personality differences, many researchers have tried to avoid the issue altogether by employing large numbers of subjects so that differences between individuals will be randomly distributed over all conditions. This is in some ways a justifiable methodology since efforts to isolate relationships between performance and personality measures have for the most part been neither direct nor consistent (Lazarus et al, 1952; Davis, 1969; Mischell, 1973). Davis (1969) argues that such attempts are naive; that the number of factors (both internal and external) which impinge on performance in any given situation would understandably mask the effects due to any single personality feature. In addition, the problem of investigating these factors in evaluative situations seems particularly difficult. Apparently, although individuals <u>perceive</u> ego-threatening situations in a fairly uniform manner, <u>reactions</u> to this type of stress are more variable between individuals than is the case with other perceived stressors such as anticipated pain or punishment (Magnusson and Ekehammer, 1975).

However, Eysenck (1966) has criticized experimental psychology in particular for failing to give proper attention to individual differences, especially in the cause of increased experimental rigour:

> "No physicist would dream of assessing the electric conductivity, or magnetic properties, or the heat resisting qualities of random samples of matter or 'stuff-in-general', ... Some conduct electricity, others do not or do so poorly; we do not throw all these differences into some gigantic error term and deal only with the average of all substances." (p.2)

More recently, Underwood (1975) has pointed out the need for bringing individual differences into the mainstream of theory construction, and particular emphasis has been placed on identifying the role of these factors in social facilitation research (Underwood, 1976). Roessler (1972), after elaborating on the imprecision of experiments manipulating personality factors, ends by defending such investigations on the grounds that "the consistent results that have been obtained are likely to be valid simply because there are so many potential sources of error variance." (p.316).

The major dangers of ignoring individual differences would appear to be that (1) significant effects apply only to mean differences and cannot be applied to individual cases, (2) insignificant results may be obtained because different types of individuals react in opposite directions to the experimental manipulations and (3) theories may be constructed from data based on very selected samples and have little generality to other groups of people (Eysenck, 1966).

The following discussion is centred on those measures of personality which have generated the most systematic research, these being anxiety, extraversion/neuroticism and need for achievement. General findings will be discussed and similarities between theories and findings will be pointed out when possible. The literature concerning anxiety and performance would seem to be particularly relevant in this discussion since many of these experiments have also varied instructions.

2.3.2 Anxiety and Performance

Two theories concerning the relationship between anxiety level and performance have dominated the literature since 1952; the drive theory (Taylor and Spence, 1952) and the association theory (Mandler and Sarason, 1952). The two have an essentially different thrust in that the drive theory predicts differences in performance based on stable, endogenous differences between subjects (which operate in all situations), whereas the association theory makes predictions based on learned anxiety responses which are evoked only in specific situations. This constitutes the basic difference between trait and state anxiety, an issue which has consumed considerable attention in the anxiety literature (Alpert and Haber, 1960; Spielberger, 1966; Endler, 1975).

The drive theory assumes that 'emotionality' (as measured by the Manifest Anxiety Scale, MAS) has drive properties. The predictions arising are the same as those discussed in Chapter I concerning arousal level and the emission of dominant responses. The association theory maintains that performance situations act as stimuli which elicit either task relevant or task irrelevant (anxiety) responses, depending on the individual's reinforcement history in performance situations. Stated simply, when task responses are evoked, performance will be facilitated whereas anxiety responses are characterised by interfering thoughts about failure or inadequate performance and result in performance impairment (Sarason, 1975).

Investigations involving manipulations of MA have ranged over all sorts of experimental tasks and real-life situations. Initial research generally supported the prediction that high drive (MA) would be associated with superior performance on simple learning tasks e.g. conditioning or those low in response competition, but would impair learning on difficult tasks or ones where response competition was high (Farber and Spence, 1953; Spence et al, 1956; Taylor and Spence, 1952). However, later investigations by and large failed to replicate these findings. Costello (1964) and Feldman (1964a) conclude that little unequivocal evidence exists in support of the theory and earlier Spence (1958) pointed out that, at least in complex learning situations, endogenous drive appeared to be a relatively unimportant determinant of performance.

However, all of the early experiments employed learning tasks (serial, stylus maze and paired associates) for which subjects' responses were public (oral) and recorded by the experimenter. In some studies (stylus maze) subjects were told by the experimenter whether the response was correct or not at each choice point. Considering the research discussed in Section 2.2, it seems likely that these situations were suggestive of evaluation and not a pure test of individual differences in basic drive as was intended. In fact the situations appear roughly analogous to those more recently employed in investigations of potential evaluation and social facilitation. In a later review of their work Spence and Spence (1966) reach the conclusion that situational variables are more effective determinants of drive than is manifest anxiety, although MA appears to predispose individuals to situational stress (Sarason, 1960; Saltz, 1970).

The association theory has undergone several reformulations in light of experimental findings, most of these incorporating features of other theories, e.g. an inverted U-shaped relationship between anxiety level and performance. The Test Anxiety Scale or Questionnaire (TAS and TAQ) have been the most widely used instruments, although the MAS is positively correlated with test anxiety, as are most measures of fear of failure. In fact, test anxiety has come to be considered synonomous with fear of failure and the TAQ has been employed in research on Atkinson's dual achievement motivation model (Atkinson and Litwin, 1960) and Birney et al's (1969) work on Hostile Press (a projective technique yielding a factor similar to fear of failure).

Obviously, there is considerable overlap amongst all these theories and the measures developed to test them. Therefore, in the following review more emphasis is given to the particular tasks used rather than the anxiety measure, since it seems likely that the former will be a more important source of error variance. Particular emphasis is placed on Sarason's work since it is the most systematic and comprehensive in the area; nearly all these experiments have employed the TAS.

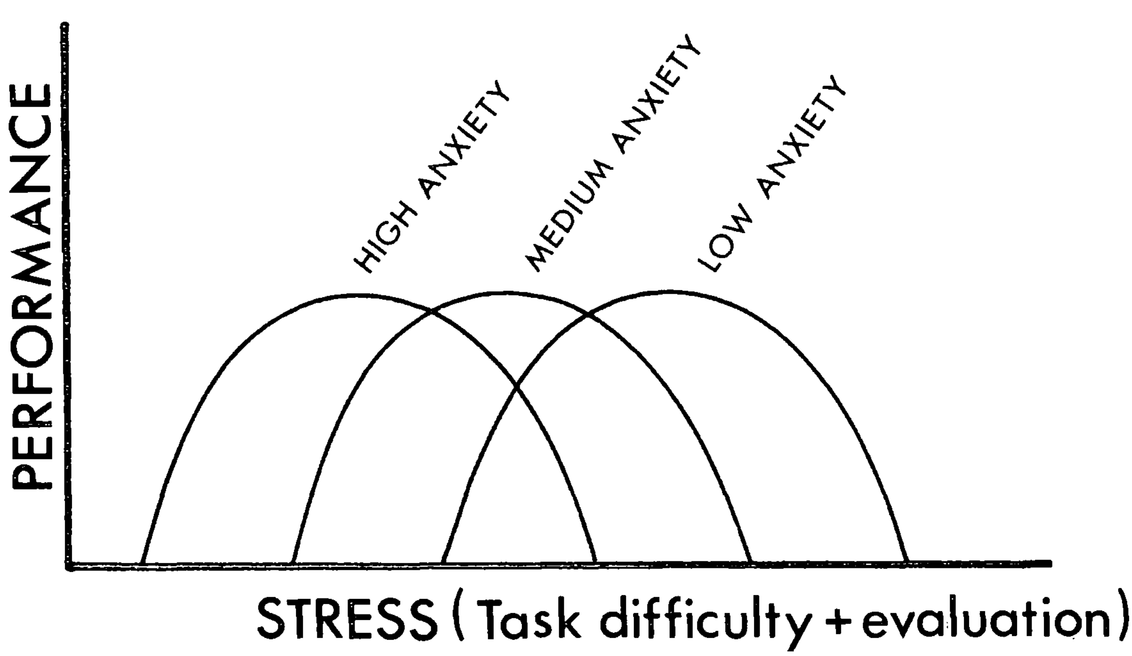
In general, the experimental evidence suggests that for both routine cognitive and learning tasks, highly anxious subjects are negatively affected by ego-threatening situations, while subjects low in anxiety do relatively well under these conditions but show inferior performance in neutral or relaxed situations (Mandler and Sarason, 1952; Sarason, 1957; Sarason, 1956; Glover and Cravens, 1974). It also appears that highly anxious subjects are more likely to show decrement on difficult tasks when under stress while low anxiety is associated with performance impairment of easy tasks in neutral conditions (Sarason and Palola, 1960; Sarason et al, 1952).

It is noteworthy that performance on tasks requiring creativity or problem solving, which have been unaffected by evaluation stress (Section 2.1.2), do show effects due to ego-threat when subjects' anxiety level is taken into account. For both anagram solutions and identifying number series rules highly anxious individuals were negatively affected by ego-threat while low anxiety subjects showed decrement under neutral conditions (Sarason, 1961; Sarason and Palola, 1960). In another study (Dunn, 1968), low anxiety was associated with superior performance regardless of the instructions, although the task used was a subtest of the WAIS and was probably recognized as an IQ test by most subjects.

The interaction of task difficulty and situational stress is interesting in that it suggests a complex curvilinear relationship between anxiety level and performance. The Yerkes-Dodson Law (Yerkes and Dobson, 1908) predicts an interaction between task difficulty and arousal level, although with anxiety level taken into account, a relationship should emerge resembling that diagrammatically represented in Figure 2.1.

FIGURE 2.1

INTERACTION BETWEEN ANXIETY LEVEL AND SITUATIONAL STRESS



Stress (or arousal) in this case is conceived of as the combination of situational and task variables. Although the association theory is based on differential responding to environmental cues between subjects rather than arousal level per se, the relationship depicted could be explained on the basis of an increasing likelihood for evaluation sensitive individuals to respond with task-irrelevant anxiety responses under progressively more stressful conditions.

It is thought that anxiety responses of this nature take the form of misdirected attention, leading the individual to either focus on himself and his potential failure or to search the environment for cues regarding appropriate behaviour (Wine, 1971; Sarason, 1975, 1972); both may be expected to interfere with performance. A series of studies has shown that subjects who score high on 'worry' items (thoughts about failure) of the TAS rather than 'nervousness' (physical feelings of arousal) are debilitated under evaluation stress (Morris et al, 1975; Doctor and Altman, 1969). Furthermore, highly test anxious children have been noted to engage in more off-task behaviour, i.e. glancing around the room or at the experimenter, and this same group of subjects demonstrated relatively poor performance on an anagram solution task (Nottelmann and Hill, 1977). Both associative and drive factors may of course be operational in these situations, although the role of attentional factors has been more directly assessed and the evidence obtained appears more conclusive.

Studies from real-life settings have consistently demonstrated negative relationships between test anxiety and performance on intelligence tests (Brown, 1974; Sarason, 1959b; Zweibelson, 1956). However, Birney et al (1969), in an extensive review of the fear of failure literature cite several examples in which evaluation sensitivity has been positively related to actual academic performance: other reports have noted similar findings using general anxiety measures (Dubey, 1976; Spielberger, 1975). Also, there is some experimental evidence that the negative effects of test anxiety can be overcome if subjects are given forewarning of the test (Sarason and Ganzer, 1970). At least for highly test anxious subjects forewarning may lead to overpreparation and subsequent superior performance. If this is the case then it may only be possible to generalize the findings from the laboratory studies discussed above to situations in which short notice of the test is given or cases where preparation, by nature of the task, is impossible.

There are only a few experiments which have studied the effects of anxiety in response to differing social situations of testing. High manifest anxiety and stress (time pressure) have been found to be more detrimental to performance on anagram solution when subjects work in pairs rather than alone (Kanekar et al, 1975). Similarly, observation by the experimenter has been associated with restricted cue utilization for high relative to low test anxious subjects (Geen, 1976). Low self-esteem, SE, (thought to be related to fear of failure) has also been shown to be sensitive to audience presence. In one case subjects low in SE became more conservative in risk-taking when confronted with an audience while those high in SE were unaffected (Cohen and Sheposh, 1977). Another study involving concept attainment demonstrated that individuals who reported low confidence in their ability to do the task made relatively more errors when performing before an audience rather than alone (Shrauger, 1972). These results are in keeping with findings from studies employing more explicit manipulations of evaluation stress. It is interesting that only one of these reports obtained a main effect due to the presence of others (Kanekar et al, 1975) and this was probably due to the fact that paired subjects worked in cooperation, thereby minimizing errors.

Although some general trends are apparent in these studies, the relationship between anxiety, the social environment and performance is not well understood. At least one experiment found that high, rather than low anxiety (MA), facilitated the learning of a coincident timing task in an audience condition (Martens, 1969a), although both anxiety groups learned more efficiently when alone. Another report showed birth-order to be more predictive than test anxiety of poor performance before an audience (Quarter and Marcus, 1971). Pederson (1970) found that neither anxiety nor social conditions affected performance on a learning task or simple multiplication, while coaction facilitated cancellation performance. However, most studies suggest that anxiety, particularly fear of failure, may merit systematic attention in future experimentation in the area of social situations and performance.

2.3.3 Extraversion, Neuroticism and Performance

Eysenck's theory of extraversion/neuroticism bears some similarity to the Spence drive theory, at least in its implications for performance. Basically the model proposes that introverts and extraverts differ in their intrinsic level of arousal, introverts being the more highly aroused (Eysenck, 1957). In addition, it is hypothesized that introverts build up cortical inhibition more slowly than extraverts, resulting in better performance over time on tasks requiring sustained attention.

That extraverts are characterised by initially lower levels of arousal than are introverts is supported by data from the performance of simple experimental tests such as spiral after-effects, vigilance and unpaced reaction time, as well as simple cognitive tasks such as symbol substitution and cue utilization. However, with more stressful tasks (the Stroop test and paced serial reaction time) an inverted U-shaped relationship has been noted (Claridge, 1967): stress in terms of situational manipulations (incentives) has produced a similar survilinear trend (Corcoran, 1965). More recently verbal learning has shown basically the same pattern predicted by drive theory, i.e. extraversion has been associated with more efficient learning of competitional material, and an inverted U-shaped relationship has been obtained for performance ranging along the dimension from stable-extraversion to unstable-introversion (Allsopp and Eysenck, 1974).

Eysenck's theory is too complex to be treated adequately in a brief discussion, particularly the data concerning the build up and dissipation of inhibition; some of the more controversial issues are discussed by Gray (1971). However, it is fairly well established that extraversion relative to introversion is associated with performance decrement on monotonous, repetitive tasks (Davies and Hockey, 1966; Thackray et al, 1974). This pattern of performance is less obvious with more difficult tasks such as pursuit rotor learning (Feldman, 1964) for which there is a complex interaction between learning rates, practice schedules, and the duration and frequency of rest periods.

Neuroticism, although originally thought to be orthogonal to extraversion, is more recently considered to also have drive properties (Eysenck, 1967). In a 1972 paper concerned with personality and academic achievement, Eysenck discusses this dimension as if it were synonomous with anxiety, and in fact the correlation between N scores and anxiety measures has been shown to be as high as .88 (Roessler, 1972).

Of particular interest to the present discussion is the proposed positive relationship between introversion/neuroticism and academic performance (Lynn, 1971), since an academic test clearly involves ego-threat. The argument is that introverts, being less sociable, will apply themselves to their studies with more efficiency and will have an actual learning advantage because of their relatively slow rate of inhibition accumulation over prolonged study periods. It is further proposed that individuals suffering from debilitating neuroticism will fail to reach university level, but that those who are not impaired will benefit from their neuroticism due to its drive properties. There is a clear parallel between this view and that concerning anxiety and overpreparation (Spielberger, 1975).

Eysenck (1972), considers the relationship between introversion and achievement to be well substantiated. More specifically, positive correlations have been found between introversion and persistence (Lynn and Gordon, 1961; Gupta, 1973), neuroticism and vocabulary and moderate neuroticism and Raven's Matrices performance (Lynn and Gordon, 1961). The results concerning neuroticism are more ambiguous than those for extraversion, although it has been argued that N scores of about one half a standard deviation above the mean have beneficial effects on academic performance (Eysenck, 1972; Lynn and Gordon, 1961). However, in an examination of performance over a 5 year period, Kline and Gale (1971) found no significant correlations between these measures of personality and academic achievement. They also point out that those obtained in other investigations have been too low to be useful for either selection or prediction.

It is rather surprising that a measure as well known and widely used as the EPI has not received more attention in investigations of performance in social situations. For example, there is considerable literature dealing with the relationship between physical stressors or distraction and extraversion and/or neuroticism. The data indicate that the performance of extraverts is facilitated by distraction with introverts showing the opposite pattern (Morgenstern et al, 1974). Apparently the former (because of their lower intrinsic level of arousal) seek stimulation (Davies et al, 1969; Eysenck, 1966; Hill, 1975) and are therefore favourably affected by it, while the latter are debilitated by superoptimal arousal. However, Gibson and Curran (1974) found that the two types of individuals employed different strategies which resulted in equal efficiency of overall performance, extraverts speeding up performance during distraction trials and introverts slowing down. Apart from actual performance differences, it has been noted that high N scorers are more apt to report annoyance over environmental stressors such as aircraft noise (Broadbent, 1972), and generally, introverts have a lower tolerance for aversive stimuli (Gray, 1971).

The effects of social situations on the performance of these temperamentally different types of people seems a logical extension of the above experimental work, especially considering Zajonc's (1965) proposal that others are a source of arousal, which, if true should result in an interaction with both extraversion and neuroticism. Colquhoun and Corcoran (1964) have reported a positive relationship between introversion and output on a cancellation task when subjects were tested in the morning and in isolation, although no similar tendency was found in afternoon sessions. Such results suggest that these personal variables may be sensitive to fairly subtle variations in the social environment.

2.3.4 Achievement Motivation and Performance

Murray (1938) regarded the need for achievement (nAch) as a basic personality characteristic, and nAch has received considerable experimental attention since McClelland et al's (1953) investigations, which employed projective test techniques to discern achievement motivation. The literature on nAch differs from that dealing with anxiety and extraversion/neuroticism in that interest lies mainly on behaviour in performance situations rather than performance per se. The topic receiving most attention has been risk-taking (level of aspiration or goal setting) and the effects of task difficulty and/or probability of success on risk-taking. There is however a substantial literature which suggests the presence of a low positive correlation between nAch and academic performance (Smith, 1964): the relationship being strongest for school aged subjects (Brown, 1974).

One extrapolation from McClelland's theory proposes a parallel dimension to nAch, the motive to avoid failure (Atkinson and Litwin, 1960). For nAch persons the motive to succeed is stronger than the motive to avoid failure (the latter is usually measured by the TAS), while the opposite is true of persons with high fear of failure. The outstanding finding from this literature is that nAch individuals have a preference for moderate risks i.e. set realistic goals based on past performance and the expectancy of success (Heckhausen, 1972; Weiner, 1970; DeCharms and Dave, 1965; Teevan and Smith, 1975), whereas failure avoidant subjects typically set goals which are either unrealistically high or low - a defensive gesture which effectively eliminates any confrontation with a realistic test of ability (Birney et al, 1969).

However, although these two behavioural patterns have been extensively investigated, their effects on actual performance are unclear. Intuitively a positive relationship would be expected between the goals an individual sets and his attained performance level. It has been demonstrated that performance is sensitive to goals set by both the experimenter (Locke, 1966a) and the subject (Locke et al, 1968), although there is little understanding of the actual processes involved. For example, a moderately high goal may affect speed of execution but fail to influence cue utilization or ability to discriminate. Also, unrealistic goal setting in the case of failure avoidant subjects would seemingly confound any relationship between goal setting and attainment since such behaviour reflects neither the subject's ability nor the attractiveness of the goal.

That these divergent orientations do exist however finds additional support from experiments studying the Ziegarnik effect, as subjects high in nAch consistently recall uncompleted tasks whereas failure avoidant individuals recall completed ones (Heckhausen, 1972; Gellerman, 1963; Weiner, 1970). Smith (1964) argues that nAch subjects display greater persistence in reaching their goals; these studies in the Ziegarnik situations imply that unfinished or inadequately done tasks have motivational value for nAch subjects which may be expressed in sustained effort over time.

The social situation is not usually considered in investigations of achievement motivation. However, in terms of instructions, nAch subjects seem to respond favourably only if the task is seen as either demanding or appealing to personal pride (Heckhausen, 1972; Gellerman, 1963; Weiner, 1970); when extrinsic rewards are offered differences between motive groups disappear (Atkinson and Reitman, 1956). Similarly, high achievement motivation is associated with better performance if the task is presented as moderately difficult or if reports of failure are given, whereas failure avoidant individuals improve after success reports (Karabenick and Youssef, 1968; Feather, 1965; Weiner, 1970; Karabenick and Marshall, 1974). In terms of performance in group situations, there is some evidence that nAch subjects perform better than those who are failure avoidant under competitive conditions (Ryan and Lakie, 1965; Weitzenkorn, 1974).

A recent and controversial development in the area of achievement motivation is Horner's (1974) work on the motive to avoid success, thought to be confined mainly to females. Fear of success is believed to arise due to conflict over the rejection of the traditional female role in achievement situations. Presumably males and females are differentially reinforced for success in such settings. Teevan and Smith (1975) have offered this explanation for their failure to obtain the usual relationship for females between goal setting and fear of failure. Other experiments using projective techniques have corroborated Horner's findings (Brown et al, 1974).

It is not uncommon to find that females yield poorer performance when competing against males (Allen and Boivin, 1976) and there are reports in which females who were judged to be high in fear of success (based on projective measures) have shown inferior performance in situations which could be construed as competitive or unfeminine (Karabenick et al, 1976; Makosky, 1976). However, other experiments report no sex differences under these conditions (Karabenick, 1972; Stake, 1976) and Sorrentino and Short (1974) found that women high in fear of success actually performed best on a male oriented task, suggesting that the measure needs refinement and perhaps more attention to the particular population under study (housewives, school girls, university women). Tresemer (1976), in an extensive review, concludes that fear of success is neither very widespread amongst females nor related substantially to performance. However, a better understanding of the differences between subject populations, the measures used in assessing fear of success and the definition of 'success' itself is needed before a conclusive evaluation of Horner's theory can be made.

2.3.5 Comments on Personality Measures and Performance

Two things are striking about the literature in this brief review. Firstly, the measures used to describe essentially different characteristics are quite often substantially related. Correlations between manifest anxiety and test anxiety range from .32 to .53 (Alpert and Haber, 1960;

Raphelson, 1957) and values up to -.43 have been obtained between nAch and test anxiety (Smith, 1964; Raphelson, 1957). In addition, neuroticism is positively related to most measures of anxiety (Roessler, 1972; Eysenck, 1972).

Secondly, although these measures are sometimes highly correlated, the experimental literature surrounding them is concerned with quite different phenomena. Text anxiety studies typically look at the effects of situational stress on performance, those using extraversion/neuroticism focus on differences in basic processes such as arousal and inhibition between types of people and achievement motivation studies elaborate on behavioural patterns in performance situations.

Some attempts at synthesis have been made. For example, Birney et al (1969) have considered the literature from several areas and have generated a description of the high fear of failure person as one who is overly susceptible to conformity pressure, unlikely to participate in achievement activities unless practice is allowed, using whatever defences are available to avoid failure, suffering performance impairment on complex or novel tasks (especially in stressful/competitive situations) and characterized not so much by unrealistic goal setting as by behaviour which avoids aspiration altogether.

Such a synthesis is useful and enables some prediction of these individuals' behaviour across situations. However, most personality measures lack this generality. It is not the aim here to provide a synthesis of the diverse literatures discussed above. However, it should be clear that personality variables are largely defined according to the individual's reaction to other people and in various social contexts. To investigate the relationship between these variables and performance without consideration of the social conditions of testing may provide some insights, but is a limited approach, as is that of studying performance irrespective of individual differences.

2.4 AN OVERVIEW OF SOCIAL SITUATIONS AND PERFORMANCE

Considering the various social situations discussed and the multitude of influences which may operate within them, some synthesis of these different literatures within some common framework would be useful. It was alluded to earlier in this chapter that ego-threatening situations as well as those involving mere presence could be conceived of as varieties of stress stimuli. Indeed, although the stress literature is by no means cohesive, most definitions of stress would accommodate the social situations discussed here.

The most primitive stress model is that extrapolated directly from engineering, i.e. of applying an external force (McGrath, 1970a). Although it should not be assumed that individuals (like machines) enter any situation in a resting or unaroused state (Atkinson and Cartwright, 1964), the experimental manipulations discussed in the previous review are applied in addition to or superimposed over a neutral or less stressful condition.

Another model conceptualizes stress as an imbalance between the demands of the situation and the response capacity of the organism (McGrath, 1970a). The imbalance may only be anticipated rather than actually experienced (Lazarus, 1966). Selye (1950, 1976) considers stress to be basic to all adaptive reactions: a stressor is experienced, physical and psychological mechanisms are called into operation and an adaptive response follows, or maladaptive responses if mechanisms are inefficient (Dohrenwend, 1961).

Therefore, a stress situation is one in which external forces or stimuli are imposed, distress or anticipation of distress may be experienced and this will result in adaptive or maladaptive behaviour.

It is important to note that stress need not be experienced as distress or threat; the individual may be unaware of any 'felt' stress (Lazarus, 1966) or may even feel euphoric (Selye, 1976). All of these characteristics of stress describe the social situations in this discussion.

In some ways an attempted assimilation of these situations may be viewed as contradictory to the basic approach taken in the preceding review, which focused on the diversity between social conditions rather than similarities or common factors. It will be recalled that a major criticism of social facilitation research and theory was the attested equivalence of audience and coaction situations. On the other hand, the separate reviews of studies on mere presence, competition and ego-threat, as well as the brief discussion of personality factors, shows clearly the waste and replication which can result when similarities (as well as differences) between variables are ignored. In addition, whole theories may be built around a narrow range of situations in which only quite specific behaviours are examined. Spence (1958), while admitting to deficiencies in his drive theory, argued that its primary purpose was to "provide for the unification of what, without the theory, would be a multiplicity of isolated or unconnected facts and laws." (p.140). It would appear that some middle ground is required in the study of social situations and performance which, while recognizing the common elements between social environments, also highlights those factors which differentiate them in significant ways. The conceptualization of all these situations within the framework of stress research may provide a starting point for further synthesis and refinement.

Further refinement is clearly necessary as there are great differences between the perception of and response to stress in the form of disaster, test failure or interpersonal conflict. Even on the basic level of physiological responding, Lacey (1967) maintains that there are at least three different kinds of arousal; autonomic, electrocortical and behavioural. Although all three often occur together, they do not always do so and measures from one response system cannot be used to predict responses from the others. It is unlikely that psychological reactions to different stressors would be any less complex.

One characteristic of the social stress situations described in these two chapters which differentiates them from other types of stressors is the role of cognitive mediation or appraisal (Speisman et al, 1964; Lazarus, 1966). In all of these situations the individual must perceive the relevant stimuli, e.g. instructions, presence of others, and then put an interpretation on the situation, e.g. 'this is a failure situation', 'an occasion to meet new people' or 'a chance to demonstrate skill'. This awareness and interpretation of events need not occur when stress is in the form of heat, light or task difficulty. McGrath (1970a) argues that it is the feature of cognitive appraisal which differentiates psychological from physical stress. This is also what makes the study of psychological stress difficult, as individuals' appraisals vary and the responses based on them differ accordingly. Unfortunately, it is often the case that appraisal and reaction processes cannot be experimentally separated. For example, some persons cope with extreme stress by denying it (Friedman et al, 1963; Katz, 1970). This confounding of appraisal and coping plus the subjective nature of most research on cognitive appraisal makes the investigation of appraisal strategies problematic (Mechanic, 1970).

McGrath (1970d) has provided a useful classification of stress research into the categories of 'real' and 'experimentally contrived' situations. One of the three 'contrived' categories is labelled 'social-psychological' conditions. These consist of: (1) evaluation threat, (2) task failure, (3) anomalies in social reinforcement, (4) interpersonal disagreement (including conformity studies), (5) role and status ambiguity and (6) vicarious stress based on viewing threatening social stimuli. Presumably a subcategory for mere presence conditions could be included. The effects of others in the environment would seem to be mediated by what the individual perceives them to be doing, their role in relationship to him and whether they constitute a source of threat or support. Hence the specification of cognitive mediation is met, as are other characteristics of stress, i.e. an external force, anticipation of distress, a stimulus to adaptive/maladaptive behaviour.

Furthermore, in scanning McGrath's six social-psychological stressors, it can be seen that three of the subcategories involve to some extent failure or potential evaluation; evaluation threat, task failure and interpersonal disagrement. Considering that these three diverse situations as well as mere presence conditions are related by a common factor, actual or potential evaluation, it seems legitimate to consider them as discrete units within the more general category of social-evaluational stress. Such a category could encompass all of the situations discussed separately in these two chapters.

Clearly, none of the above social situations is 'pure' in that influences from other categories are absent. As mentioned previously, coaction situations are likely to involve varying amounts of competition. In addition, considering the literature on achievement motivation/fear of failure, it seems unlikely that individuals at the extreme ends of the distribution differentiate between subtle instructional variations. Persons high in fear of failure are likely to perceive any performance situation as one of possible failure, and the opposite set is probable for high nAch individuals. However, a number of environmental features which are likely to affect the evaluative salience of any given situation have been pointed out. The recognition and control of these factors should enable closer approximations to the discrete social-evaluational stress situations described above and a better understanding of their effects on performance.

CHAPTER III

EXPERIMENTAL RATIONALE AND METHODOLOGICAL PROBLEMS

3.0 PLAN OF EXPERIMENTS

The Social Situations

The following experiments examine the effects of four types of social conditions: coaction, audience, competition and ego-threat. These situations are superficially similar in the amount and quality of interaction possible, the presence of evaluative stimuli or the possibility of mere presence effects. However, beyond these gross similarities, they differ markedly in the sources of performance feedback available and the explicitness of competitive and evaluative stimuli. Few experiments have manipulated all of these situations within the same experimental framework, making it difficult to ascertain how different or similar they are in their effects on performance. The primary aim of the present work was therefore to vary the social situation in these ways and examine the separate and interactive effects of each on performance.

Tasks

It is often the case with research in experimental social psychology that performance tasks are chosen on the basis of the ease with which they can be administered and subjected to quantitative analysis - probably because the task usually only serves as a means for studying a quite unrelated social phenomenon such as conformity, risk-taking etc. This need not be a fault, provided consideration is given to the complex ways in which features of the task interact with social stress variables. Recent research has tended to give more recognition to the importance of specific performance tasks, although there is still little information available concerning the separate effects of social stress manipulations on speed, accuracy and variability, and the author has located no studies which use tasks demanding the simultaneous utilization of information from two different sources, as is required in commonplace skills such as driving a car. Therefore, the second aim of this research was to provide detailed analyses of performance using carefully selected tasks. The publicness of the task and the availability of performance feedback (related to both evaluation potential and competition) were also considered, and three of the experiments directly manipulated these variables.

Individual Differences

The third aim of the present work was to incorporate measures of individual differences in personality into the analysis of performance data. This was an exploratory effort as resources were not available to permit a detailed, carefully controlled manipulation of personality variables. Indeed, the literature on social situations provides few guidelines as to thedirection that such investigation should take. The particular measures used were chosen largely on an intuitive basis, and it was hoped that trends arising from initial experiments would suggest lines for further research.

3.1 EXPERIMENTALLY INDUCED SOCIAL STRESS

3.1.1 Generalizing to Real Situations

One of the main issues surrounding all of social psychological research is that the situations studied are usually contrived solely for the purposes of the experiment: field studies are relatively rare because of the difficulties inherent in controlling the variables of interest. The artificiality of the social psychology laboratory and the applicability of laboratory findings to real situations has been discussed thoroughly elsewhere (Argyle, 1969; Appley and Trumbull, 1967; McGrath, 1970c), therefore the topic will not be dealt with in depth here.

It is recognized that rigorous experimental procedures may in fact affect the processes under investigation, and that the responses emitted by subjects and noted by experimenters are limited by the constraints of the controlled situations. It is not clear however how seriously results are affected by these constraints. Argyris (1968) has drawn several parallels between the social structure of the experimental situation and that of organizations, i.e. subject/experimenter, employee/manager, etc. - suggesting that, although the experimental setting may impose unnatural constraints, so might many natural settings. As with any laboratory based research, results must be replicated in real-life settings before generality can be assumed. This caution does not discredit the experimental method, but does highlight the need for careful attention to the procedures by which stress is created and the differences which are likely to exist between the laboratory and the natural situation in terms of the variables investigated.

3.1.2 Evaluation Stress

A common problem with experimental work involving potential or anticipated evaluation is the difficulty of creating a situation which is realistic enough to evoke a stress or anxiety response (Weick, 1970; Appley and Trumbull, 1967). Typically, experiments have employed instructions such as those described in the previous chapter. But, it is doubtful that these contrived situations have the same impact as real evaluative settings. Subjects are usually given no forewarning and thus anticipation of evaluation is minimized. In addition, the tasks used often seem irrelevant to subjects. In such situations the consequences of failure are only the momentary loss of self-esteem or the embarrassment of being judged incompetent by the experimenter. There are no doubt subjects who will fail to be influenced by this type of social pressure. In real-life situations failure has real consequences quite apart from and in addition to social embarrassment.

One specific problem with the use of evaluative instructions is that of highlighting the evaluative cues for the subject without creating explicit demand characteristics.[1] Since experimental tasks are usually novel, subjects may focus on the practical features of the instructions which explain how to do the task and miss the more subtle evaluative cues. Over-emphasizing evaluation is likely to reveal the nature of the experimental hypothesis.

The procedure adopted in the present work was to present the task as one related to an ability and also provide a legitimate reason why performance level was being studied under the experimental conditions. The comparison of subjects' scores with those of others was also mentioned, reinforcing the idea of evaluation without needless repetition. (This is in line with Long and Bessemer's (1971) findings that, when varying several sets of instructions, only those which mentioned evaluation of an ability and comparison with others produced any changes in performance.) Even though such precautions were taken, it is suspected that some subjects failed to be motivated by the stressful instructions, and the situations created, although stressful for some, were generally less potent than real performance conditions in which failure has practical and sometimes long lasting implications.

Care was also taken in the creation of neutral instructions so that subjects were given a realistic explanation of the situation which de-emphasized personal evaluation. This is in contrast to the procedure of simply providing instructions about how to do the task in the neutral situation (allowing subjects to make their own interpretation of the situation) or to the rather naive instructional sets of "this is a test/this is not a test". However, evaluative cues may

---

1. Demand characteristics refer to cues in the experimental situation which suggest to the subject what behaviour is expected. The danger is that subjects will wish to behave appropriately and will react to these cues in a way which confirms the experimental hypothesis.

exist in many forms, including the experimenter's manner and probably several environmental stimuli which have not been identified. The experimenter tried to interact in a standard fashion with all subjects; this was friendly though reserved. Conversation was limited to questions about difficulty in finding the room, the weather, etc. Based on post-experimental interviews this strategy appeared to be successful. The more serious problem was that of getting subjects to accept the stressful nature of the situation rather than its neutrality, as is usually the case in social stress research (McGuire, 1969).

3.1.3 Social Stress and the Presence of Others

The creation of the social situation is less problematic than that of evaluation stress as either others are present or they are not; deception is not involved. Still, the situations under consideration here are rather unlike most naturally-formed social groups. Interaction is limited, selection of group members is random, future interaction is unlikely or at least would probably not be influenced by the experimental situation. Thus the pressures which function in most natural groups are absent or at least modified in these minimal interaction situations. Indeed, there is controversy as to whether such groupings of individuals constitute true social situations. The reader is referred to Cartwright and Zander (1968) or Argyle (1969) for a detailed discussion of this issue. However, the question being considered in this work is the effect of minimal social conditions on individual performance. The identification of these fundamental social processes, if they exist, should form the necessary base on which theories dealing with more complex group structures could be elaborated.

Several different social situations were used in the present research and these are defined as follows:

alone (A) - the subject works on the task with no other persons in the room.

experimenter only (E)	-	the subject works on the task with no other people in the room but the experimenter, who does not observe the subject per se but carries out part of the experimental procedure.
audience (Aud)	-	the subject works on the task while being directly observed by one or more persons who are not involved in conducting the experiment. Observers simply watch the subject and do not speak to or distract him in any way. In one case a video camera was employed as an audience.
implicit competition (Im-Comp)	-	several subjects work individually on the same task either simultaneously (coaction) or in turns. Verbal interaction is minimal. Groups vary in size between experiments but never are less than 3 or larger than 5. Task instructions are neutral, i.e. evaluation is minimized.
ego-threat (T)	-	the subject works on the task either alone or with the experimenter present but is given explicit instructions that his performance is going to be evaluated and compared with others.
explicit competition (Ex-Comp)	-	defined in the same way as implicit competition except that instructions are ego-threatening rather than neutral.

Ideally, alone situations should be used when making comparisons between social and non-social conditions, and this procedure was followed whenever possible. However, in several experiments it was necessary for the experimenter to remain in the room to monitor equipment or record scores. During such times, she avoided observation of the subject, remained as far away as possible, and appeared to be occupied with other duties. The presence or absence of the experimenter was always constant across conditions and was only permitted after it had become obvious (based on previous experiments and feedback from subjects) that she had little effect as an audience.

3.1.4 Quantifying Stress

Finally, a primary problem with experimentally-created social stress is that of quantifying the intensity of the resulting stress. Prediction is based on the degree to which subjects are affected by the various experimental manipulations, and with no truly reliable independent measures for ascertaining resulting stress or arousal, predictions are usually speculative and post hoc interpretations often arise.

The effectiveness of the experimental manipulations were always informally discussed with subjects during debriefing sessions, and occasionally self-report questionnaires were issued - although these were judged to yield about the same quality of information as the informal interviews.

The use of physiological measures was considered but decided against in view of the problems with obtaining reliable readings and making interpretations based on these (Lacey, 1967). Also, although independent validation of arousal level using self-report, performance and physiological data would have been desirable, there are few examples in the literature where such validation has been successful (McGrath, 1970c). Lazarus (1967) argues that coping processes affect response systems in different ways, carrying the implication that looking for validation across response systems may be a misguided effort in the first place.

Even so, information on the ways in which these response systems differ in relation to specific stressors would have been an interesting contribution, and monitoring of all three response systems would have been attempted had any feasible means of doing this been apparent. However, physiological responses can be sensitive to almost any change in the experimental procedure. In order to establish valid individual baseline readings and allow for adaptation to superfluous changes in the procedure, repeated exposure to the experimental situation may be necessary. This would have been an insurmountable problem in several experiments which employed large numbers of subjects (60 to 100). In addition, there are few physiological measures which lend themselves easily to application in a group situation where several subjects require simultaneous monitoring. Apart from these practical difficulties, a major consideration against the use of physiological recordings was that of further accentuating the artificial, laboratory nature of the situation. Since there was some doubt that valid physiological readings could be obtained in the situations of interest anyway, a more realistic performance situation was opted for.

In part, it was hoped that the addition of individual differences measures would provide some independent confirmation of the effectiveness of the stress manipulations. Measures of both intrinsic and situational susceptibility to stress were employed. Thus any performance differences between subjects varying in self-reported anxiety or endogenous arousal level could be interpreted in terms of the conditions of performance and compared with the existing literature concerning these personality dimensions and their expected relationship to performance and stress.

3.2 SUBJECT RECRUITMENT AND EXPERIMENTER EFFECTS

3.2.1 Demand Characteristics

Another facet of the controversy over 'artificiality' is that of the 'naive' subject and demand characteristics. It has been argued that a psychology experiment constitutes a special social situation (Orne, 1962, 1969) defined by the relationship between the subject and the experimenter. According to Orne (1962) subjects try to guess the hypothesis under investigation (based on cues present in the situation, i.e. demand characteristics), and, having done so, comply with the experimenter's expectations. The problem becomes more serious as subjects become experimentally sophisticated and better able to recognize deception techniques.[2]

However, there is no general consensus on the extent to which demand characteristics are influential, as some experiments have failed to find "compliance" (Page and Scheidt, 1971), while others have identified an opposite tendency on the part of subjects, "sabotage" (Argyris, 1968; Berkowitz, 1971). Berkowitz (1971) questions whether the majority of subjects are sufficiently insecure to modify their behaviour in response to the mild pressure present in most experimental situations. But, there is little doubt that as subjects become more experimentally sophisticated they also become more alert to the possibility of some 'real' purpose behind the experiment (Page and Scheidt, 1971; Simons and Turner, 1976). This means, of course, that a wrong hypothesis may be formed which can affect subsequent behaviour (McGuire, 1969). Weber and Cook (1972), in a comprehensive review, concluded that the experimental evidence for "compliance" and/or "sabotage" in subjects is equivocal and that the most obvious biasing factor is the subjects' concern about how their performance will be evaluated.

There was some evidence from the present experiments, based on post-experimental interviews, that subjects did try to guess the 'true' nature of the experiment. A few expressed concern that they had "messed-up" the experiment, others stated that they had formed hypotheses which they had then tried to act against. However, it is unlikely that the experimental findings were biased in any particular direction by

---

2. Demand characteristics are not to be confused with the experimenter expectancy effects noted in classroom settings (Rosenthal, 1969; Beez, 1972). The latter is not "compliance" as Orne uses the term, as it refers more to the differential cueing of subjects by experimenters (teachers).

these reactions; rather, they constituted another source of error variance. Although most subjects reported awareness of the competitive/evaluative nature of the social situations, none could have predicted what specific changes were expected on any particular task, i.e. speed, accuracy or variability, or how performance might be modified by the presence or absence of others.

3.2.2 Sampling Biases

Most of the experiments employed university undergraduates as subjects. Therefore, another possible criticism is the unrepresentative nature of the sample. For example, it is likely that university students are more oriented toward competitive/evaluative stimuli and may respond in ways which are not characteristic of a less select group. Also, having had more experience and relatively more success in performance situations, students have had more opportunity to develop strategies for coping with evaluation stress and may be more skillful competitors than the general population.

However, in terms of practical utility, findings about the effects of social-evaluative stress are most relevant to those persons who are frequently exposed to such stress; university students would be one of the target populations to which the results would be expected to apply. On the other hand, because they are perhaps more likely than most people to perceive performance in social situations as competitive/evaluative, they are probably not the ideal population for studying the independent effects of the mere presence of others.

To test the generality of results, three different populations were used; university undergraduates (mainly psychology first year students), Open University (OU) students and fifth and sixth form school children. The OU students provided a sample with a greater range of age and background, and the school children were less selected on the basis of ability, academic motivation and experience in coping with evaluation stress. These populations were used when available to test the generality of certain results obtained with university students. No experiments were exactly replicated with 'cross-cultural' comparison in mind, although the basic procedures, social manipulations and evaluative stimuli were similar, enabling some general comparison.

Subject availability was a constant problem and in some cases set practical limits for the type of experiment conducted. All subjects were volunteers in that they had the right to refuse participation. Most were approached individually or in their classrooms and asked to participate. A few signed recruitment lists posted in various university buildings. One experiment (II) offered a small monetary incentive, but this attracted few subjects and was not employed in subsequent experiments.

It has been suggested that volunteer subjects might be a select group on several dimensions of personal adjustment (Rosenthal and Rosnow, 1969, 1976). But, this was not really a problem in the present series of experiments as there were in fact very few spontaneous volunteers. Nearly all subjects agreed to participate after being personally asked, and the refusal rate in these circumstances was low, possibly due to the "conformity" phenomenon noted by Rosenthal and Rosnow (1969) in public volunteering situations. In addition, efforts were made to reduce sampling biases due to 'no-shows' by sending personal reminders to all subjects and rescheduling those who failed to report for the experiment the first time.[3]

3. 'No-shows' refer to volunteers who fail to report for the experiment. There is speculation that these may be people who volunteer in compliance to conformity pressures at the time of recruitment but who would not spontaneously volunteer. If they are omitted from the experiment, the sample then more closely resembles one of pure volunteers.

3.2.3 Experimenter Effects

Since only one experimenter was used throughout the series of experiments it was not possible to ascertain what effect the personal characteristics of the experimenter might have had on the findings (Rosenthal, 1969).

That previous exposure to the experimenter may have confounded results was considered. Such an effect has been found when exposure occurs in the same situation, such as repeated IQ testing (Sacks, 1952). Although other authors have failed to obtain differences between subjects with varying amounts of previous contact with the experimenter (Rosenkrantz and Van de Reit, 1974; Thomas and Fortson, 1975). Such effects were judged to be minimal in experiments involving school children and OU students, as previous contact was only momentary (at recruitment). When university students were used care was taken that none be included in the sample with whom the experimenter had more than a passing acquaintance or for which the experimenter might be involved in actual academic assessment. None of the subjects had any previous contact with the experimenter in any performance situations for which some generalization of affect might have been anticipated.

The sex of the experimenter was not of course controlled for. However, all designs were balanced for subjects' sex. Therefore, any differences found between males and females which were not anticipated and could not be accounted for on the basis of known sex differences for particular tasks, e.g. motor tasks, could be due to a possible experimenter sex effect. In any event, such a situation never arose, since sex differences, as will be seen, were minimal.

3.3 TASK VARIABLES

The importance of task characteristics in performance situations has already been discussed (Section 1.3). The availability of feedback and the clarity of performance goals were felt to be particularly important variables in the present work. The social situation would be expected to be less influential with highly involving tasks, due to increased attention to the task itself. Also, when performance goals are clear, information provided by other group members about performance standards is redundant. Even in cases where goals are not specifically set it is sometimes, by definition of the task, obvious what is meant by good or bad performance. For example, good performance on the pursuit-rotor is clearly a high time-on-target score. The goal is less conspicuous, however, on a self-paced cancellation task, in which there is a speed/accuracy trade-off and no standards need be available as to how much work should be completed.

Several different tasks were employed in the following experiments, partly to identify what type of task would be most sensitive to the particular stressors being investigated, and also because the generality of effects across tasks was a point of interest. For the most part learning tasks were avoided, as these involve special problems with subjects learning at different rates, training schedules, warm-up and reminiscence effects and the overlap of learning and performance phases. Measures of pure performance were aimed for, the rationale being that these would provide more stable baselines and show most clearly the immediate impact of changes in the social environment.

Specific descriptions of the tasks will be given in the experimental chapters, but generally, those used required little training but demanded constant concentration and effort. Tasks for which execution quickly became automatic (such as simple addition) were rejected, since only quite extreme social stress would be expected to alter performance on such routine and well learned activities. Except for the last two experiments (Chapter VI), cognitive tasks requiring sustained attention were employed, these providing measures of speed, accuracy and in some cases variability. It was considered that these met the criteria for 'sensitive' tasks.

Tasks having a heavy motor component were employed in the final two experiments in order to suggest a more realistic competitive situation. Both allowed for public evaluation and feedback of results, accentuating comparison between subjects. Speed was not a factor with either of these tasks, although accuracy and variability were measured. Both required considerable concentration and precision for efficient performance.

Although measures of pure performance were desired this proved to be quite difficult to achieve, as practice effects were noted even on tasks for which a minimal amount of training was needed. In addition, although it was always intended that subjects should not be given knowledge of performance standards, this was easier to achieve with some tasks than with others; this point will be discussed more fully in the experimental chapters.

3.4 INDIVIDUAL DIFFERENCES

Despite the multitude of measures available, personality tests have proved to be only moderately successful in providing reliable and meaningful information about the relationship between dispositional differences and performance (Eriksen et al, 1952; Argyle and Little, 1972; Mischel, 1968). Part of the difficulty surrounds the assessment of personality, which has relied heavily on questionnaire type instruments. The problems with such measures have been thoroughly discussed elsewhere (Crowne and Marlowe, 1964; Anastasi, 1968; Tyler, 1956; Sarason and Smith, 1971); these relate to response sets, faking of responses, differences in the interpretation of questions and the possible irrelevance of the imposed categories for any particular individual. Such instruments provide only gross descriptions of an individual's personality and may perhaps only be useful when a score departs radically from the mean.

Failure to find stable relationships between personality test scores, performance and social behaviour has led some theorists to suggest that the whole concept of a stable personality is erroneous and that behaviour is 'situation bound' (Mischel, 1968). Several recent papers discuss the controversy over "situationism" versus "individualism" and opt for a view which minimizes the importance of personality differences and stresses the powerful effects of specific situations (Secord and Backman, 1974; Bem and Allen, 1974, Bowers, 1973). Mischel (1973) however takes issue with this extreme view, pointing out that the clarity of situational cues determines the power of the situation, and that individual factors are obviously important but their influence is mediated by a multitude of variables e.g. sex, age, IQ. He concludes that predictions can be made concerning subject variables and relevant behaviours but there are severe limits on the range and level of the relationships which can be expected. Argyle and Little (1972) argue that a larger proportion of the variance in behaviour can be accounted for by the interaction of subject and situational variables than by the separate effects of either.

In the present work information was sought about the role of individual differences, but the methods used were exploratory. Firstly, although the inadequacies of personality questionnaires were recognized, there seemed no other feasible way of assessing the large numbers of individuals employed in several of the experiments. Secondly, there were no strong guidelines to suggest which dimensions of personality might be most important in the situations of interest. Recent discussions favour the use of specific rather than general measures of personality for comparisons with performance data, due to the relatively greater success of the former in accounting for differences in performance (Sarason, 1961, 1959b; Lamb, 1976). However, tests measuring both traits and specific characteristic were used, and most of those selected already have a substantial literature surrounding them. The personality dimensions chosen were those which intuitively seemed relevant to social-evaluative situations and which have in the past shown relationships to either social behaviour or performance measures, these being:

Fear of Failure	-	Test Anxiety Scale (Sarason, 1972), Argyle-Robinson Achievement Motivation Questionnaire (Argyle and Robinson, 1962).
Audience Anxiety	-	Audience Sensitivity Inventory (Paivio, 1965).[4]
Sociability	-	The Eysenck Personality Inventory, E scale (Eysenck and Eysenck, 1964).
Endogenous Arousal	-	The Eysenck Personality Inventory, E scale and N scale (Eysenck and Eysenck, 1964).

It was not expected that strong relationships would be obtained between these measures and those of performance. However, by using well-established tests, any suggested trends in the data could be compared with the existing literature for confirmation and interpretation.

Because of the large number of subjects needed, it was not possible to pre-select individuals on the basis of extreme scores, nor was this clearly desirable. Extreme scoring subjects are by definition atypical

4. The Audience Sensitivity Inventory is a specific anxiety measure similar to the Test Anxiety Scale but focuses on anxiety in audience rather than testing situations. Relationships have been found between audience presence and ASI scores for both overt behaviours (Paivio, 1965) and autonomic funtioning (Simpson and Molloy, 1971).

and may reveal little in terms of the effects of the measured dimension in the population at large. Also, with no clear predictions in mind, this seemed a costly procedure, at least for initial experiments. Furthermore, an extreme groups design is only justified when relationships between the variables of interest are known to be monotonic, and this is not the case with anxiety measures (see Figure 2.1).

3.5 DESIGN OF EXPERIMENTS

Because of the considerable individual variation in reactions to stress stimuli, most experiments in this area employ a repeated measures design. In this way subjects serve as their own controls and difference rather than absolute scores are analysed, filtering out much of the individual variation.

However, a major problem with this type of design is the possibility of carry-over effects from one condition to another. Not only is there likely to be a practice effect between subsequent testing sessions, but exposure to the experimental setting in one condition may modify the subjects' reactions on following days. This issue has been discussed by Poulton and Freeman (1966) and is of central importance in stress research.

There are two procedures commonly used to control for carry-over effects; counter-balancing and the use of a pre-test.

In the counter-balanced design, half the subjects are exposed to the stress condition first followed by the control condition, while the other half are tested in the reverse order:

$$A1 \longrightarrow B2$$

$$B1 \longrightarrow A2$$

In the analysis of results, main effects for the treatment (within subjects) can be examined as well as the interaction with order of testing (between subjects). Carry-over effects can be asymmetrical and this should be reflected in a significant Order by Treatments interaction. Such an effect is obtained when subjects show an unexpected advantage or disadvantage in the second session which is related to the particular conditions in which they were first tested, i.e. (A2 - A1) ≠ (B2 - B1). If such an interaction occurs the effects ascribed to the treatments will be obscured or exaggerated.

When a pre-test is used, all subjects are tested in the control condition first and are then divided into control and experimental groups for subsequent tests:

$$A1 \longrightarrow A2$$

$$A1 \longrightarrow B1$$

This design assumes practice and other carry-over effects to be the same for all subjects in subsequent testing sessions, differences observed between control and experimental groups in later sessions are then attributed to the experimental manipulations. Of course, this design does not allow for examination of the treatment without prior experience in the experimental situation, since the analysis is conducted on the data from the second session only (Lana, 1969). But this would not seem to be a serious problem in performance research, as, regardless of the design, subjects are usually given practice with the task in order to establish a stable baseline for the experimental treatments.

A repeated measures design was employed in all but one of the following experiments, partly to reduce individual variation, but also to make the analysis of personality differences data more meaningful. Regarding this latter point, when extreme groups are employed in personality research, independent groups can be used with personality as a factor in the design; differences in ability and other variables would be assumed to be randomly distributed over both personality groups. However, as mentioned above, subjects in these experiments were not pre-selected, and the resulting scores did not fall into groupings which would enable a meaningful comparison of mean differences. Therefore, correlational analyses were conducted on the data; the correlation statistic allows for comparisons of the direction and magnitude of relationships across experiments as well as permitting an examination of the shape of the distribution (should curvilinearity be suspected). However, since the range of initial ability was great on many of the tasks, the correlation between absolute performance scores and personality data would naturally be confounded by the subjects' ability to do the task. As relationships were expected to be low in the first place, the difference score, which is less likely to be contaminated by differences in ability level, was judged to be the only performance measure which could be meaningfully compared with personality test data.

The only experiment in which independent groups were used (Experiment II) was conducted primarily to examine the effects of the social manipulations without the confounding effects of practice. In all other experiments counter-balanced or pre-test designs were employed. The latter was favoured in later experiments because of the length of time between experimental sessions (sometimes more than a week) and the possibility of subjects communicating with each other about the experimental procedures during this interval. Also, the effects of order proved to be so pronounced that there seemed little advantage in employing the counter-balanced design in these situations. Only in the final exploratory experiment involving darts throwing was such a control not employed. In this case performance had reached a stable baseline prior to testing (eliminating practice effects), and subjects were tested repeatedly in both control and experimental conditions, which should have highlighted any carry-over effects between sessions.

All experiments employed factorial designs with provisions for analysis of performance over time. Parametric analysis of variance was used as the sample sizes were sufficiently large to minimize any effects due to irregularities in the distribution. Personality data are displayed in correlation matrices. Analysis takes the form of looking for consistent patterns in terms of both direction and magnitude over the whole series of experiments. Little weight is placed on single, significant correlations, since with multiple correlations a percentage of the comparisons would be significant due to chance. In any case, statistical significance may be misleading, as experiments varied in the size of the samples employed and an examination of only significant relationships could obscure consistent and interesting patterns in the correlations between experiments. Because these data are examined across experiments the results are not discussed in the experimental chapters but are summarized in Chapter VII; the data are presented in full in Appendix I.

CHAPTER IV

EXPERIMENTS I, II AND III

4.0 INTRODUCTION

The three experiments reported here are concerned with varying combinations of coaction, competition, audience and instructional manipulations. Experiment I is exploratory, designed to investigate the sensitivity to social conditions of the cancellation task which is employed in several subsequent experiments. Experiment II expands the findings of the first study and examines the separate effects of all of the above mentioned social situations. Experiment III attempts to explore further one specific finding arising from Experiment II, concerning shifts in attention under different social conditions.

Considerable care was taken in the choice of a task for this initial series of experiments. Certain features were felt to be requisite. Firstly, the task must be 'sensitive'. Ideally it should allow for the measurement of speed, accuracy and variability. Secondly, it should be easy to administer and require a minimum of prior training. There are obvious practical advantages to tasks with these latter specifications, but apart from these, it was felt that if subjects had little prior experience in the experimental setting they would be more likely to respond to the social stress manipulations. Novel situations can be structured more easily to arouse social anxiety and subjects have relatively little time to develop defensive or coping strategies. A third requirement was that the task should demand continuous attention so that subjects would be taxed to perform well, and finally it should be boring, as one which is intrinsically interesting may well be less susceptible to influences from the social environment.

All three experiments use the cancellation task employed in Experiment I as the primary measure of performance. This is basically a measure of attention, usually considered in relation to perceptual selectivity or visual search (Poulton, 1970). The task permits analysis of speed, accuracy, and changes in performance over time. Within subject variability is determinable but proved to be fairly low on this routine activity and was not therefore calculated.

Experiments II and III use cancellation in combination with a running digit span task (RDS). Attending and responding to two sources of information simultaneously offers a more sensitive performance situation than that of simple cancellation; more measures are available for analysis and the possibility of shifts in attention from one aspect of the task to another can be investigated. In general, the dual task is more demanding, stretching the individual's performance capacity, and is therefore more likely to be sensitive to environmental stressors. In previous studies using analogous tasks, subjects have shown reduced attention to secondary sources of information in response to various experimental treatments including noise (Hockey, 1970), heat (Bursill, 1958), threat of shock (Wachtel, 1968) and manipulations of endogenous drive (MA) (Zaffy and Bruning, 1966). These studies support Easterbrook's (1959) theory of restricted cue utilization under conditions of high arousal.

All of the experiments employ Durham University students as subjects, mainly first year psychology students.

In Experiments I and II subjects were administered the Eysenck Personality Inventory (EPI) (Eysenck and Eysenck, 1964) and the Argyle-Robinson Achievement Motivation Questionnaire (q Ach) (Argyle and Robinson, 1962). The latter has subscales measuring need for achievement and fear of failure which would both seem to be relevant in competitive/evaluative situations. However, the range of scores on both subscales was too narrow for a meaningful analysis. In Experiment III the q Ach was replaced by Sarason's (1972) Test Anxiety Scale (TAS), which measures only fear of failure and is specific to testing situations (not a test of general personality disposition as is the q Ach). The EPI was not used in Experiment III, as the study dealt quite specifically with audience and competition effects. Paivio's (1965) Audience Sensitivity Inventory (ASI) was substituted providing a measure of specific anxiety in the presence of an audience.

4.1 EXPERIMENT I

4.1.1 Aims

The primary aim of the present experiment is to investigate performance on a cancellation task under conditions of alone (A) or Explicit Competition (Ex-Comp). Ego-threatening instructions are used in the social situation in order to maximize the differences between the two conditions of testing. It was left to later experiments to determine the separate effects of instructions and mere presence, should an effect be found.

As the task is routine and involves little learning, it might be expected that social-evaluative stress would facilitate performance. Some degradation over time could be anticipated, as the activity is monotonous and requires considerable effort and concentration in order to maintain a high performance level. However, the social condition might serve to maintain higher performance over a longer time span than the alone situation.

4.1.2 Task

The task is a 'proof-reading' (or visual search) exercise which requires line by line scanning of printed material, crossing each 'e' with a single diagonal line. Material was obtained from *Moby Dick* and was chosen for its uniformity of subject matter and descriptive content. As the task was designed to be monotonous, passages which contained elements of plot or character description were omitted.[1] To further reduce the possibility of subjects becoming involved with the story, all punctuation, except fullstops, was omitted and pages were presented

---

1. The sections chosen actually consist of a detailed and laborious description of the whale's anatomy. *Moby Dick*, The Bobbs-Merrill Company Inc., 1964, pp. 427-435.

in random order.

In analysing performance on this task it is usual to use the number of lines completed as the primary performance measure (e.g. Colquhoun and Corcoran, 1964), but this method results in some difficulties when interpreting results. It was expected that Output might be more sensitive to social influence, as subjects can increase speed at the expense of accuracy on this task, thereby accomplishing more work (number of lines) while locating the same number of or even fewer actual targets. Overall efficiency is then difficult to gauge because of a speed/accuracy trade-off.

For analysis purposes three performance measures were taken:

(1) Output - the number of targets (e's) present in the length of material completed. Output was estimated by multiplying the number of lines checked by the average number of 'e's' per line (6.29).

(2) Omissions - the number of omissions ('e's' not crossed) divided by the number of lines completed (no errors of any other type were observed in the data).

(3) Hits - Output minus the total number of omissions. This is a measure of the actual number of targets detected in each time period.

4.1.3 Method[2]

The design is a 2 X 2 X 2 X 4 factorial. 'Between' factors are Sex and Order of Testing and 'within' factors Social Conditions and Time Periods. Subjects were 8 male and 8 female undergraduates who 'volunteered' in the manner described in Chapter III. All subjects were tested once in Explicit Competition (Ex-Comp) and once Alone (A). The design was counter-balanced for order of testing.

In condition A subjects were tested alone with neutral instructions.

---

2. Complete instructions for this and all subsequent experiments can be found in Appendix II.

They were told to work as quickly and carefully as possible and that their scores would later be compared with personality tests they were to take. Subjects in Ex-Comp were tested in coacting groups of four, seated around a single table and facing each other. Instructions were the same as those in A but included a statement that the experimenter was interested in seeing which subjects did the best. All groups were single sexed.

Subjects reported on Day I to the experimenter's office, where the testing took place and were read instructions as to how to do the task followed by 2 minutes of practice. A 5 minute rest period was interpolated between the practice and the test in order to familiarize the subjects with the task and mentally prepare them for the experiment, since on Day II they would come knowing more or less what to expect.

After the rest period, the experimenter read either the neutral or ego-threatening set of instructions and gave the signal to begin working. She then left the room and remained outside until the end of the session when she re-entered and told the subjects to stop. The testing session lasted 20 minutes. A buzzer was sounded at 5 minute intervals, at which time subjects indicated with an 'X' how far they had progressed. This divided the performance session into four time periods for analysis.

Subjects reported for the Day II session between 3 and 7 days later, depending on when an agreeable time could be arranged. Day II was procedurally like Day I, except that no practice or rest period was given. Subjects were told to continue working as they had in the previous session and were then read the alternative set of neutral or ego-threatening instructions.

Following the second session the EPI and qAch were administered. Subjects were instructed to "work quickly and not ponder over questions". In addition they were assured that neither of the tests measured anything "bad" about them and that results would be entirely confidential. When the personality tests were finished a debriefing session was held and the experimental manipulations were discussed.

4.1.4 Results[3]

Results for the three different performance measures will be discussed separately, the data for Hits being considered last since it is derived from Output and Omissions data.

Output

Table 4.1 shows mean performance over time for Output in the Ex-Comp and A conditions of testing.

TABLE 4.1

Mean Output for Social and Non-Social Conditions Collapsed over Order of Testing

Period	1	2	3	4	Totals Mean	S.D.
A	260.3	255.8	262.3	265.2	1043.6	169.38
Ex-Comp	297.2	270.8	269.7	270.6	1108.3	159.34

The table indicates that for Output subjects tested in Ex-Comp performed at a consistently higher level than the same subjects when tested alone. Consistent with this general pattern the analysis of variance shows a significant main effect for Social Conditions ($F=10.98$, $df=1, 112$, $p<.01$). However, although such an effect was expected, the results for Order complicate the picture. Figure 4.1 displays the data for Output in the two conditions but separated into the two orders of testing. It is clear from the Day I data that, apart from Period 1, there is little difference between the performance of

3. Analysis of variance summary tables, means and standard deviations for all performance measures and all experiments are presented in Appendix III.

Ex-Comp and A subjects.

The data from Day II present a different pattern. Firstly, there is improvement in both conditions from Day I to Day II, an effect due to practice which partly obscures the group effect. However, the Ex-Comp condition shows relatively more improvement in Day II, resulting in better performance throughout the session than condition A. The analysis of variance confirms this observation with a highly significant Order X Social Conditions interaction ($F = 41.0$, $df = 1, 112$, $p<.001$). However even when considering only Day II data, a conclusion that this higher performance level in Ex-Comp indicates a true facilitation effect of the social situation must be regarded with caution. These same subjects began showing relatively higher Output in Period 3 of Day I (when in condition A) and may have been, by chance, a better group of performers.

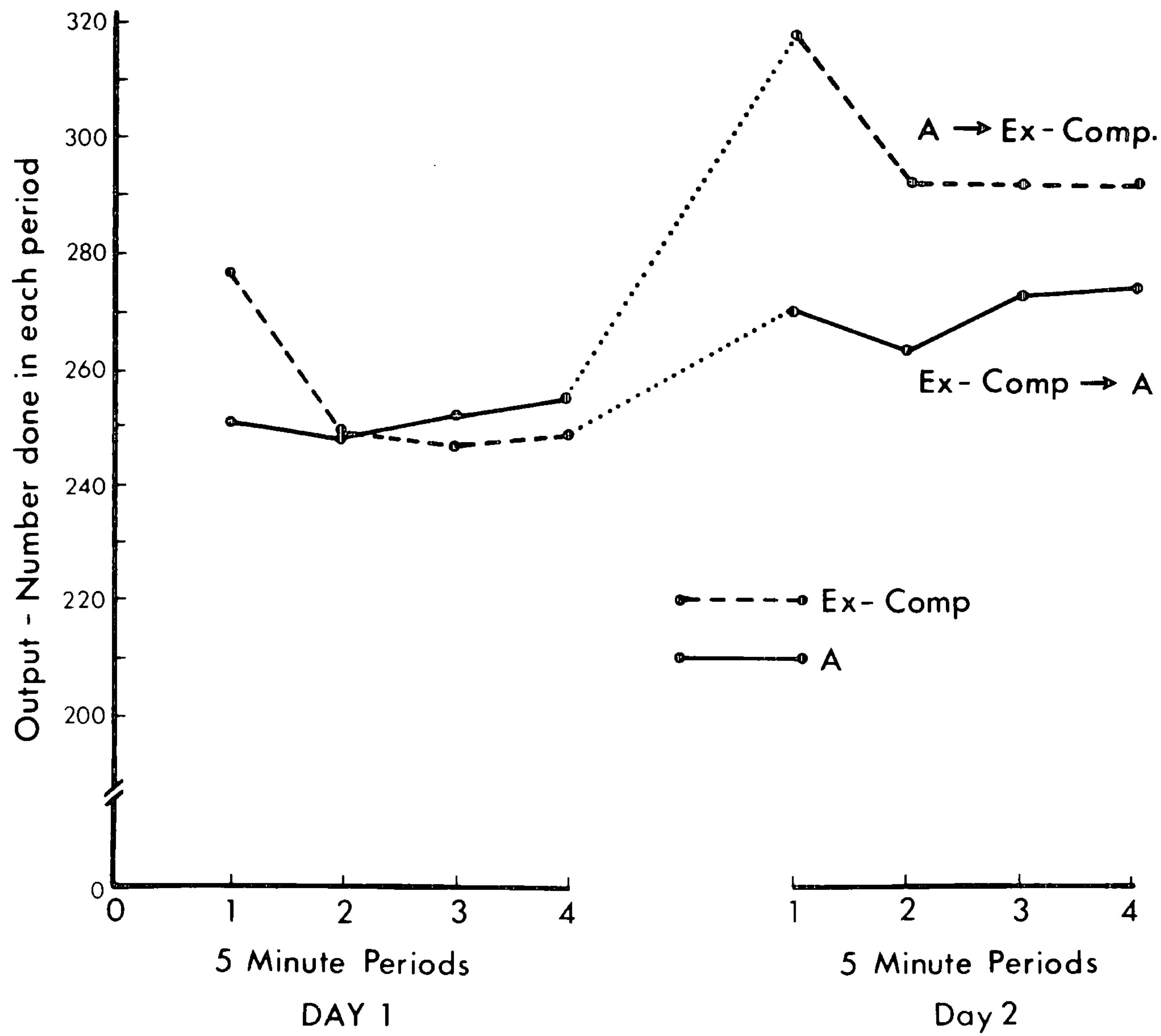
The most convincing data suggesting social facilitation is that for performance over time. In addition to a significant main effect for Time Periods ($F = 5.15$, $df = 3, 112$, $p<.005$), a significant interaction for Social Conditions X Time Periods is present ($F = 9.52$, $df = 3, 112$, $p<.001$). It can be seen from Table 4.1 that subjects in both conditions of testing demonstrate some decrement in Output between Periods 1 and 2, this being more striking for Ex-Comp. Judging from Figure 4.1 this effect would appear to be independent of the order of testing. Thus, on both Day I and Day II subjects show a marked reduction in Output after the first 5 minutes of testing when in the social situation. Apart from this initial drop in performance level in the Ex-Comp condition, there is little further fluctuation over time.

Omissions

Figure 4.2 shows the error rate (mean omissions per line) for the

Figure 4.1.

Mean Output according to Social Conditions and Order over time.



two conditions of testing separated by Order and plotted over time. Note that a low score on this measure indicates good performance. High scores mean that many targets were missed relative to the total number attempted.

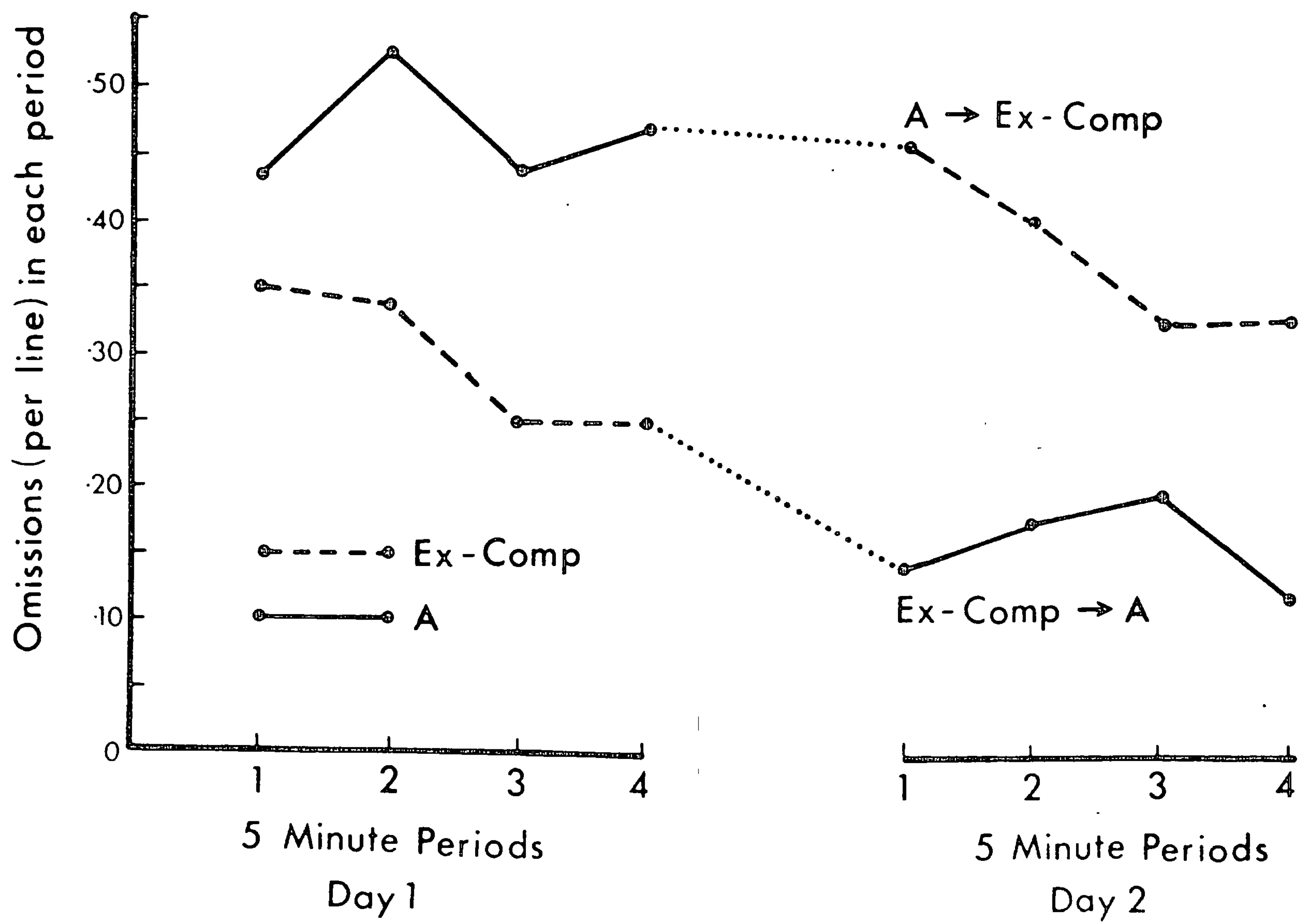
The most striking feature of this figure is the marked superiority in accuracy for subjects tested in the Ex-Comp → A order. Regardless of the social situation these individuals give more accurate performance at all times during both days of testing and even show slightly more improvement from Day I to Day II relative to those in the A → Ex-Comp order. The analysis of variance indicates a significant main effect for Order ($F = 4.65$, $df = 3, 112$, $p<.05$). It is possible that due to the sample size subjects tested in the Ex-Comp → A order were by chance better on this measure than those tested in the reverse order. However, it is also conceivable that being in the social situation had the effect of making subjects more attentive, producing a somewhat more careful performance strategy that carried over into the second testing session, i.e. an asymmetrical transfer effect (see Section 3.5). However, it is difficult to distinguish from the data which of these two explanations is most feasible.

Also, the interaction for Order X Social Conditions is significant ($F = 10.07$, $df = 1, 112$, $p<.01$). This can clearly be seen in Figure 4.2; subjects tested in the A → Ex-Comp order give more accurate performance in the social situation whereas those tested in the Ex-Comp → A order are most accurate when alone. Since both orders show improvement in accuracy on Day II to about the same degree, this effect would appear to be due more to improvement with practice than to the social conditions.

The analysis of variance also yields a marginally significant main effect for Time Periods ($F = 2.75$, $df = 3, 112$, $p<.06$), and Figure 4.2

Figure 4.2.

Mean Omissions according to Social Conditions and Order over time.



shows a general, though sporadic, tendency to improve accuracy over time. The pattern appears more clearly in the Ex-Comp conditions, although no significant interaction was found for Time Periods X Social Condition.

That the two orders of testing show this difference on the Omissions measure has possible implications for the other measures of performance. Specifically, there is a low positive correlation between Output and error rate on this task (r = .15 to .28). A speed/accuracy trade-off would partially account for the discrepancy between the Day I and Day II results for Output, since the more careful subjects (Ex-Comp → A) would be expected to show somewhat lower Output scores. Assuming there was a group effect, this difference in accuracy would minimize the differences in Output between the two conditions on Day I but would accentuate them on Day II. The data for Hits, which takes account of both speed and accuracy, shows a more consistent difference between the two conditions on both testing days.

Hits

Table 4.2 shows the results for Hits collapsed over the two orders of testing.

TABLE 4.2

Mean Results for Hits for Social and Non-Social Conditions Collapsed over Order of Testing

Period	1	2	3	4	Totals Mean	S.D.
A	248.0	241.0	249.6	258.6	997.2	159.17
Ex-Comp	277.6	254.8	257.0	265.3	1054.7	150.22

As can be seen the means follow much the same pattern as that in Table 4.1 (Output). Similarly, the analysis of variance shows a significant main effect for Social Conditions ($F = 7.12$, $df = 1,112$, $p<.02$). Performance over time also follows a similar pattern to Output except that improvement in Periods 3 and 4 is more marked, due probably to the progressive increase in accuracy discussed above. Once again there is a significant main effect for Time Periods ($F = 5.71$, $df = 3, 112$, $p<.005$) and a significant Periods X Social Conditions interaction ($F = 3.15$, $df = 3, 112$, $p<.05$). The latter would appear to be due to the decrement following Period 1. The interaction with Social Conditions is less pronounced on this measure, although the decrement from Period 1 to 2 in the Ex-Comp condition is still, in absolute terms, roughly three times that in the A condition.

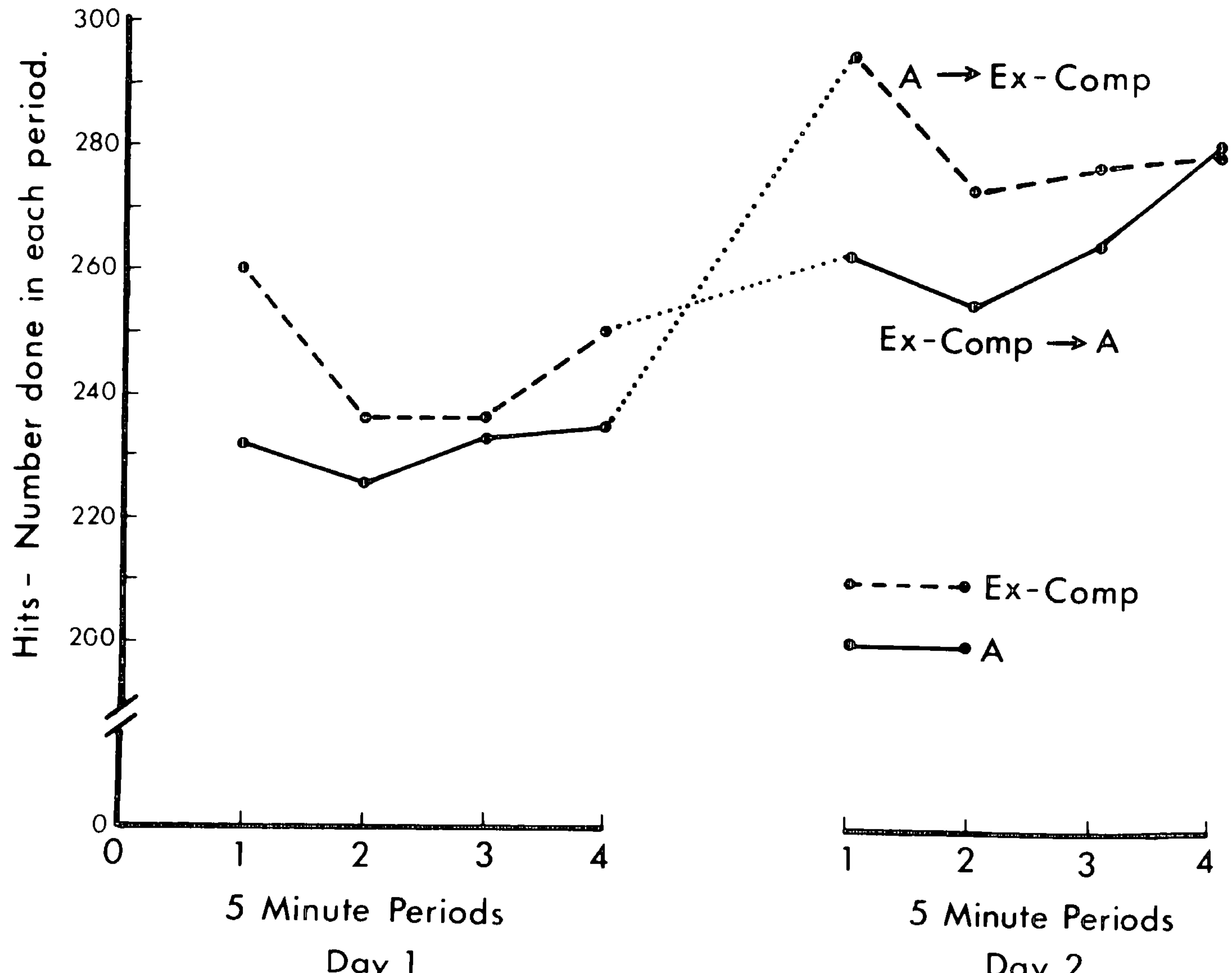
Figure 4.3 displays the results for Hits in the two conditions separated by the order of testing. The analysis of variance, as with the Output measure, reveals a significant Order X Social Conditions interaction ($F = 42.74$, $df = 1, 112$, $p<.001$). From the figure it can be seen that subjects in both testing orders improve their performance on Day II, but those tested in the A → Ex-Comp order show relatively greater improvement. As was suggested above, with accuracy taken into account, the differences in performance level between the two conditions are more consistent on Days I and II, i.e. <u>both</u> days of testing show an advantage for the Ex-Comp condition.

Again, the clearest single effect is the facilitation of performance in the initial 5 minutes of testing for subjects in Ex-Comp. Figure 4.3 shows that A subjects improve fairly consistently throughout the testing sessions. However, even though those in Ex-Comp maintain a higher level of performance than A subjects, at no point during either day do they match the levels of performance attained during Period 1. This initial facilitation in the social conditions would appear to be uncontaminated by either the effects of order of testing or practice.

There were no main effects or interactions due to Sex.

Figure 4.3.

Mean Hits according to Social Conditions and Order over time.



4.1.5 Summary of Results and Discussion

The main points from the preceding section can be summarized as follows:

1. On Output the Ex-Comp condition showed better performance than A, but this effect was confined largely to the second day of testing.

2. On Output performance over time showed little fluctuation except for a marked reduction in performance level shown by the Ex-Comp condition between Periods 1 and 2 on both days.

3. Performance on the Omissions measure showed no relationship to the social conditions of testing, although subjects in the Ex-Comp → A order were consistently more accurate.

4. On the Hits measure the Ex-Comp condition showed better performance on both days of testing than the A condition.

5. On the Hits measure subjects tested in Ex-Comp demonstrated a drop in performance level after the first 5 minutes of testing relative to A.

The data offer evidence of facilitation of performance for subjects tested under explicit competition, although some qualification is necessary. Firstly there were indications that either asymmetrical transfer effects or sampling bias resulted in different performance strategies for accuracy between subjects tested in the separate orders: this may have influenced the results for both Output and Hits. Unfortunately, because of practice and order effects it is not possible to discern whether this was a genuine difference in ability between the two groups of subjects or was a performance strategy initially adopted in response to the social conditions. Clearly an experimental design is called for which does not necessitate counter-balancing the order of testing.

Secondly, although a main effect for social conditions was found for both Output and Hits, facilitation was most notable during the first 5 minutes of each testing session. It was suggested that this initial facilitation may have been due to positive stimulation (arousal) resulting from the social conditions. That the effect was not maintained, however, raises several questions. For example, it is not clear whether the reduction in performance level was due to subjects initially setting a standard of performance which could not be maintained or whether they quickly habituated to the social situation, reducing the initial positive impact of the group. Of these two alternatives the latter seems more probable, as it is unlikely that fatigue would have been a substantial factor in performance at this early stage in the session. Furthermore, most subjects stated that, although they were initially aware of the sounds and movement of other subjects, they quickly became used to this and hardly noticed these things in later parts of the session. Another possible explanation for this pattern is based on the comment made by several subjects that they were very aware of the competitive instructions when in the group and felt this to be an inhibiting influence. If subjects were conscious of not appearing 'too competitive' this may have served to inhibit performance in later periods of the test, particularly for those who perceived themselves to be the better performers in their group. Both a rapid habituation to the stress and a negative reaction to the competitive presentation of the task may have combined to reduce the initially facilitative effects of the social situation.

In addition to the primary aim of examining performance in these two conditions, a secondary purpose was to investigate the sensitivity of the cancellation task to social-evaluative stress. The task as employed here has a number of positive features for this type of research. Firstly, all subjects mastered it without difficulty. Secondly, most found it to be of little interest in itself and were, therefore, not likely to become overly involved with or intrinsically motivated to perform well on it. Furthermore, and most importantly, the ability to separate task performance into the three components of Output, Errors and Hits has advantages, as this can elucidate different strategies of task execution which may or may not be clearly reflected in measures of overall performance. The major disadvantage with the task is the marked tendency for subjects to improve performance with practice. However, some practice effect is noted with almost all tasks, and this may only be a critical drawback when the experimental design calls for a counter-balanced order of testing.

As mentioned, there were no significant effects obtained due to subjects' sex. Thus, the negative reactions to competitive situations which have been noted for females (see Section 2.3) were not found with this population on this task. Different results may have emerged had groups been mixed rather than single sexed or had the task been presented as one which was predominantly relevant for either males or females. However, sex differences per se were not of central concern in this experiment.

The results as a whole provide tentative support for a facilitative effect on performance in a group testing situation. However, a more ambitious experimental design, employing a larger number of subjects and independently manipulating instructions and social conditions, is needed before a more detailed and conclusive statement can be made regarding these findings.

4.2 EXPERIMENT II

4.2.1 Aims

Based on the findings in Experiment I the present study is designed to examine more closely the separate effects of group presence and ego-threatening instructions. An independent groups design is used in this case in order to eliminate the confounding effects due to practice on the cancellation task when order of testing is counter-balanced and to enable the manipulation of two variables, which in a counter-balanced design would be unwieldy. The main design incorporates alone (A), Ego-Threat (A-T), Implicit Competition (Im-Comp) and Explicit Competition (Ex-Comp). In addition to these situations an audience (Aud) condition is included, as 'mere presence' is also of interest. Since few experiments compare the effects of audiences and coactors in the same experiment, it was decided to do so here where the experimental task and conditions of testing would be constant.

In addition to changes in the experimental design, this experiment also uses a somewhat different task; subjects perform a running digit span (RDS) task simultaneously with vowel cancellation. Some of the advantages of such tasks have been discussed in Section 4.0. In addition, there was evidence from Experiment I that subjects might habituate fairly rapidly to the mild social stress used, suggesting that a more sensitive, demanding task might be advisable if performance differences are to be detected. Furthermore, several authors have found evidence that subjects are sometimes able to maintain high levels of performance on a primary task under conditions of mild environmental stress but suffer impairment on secondary or peripheral tasks (Bursill, 1958; Hockey, 1970). This suggests that, even though central performance processes may appear to be unaffected by environmental conditions, reserve capacity may be reduced. The dual task, with RDS serving as a peripheral source of information allows for the investigation of this possibility.

Based on the findings from Experiment I, the presence of others would be expected to facilitate performance on the cancellation task, with Ex-Comp probably resulting in the best performance. However, no predictions are made as to how the audience condition would compare with the two group situations. The addition of a secondary task raises several possibilities concerning predictions. Firstly, it is not clear how the addition of the RDS task might affect performance over time. The dual task is clearly more difficult than simple cancellation, but whether this would result in greater fatigue (leading to early decrement) or greater interest (resulting in sustained high performance) is not easily predictable. Secondly, although performance on RDS would be expected to show a decrement in the more stressful conditions, the reports which have noted such an effect on secondary tasks have used rather different types of environmental stress, such as noise (Hockey, 1970), and different secondary tasks (usually vigilance). However, apart from dual task studies, there is evidence that RDS is itself sensitive to arousal level, with high levels of arousal (noise) leading to restricted cue utilization and reduced RDS (Hamilton, Hockey and Rejman, 1977) while low arousal levels (sleep loss) may actually improve performance on this task (Hamilton, Wilkinson and Edwards, 1972). Both of these areas of research would suggest that RDS might show impairment under arousal changes induced by social-evaluative stress.

4.2.2 Task

The cancellation task used in Experiment I is employed in the same form as described in Section 4.1.2. The method of component performance analysis is also the same.

The RDS task consists of 44 tape-recorded lists of random digits, varying in length between 15 and 31 and presented at a rate of approximately 0.75 seconds per item. At the end of each list is a 15 second pause during which subjects stop crossing e's and write down the last 8 digits in the series (in the proper order) in the margins of their test sheets. The random arrangement of digits was constrained so that the last 8 digits contained no repetition and no more than 2 digits in either forward or backward sequence.

In order to encourage continuous attention the list length was varied and digits were spoken in a monotone voice which avoided changes in inflection such as would normally indicate the beginning or end of a list. Subjects were told they could guess if they wished and could resume the cancellation task as soon as they had finished writing the numbers. It was stressed that cancellation was the more important of the two tasks and that subjects should write the digits as quickly as possible and then continue with the primary task.

The total time of the task was about 23 minutes. As subjects were instructed to write the digits in the margins of their papers next to the line on which they were currently working, it was possible to divide the testing period into four blocks based on the position of the 11th, 22nd and 33rd digit list. The lists were arranged such that the session divided into four approximately equal quarters at these points. RDS scores were based on subjects terminal span (Waugh, 1960) i.e. the average number of consecutive digits recalled correctly from the end of the list up to the first error in recall.

4.2.3 Method

Subjects were 50 male and 50 female undergraduates of Durham University who received a small payment in exchange for participation. Subjects were randomly assigned to one of the five conditions of testing, although this procedure had to be modified at times in order to timetable groups. All groups in the present study contained at least one member of the opposite sex. Hence, sex differences, which were minimal in Experiment I, were examined in a context more likely to produce differential reactions (see Section 2.3.4).

All subjects initially reported to the experimenter's office and were taken to the experimental room, which was approximately 6 by 8 feet and equipped with individual desks. In group conditions the desks were arranged so that subjects faced each other, and the size of the room insured that all individuals were in close proximity. In the Aud condition the observers (experimenter plus one other) sat approximately 6 feet from and facing the subject and watched him throughout the session.

All subjects were read the task instructions and given a 2 minute practice trial, after which the work was examined to make sure the instructions had been understood. Following this either the neutral or ego-threatening instructions were read by the experimenter.

In all but the Aud condition the experimenter left the room immediately after the subjects began working and did not return until the experiment was over; a taped message to "stop" followed 15 seconds after the final digit list. In the Aud condition another postgraduate student was introduced to the subject before the practice trial. It was explained that the other student was interested in watching how the experiment was conducted. The observer left immediately after the task was completed.

Following the performance period, subjects were administered the EPI and the qAch, accompanied by the instructions used in Experiment I. Subjects were then debriefed and the experimental manipulations were discussed.

Summary of Experimental Conditions

Alone (A)	- alone with neutral instructions.
Ego-Threat (A-T)	- alone with ego-threatening instructions.
Audience (Aud)	- the experimenter and one observer present with neutral instructions.
Implicit Competition (Im-Comp)	- coacting group of 5 with neutral instructions.
Explicit Competition (Ex-Comp)	- coacting group of 5 with ego-threatening instructions.

4.2.4 Results

Because the audience condition is additional to the main design, the data can not be analysed using a typical factorial analysis of variance. Instead, the data were arranged in a manner analogous to a 2 X 5 analysis of variance with Sex and Social Conditions acting as factors and Planned Comparisons were conducted on the data. This method allows for comparisons between Aud and any of the other treatments within the same analysis. Comparisons of interest are as follows:

1. Alone (A + A-T) versus Group (Im-Comp + Ex-Comp).
2. Ego-threatening instructions (A-T + Ex-Comp) versus Neutral instructions (A + Im-Comp).
3. The interaction of 1. and 2.
4. Aud versus A.
5. Aud versus Im-Comp.

The data are collapsed over time periods in this experiment as preliminary analysis suggested that performance over time did not vary due to the social conditions. Figures 4.4 - 4.7 show the mean results plotted over periods for the four measures considered. As can be seen, performance in all conditions follows the same general pattern over time. In addition, sex differences over time were small and sporadic, and interpretations based on these isolated differences would more than likely have confused rather than clarified understanding of the central issues in this experiment.

Unlike the results for Experiment 1, no initial decrement is apparent on cancellation (see Figures 4.4 and 4.6), although this is probably due to differences in the task. Part of the rationale for incorporating the RDS task was to counteract habituation by providing a more demanding activity. This seems to have been a successful strategy as no decrement occurs in any of the social conditions; if anything there is fairly steady improvement.

Output

Table 4.3 below shows the means and standard deviations for Output according to the conditions of testing.

TABLE 4.3

Means and Standard Deviations for Output According to Social Conditions

	A	A-T	Aud	Im-Comp	Ex-Comp
Mean	878.4	969.4	990.7	1102.4	1000.4
S.D.	195.9	193.8	155.0	183.0	122.7

As can be seen both coaction conditions demonstrate a higher level of performance than either of the alone conditions. However, the instructions used appear to modify this general trend to some degree. Specifically, ego-threatening instructions are associated with a higher level of performance when subjects are tested alone (A-T) but in the group conditions subjects given these instructions show lower performance (Ex-Comp). Planned Comparisons reveal a significant difference between Alone and Group conditions ($F = 10.24$, $df = 1, 90$, $p<.01$) as well as a significant interaction for Social Conditions X Instructions ($F = 5.87$,

Figure 4.4.

Mean Output for Social Conditions over time

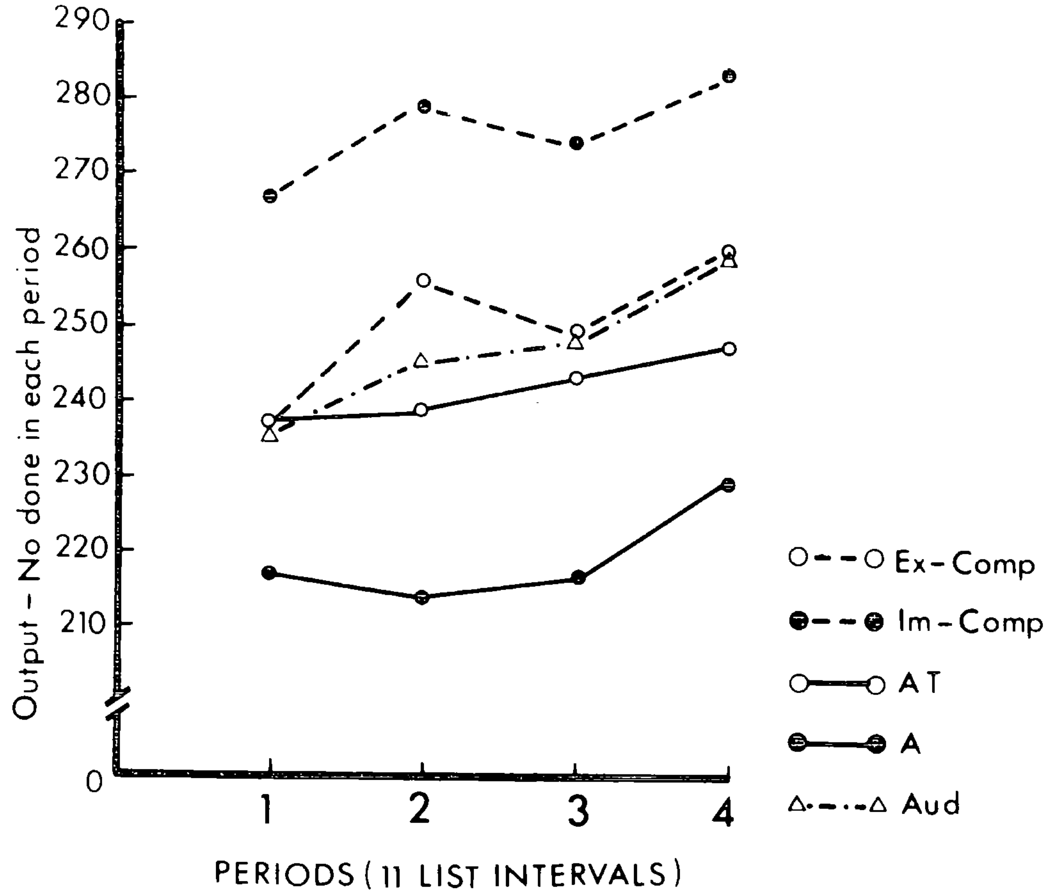


Figure 4.5.

Mean Omissions (per line) for Social Conditions over time

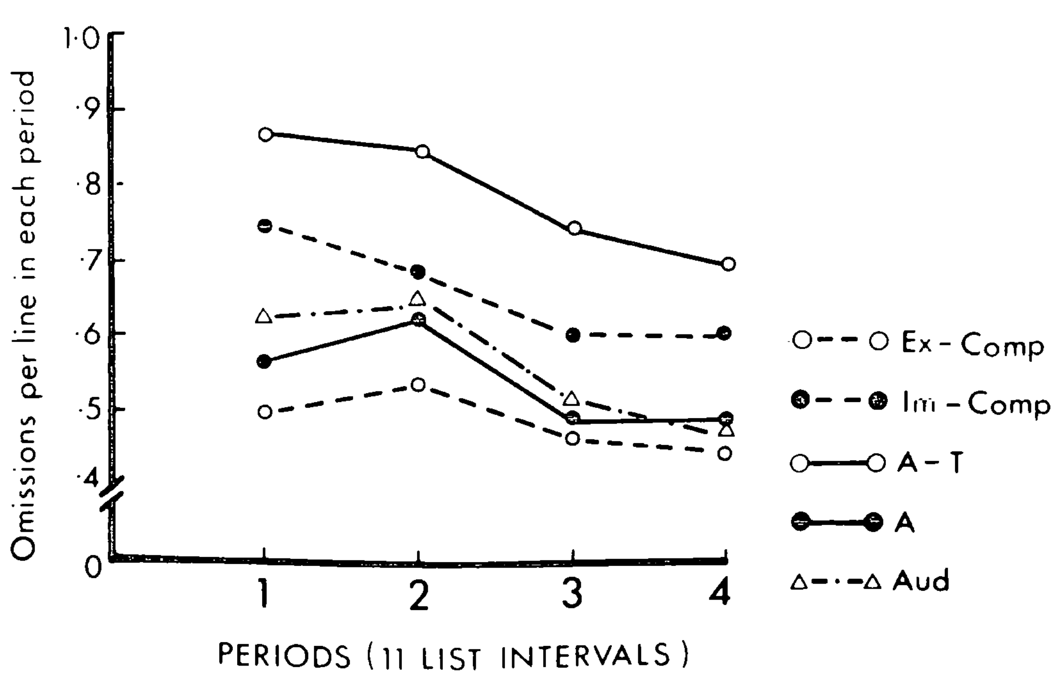


Figure 4.6.

Mean Hits for Social Conditions over time.

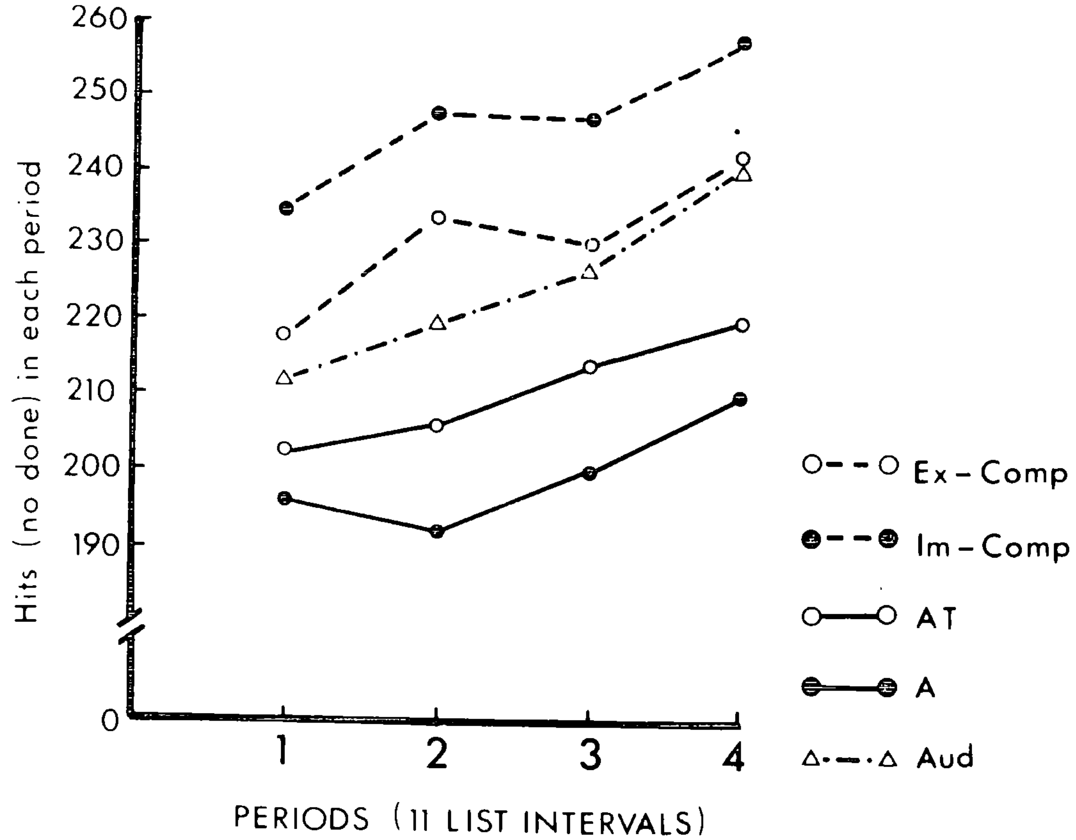
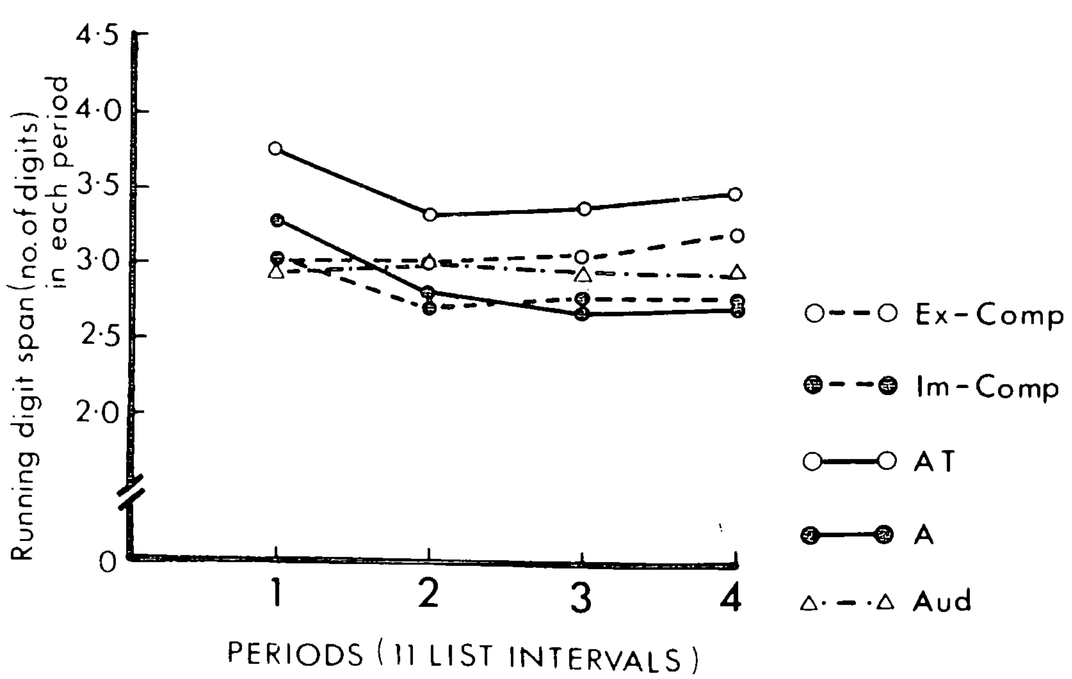


Figure 4.7.

Mean R.D.S. for Social Conditions over time.



df = 1, 90, $p<.025$). However, comparisons between the two group and two alone conditions are not significant. Thus, consistent with the results from Experiment I, subjects tested in groups performed at a higher level than subjects tested alone, although in this experiment the effect did not diminish over time. Ego-threatening instructions appear to reduce the difference between alone and group conditions; possible explanations for this effect will be discussed after data for the three performance measures have been presented.

The table also indicates that performance in Aud falls between that of subjects tested in alone conditions and those tested in groups. Individual Planned Comparisons show marginal significance for both A versus Aud ($F = 3.98$, df = 1,90, $p= .05$) and Aud versus Im-Comp ($F = 3.93$, df = 1, 90, $p = .05$).

No differences due to Sex were found to be significant on this or on any of the other performance measures. Hence, no further discussion of sex differences will be undertaken in this section.

Omissions

Table 4.4 shows the mean results for the Omissions measure in the five different conditions of testing.

TABLE 4.4

Means and Standard Deviations for Omissions (per line) According to Social Conditions

	A	A-T	Aud	Im-Comp	Ex-Comp
Mean	.52	.79	.56	.64	.49
S.D.	.43	.34	.33	.35	.29

There are two striking features about the pattern of means. Firstly, subjects in A-T make more errors of omission than subjects in any other condition. Secondly, as in Table 4.3 there is an apparent difference in the effect that ego-threatening instructions have on alone and grouped subjects, i.e. those tested in A-T have the highest error rate while those in Ex-Comp demonstrate the lowest. Consistent with these observations is a significant interaction for Social Conditions X Instructions ($F = 7.44$, df = 1, 90, $p<.01$). Thus, considering the results for Output and Omissions together, it would appear that different strategies for execution of the task were employed depending on the social environment and the instructional set. Subjects tested in Ex-Comp show a careful but slower strategy of performance, whereas those in A-T work quickly but miss more targets. It should be pointed out however, that on absolute level of Output there is little difference between these two conditions. The effect of instructions only becomes apparent when examined within the same social conditions, and, even then only A-T versus A is significantly different on the Omissions measure ($F = 6.12$, df = 1, 90, $p<.025$).

The F ratios for the Aud condition compared to both A and Im-Comp are less than 1. Thus, although speed showed some tendency to be influenced by the presence of others, this was not reflected to any marked degree in accuracy.

Hits

The mean results for the Hits measure can be found in Table 4.5 below.

TABLE 4.5

Means and Standard Deviations for Hits According to Social Conditions

	A	A-T	Aud	Im-Comp	Ex-Comp
Mean	804.6	845.8	900.0	985.1	921.3
S.D.	175.39	187.03	150.45	165.52	128.51

The same basic pattern is apparent here as in Table 4.3 (Output). However, the interaction of Instructions and Social Conditions is less marked and on the Hits measure did not reach significance. This could have been anticipated due to the speed/accuracy trade-off mentioned above; although subjects appeared to be using somewhat different performance strategies according to the social situation and instructions given, this was effectively masked in overall performance. The comparison for Group versus Alone is significant ($F = 12.59$; $df = 1, 90$; $p<.01$), based on the higher average performance level for subjects tested in the two group conditions relative to those tested alone.

For the Aud condition performance level on Hits is between that of grouped and alone subjects (as was the case with Output). However, individual comparisons with A and Im-Comp do not reach acceptable levels of significance. Considering that, of the three performance measures examined, only one (Output) shows even a marginally significant effect due to the presence of an audience, caution must be placed on any assertion that an audience had any effect on peformance of the cancellation task.

Running Digit Span (RDS)

Table 4.6 shows the mean results for RDS according to the conditions of testing.

TABLE 4.6

Means and Standard Deviations for RDS According to Social Conditions

	A	A-T	Aud	Im-Comp	Ex-Comp
Mean	2.87	3.59	3.0	2.87	3.07
S.D.	.82	.84	.88	.88	.69

An examination of the means suggests that there is little advantage on this measure to being tested in any one of the three social conditions (alone, audience or group). There does, however, appear to be some difference on RDS performance when subjects are given ego-threatening instructions, i.e. higher performance levels are evident in both A-T and Ex-Comp. Planned Comparisons show no significant effect for Alone versus Group, but the comparison for Ego-threatening versus Neutral instructions is significant ($F = 6.10$; df = 1, 90; $p<.025$). There is obviously no interaction between these two variables since the pattern is the same regardless of the social situation. However, only the comparison between A and A-T attains significance when the individual conditions are examined separately ($F = 7.56$; df = 1, 90, $p<.01$). The Aud condition is not significantly different from either A or Im-Comp.

In summary, it would appear that overall performance (Hits) is facilitated by group presence, but that particular strategies for task performance depend on the social situation and the instructions issued. The presence of observers had little effect on the performance of this task. There was some tendency for observed subjects to work faster, but accuracy of work as well as RDS performance was undifferentiated from subjects working alone.

It could be argued that A was the least and Ex-Comp the most stressful situation with A-T, Im-Comp, Aud falling between these two extremes, in that order. On this assumption, there is some evidence of an inverted U-shaped relationship between the social conditions and performance. Table 4.7 lists the social conditions along the (largely intuitive) continuum of stressfulness, and their respective means on all performance measures.

TABLE 4.7

Summary of Social Conditions and Mean Performance

	A	A-T	Aud	Im-Comp	Ex-Comp
Output	878.4	969.4	990.7	1102.4	1000.4
Omissions	.52	.79	.56	.64	.49
Hits	804.6	845.8	900.0	985.1	921.3
RDS	2.87	3.59	3.00	2.87	3.07

For both Output and Hits there is a tendency for low stress conditions to be associated with low levels of performance. Similarly, performance begins to show some decrement at the highest point along the continuum, Ex-Comp. However, the means for RDS (the measure reputed to be sensitive to arousal level) show a completely erratic pattern, while Omissions, not surprisingly, demonstrates a somewhat inverse relationship to Output.

An arousal continuum, at least ordered in this way, does not provide an adequate explanation of the results, particularly those indicating an interaction between the social situation and instructions. Explanations incorporating an inverted U-shaped relationship are probably more applicable to simpler tasks, either those which produce only one measure of performance or for which 'strategies' are not a prominent feature of task execution.

A more satisfactory interpretation of these data is that different instructions and social environments resulted in differences in the way in which attention was allocated to various task components. Ego-threatening instructions may have had the effect of making subjects more aware of their performance level, causing them to direct attention to those parts of the task for which self-evaluation was least ambiguous. For example, on vowel cancellation there is no obvious performance goal or standard against which to gauge individual competence. RDS, on the other hand, had a set goal of eight digits. As few subjects were ever able to remember all eight, they could clearly see that their score was falling below the set goal, constituting failure. This difference in the clarity of goals between the two tasks may have led A-T subjects to direct attention toward the secondary task, on which they were better able to evaluate their performance. On the other hand, Ex-Comp subjects had other individuals present with whom they could compare themselves at least on the amount of work completed.

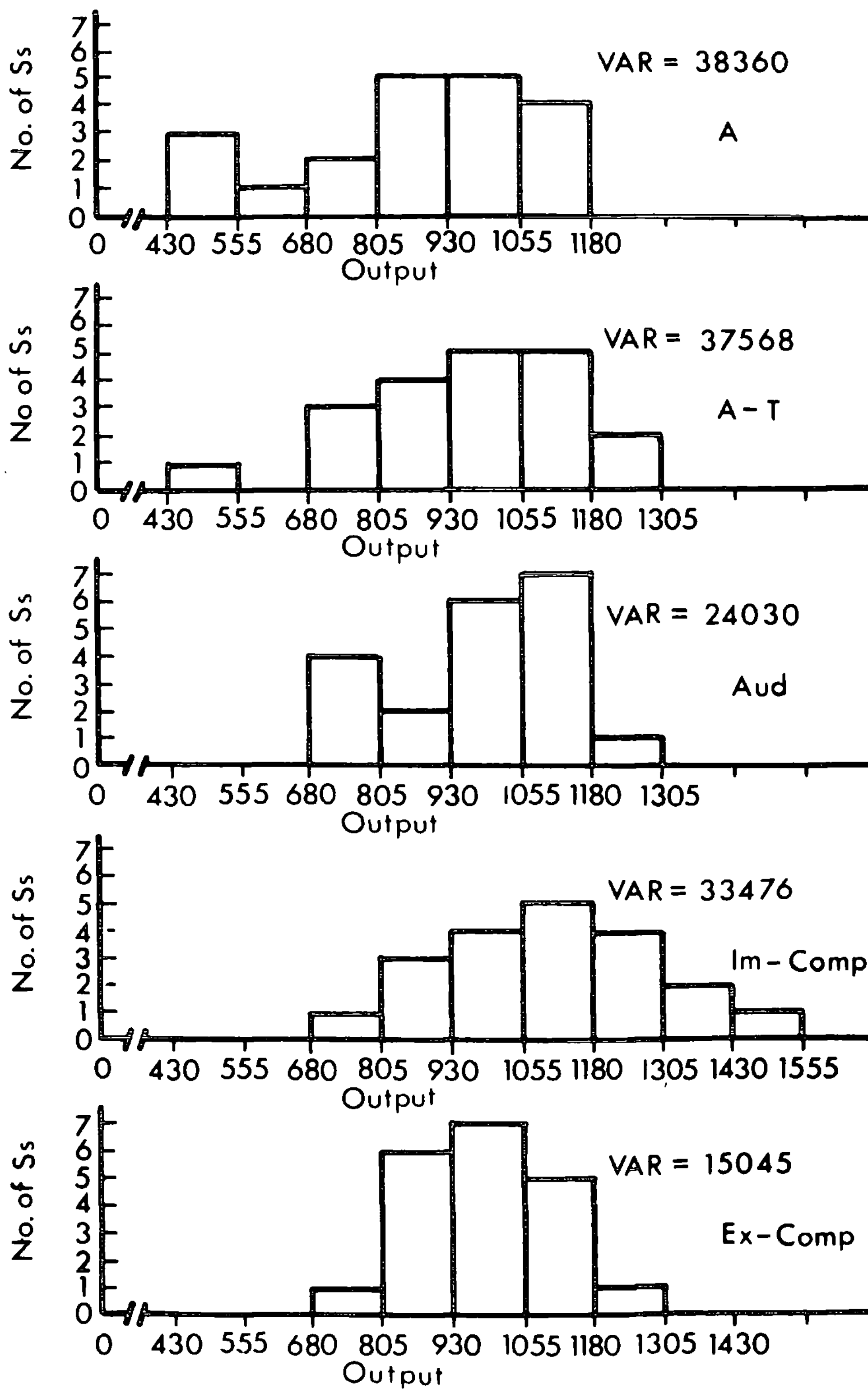
Some indirect evidence suggesting within group comparison comes from an analysis of the distribution of performance scores within conditions. Figure 4.8 shows histograms of the distribution of scores for Output in all conditions. The variances are significantly different between Ex-Comp and A-T ($F = 2.5$; $df = 19, 19$; $p < .06$), Ex-Comp and A ($F = 2.55$; $df = 19, 19$; $p < .05$) and the ratio for Ex-Comp and Im-Comp approaches significance ($F = 2.23$; $df = 19, 19$; $p < .1$).[4] From the figure it can be seen that in the Ex-Comp condition there is a greater uniformity of performance than in any other condition, an absence of both very low and very high scores and marked clustering of scores around the mean. Assuming that the ego-threatening instructions did make subjects more aware of their performance level, then it would appear that in the group condition this resulted in a tendency for individuals to pace their speed of performance with each other.

---

4. Heterogeneity of variance is a violation of the assumptions underlying analysis of variance and hence Planned Comparisons. However, Edwards (1968) has pointed out that the F test is relatively insensitive to heterogeneity of variance provided the number of subjects in each cell is the same and the sample fairly large, which was the case in this study.

Figure 4.8.

Histograms of the distribution of Output scores according to Social Conditions. *



* Intervals of 125 were chosen arbitrarily. The wide range of scores necessitated fairly large groupings in order to show clustering patterns.

4.2.5 Summary of Results and Discussion

The main findings from the present experiment can be summarized as follows:

1. On Output subjects who were tested in a group situation performed at a higher level than those who were tested alone.

2. On Output subjects tested in groups and given ego-threatening instructions performed relatively poorer than those given neutral instructions, while subjects tested alone showed the reverse pattern.

3. For Omissions, subjects tested in groups with ego-threatening instructions were more accurate than those given neutral instructions, while the reverse was true of subjects tested alone.

4. On Hits, subjects tested in groups showed a higher performance level than subjects tested alone.

5. RDS showed little sensitivity to the social conditions, but subjects who received ego-threatening instructions gave better performance than those receiving neutral instructions.

6. The audience condition was not significantly different from any of the other conditions of testing on any of the performance measures. However, performance on Output was marginally better for these subjects than for those tested in A and was marginally worse than for those tested in Im-Comp.

7. Subjects tested in the Ex-Comp condition showed less inter-subject variability on Output than those tested in other conditions.

In terms of overall performance there would appear to be an advantage to performing the task in a group situation. Not only was the averaged performance of the group conditions better for Output, but there was no marked decrement in accuracy. The Hits measure reflects this overall superiority most clearly. RDS performance, on the other hand, proved to be more sensitive to instructions than to the social situation.

The level of resultant arousal engendered by the social conditions was suggested as a possible framework for interpreting the results, and considering only the speed of performance (Output) the results roughly fit an inverted U-shaped curve. However this explanation fails to account for the different strategies for accuracy observed in the A-T and Ex-Comp conditions or the apparent shift in attention from the cancellation to the RDS task in A-T. Although level of arousal may provide a base for differences in performance between conditions, the specific performance patterns which emerged must be accounted for by other factors.

The results for RDS cast further doubt on an arousal interpretation since failure to find differential effects due to the experimental manipulations is inconsistent with other work attesting a relationship between arousal level and range of cue utilization. However, the RDS task as used in this experiment is somewhat different from those employed in other research. For example, since RDS was only given secondary status to begin with, subjects may have only given it minimal attention and would therefore be unlikely to show much decrement under any conditions. A comparison of the RDS data from this experiment with those from a study by Hockey (1973), in which subjects performed only RDS, offers some support for this notion (list lengths and rate of presentation are the same). Table 4.8 below shows the mean proportion of errors on RDS according to serial position. Position 1 corresponds to the last digit in the series, i.e. the most recent one heard. The table shows that subjects in the present study were in fact demonstrating a much higher error rate on digits early in the sequence than subjects who were doing only RDS.

TABLE 4.8

Proportion of Errors on RDS According to Serial Position

	7	6	5	4	3	2	1
A	.86	.82	.74	.61	.41	.25	.02
A-T	.82	.76	.63	.49	.26	.12	.01
Aud	.86	.81	.70	.58	.39	.22	.02
Im-Comp	.93	.87	.75	.62	.39	.22	.05
Ex-Comp	.89	.80	.69	.53	.34	.30	.03
Hockey (1973)	.91	.76	.67	.27	.17	.04	.00

The predicted relationships between arousal level and cue utilization may not apply when the error rate on the task is of such magnitude even in low stress conditions.

As an alternative explanation of the results, it was suggested that ego-threatening instructions made subjects more aware of their performance standard, and that this awareness led them to focus on that aspect of the task which provided the most tangible knowledge concerning performance level. For subjects tested alone RDS provided clearer performance feedback, since the task had a more easily defined goal. However, for those tested in a group the others present constituted a source of feedback for the primary task. This interpretation is consistent with Festinger's (1954) theory of social comparison processes. That subjects can switch attention from the designated primary to the secondary task has been demonstrated in cases where specific instructions to do so have been issued (Fisher, 1975).

One interesting finding which supports the above argument was the variance difference noted in the Ex-Comp condition. It was suggested that subjects in this situation may have sought information from others as to the 'correct' level of performance and then paced themselves accordingly, resulting in the lower inter-subject variance observed. Again this is consistent with Festinger's (1954) predictions regarding discrepant performance feedback in group situations, although 'pacing' as the term is applied in athletic contests is usually associated with better performance (e.g. Triplett, 1897). In this case there appeared to be a general gravitation toward the group mean. It should be borne in mind, however, that the Ex-Comp condition in this experiment was not an ordinary competition. There is no reason to suppose that the ego-threatening instructions actually inspired the subjects to try and beat each other; that the instructions led to heightened interest in the 'correct' performance standard seems more likely and would account for the absence in this condition of high as well as low scores.

The idea that pacing with the group mean may occur in group situations such as those employed here raises several interesting questions. Obviously, if subjects are pacing their performance with each other, predicting mean differences becomes more difficult. For example, if there are more good performers than bad in the group, the mean may go up. But, if the reverse is true, the mean may go down. Such an effect need not be related to either arousal level or the difficulty of the particular task. Considering the implications that a 'pacing' tendency may have for performance in group situations, this finding needs replication, as well as information about the specific conditions under which it is most likely to occur.

In terms of mere presence effects, there would appear to be a difference between performance in the presence of others as co-workers and as observers i.e. observers have little influence on the performance of the task. The results from this experiment corroborate other findings which have stressed the importance of the evaluative potential of observers (Paulus and Murdock, 1971; Cottrell et al, 1968). In the present experiment evaluation was 'played down' in the task instructions and the observers had no way of accurately assessing the subjects' level of performance while he was working. As a crude measure of subjects' discomfort due to the audience, subjects were asked to indicate on a scale from 1 to 3 whether or not they felt upset or distracted by the presence of the observers (1 being not bothered at all). The mean score for those so asked (N = 17) was 1.41, indicating that most subjects did not find the two observers very stressing. Those who did report some degree of discomfort stated that this was confined to the beginning of the experiment and once it had begun they virtually forgot about the audience. It is also true that the number of observers was only half the number of coactors, and that a larger audience might have been more effective. But, there are numerous reports indicating an audience effect with only the experimenter or one person present (e.g. Pessin, 1933; Putz, 1975; Matlin and Zajonc, 1968). Even so, it should be remembered that 'mere presence' was the issue under investigation. Although a larger audience may have been more stressing and led to more substantial performance differences, this is a different issue and relates to specific features of audiences rather than to their absence or presence.

Finally, as in Experiment I, no differences emerged due to subjects' sex. Therefore, it seems reasonably clear that with this subject population and on this task, males and females show no differential reactions to social-evaluative stress, even when competition is explicit and groups are comprised of both males and females.

4.3 EXPERIMENT III

4.3.1 Aims

The present experiment is designed to explore in more detail the attentional shifts noted under different social conditions in Experiment II on the cancellation/RDS task. It was suggested that ego-threatening instructions led to increased awareness of performance level in subjects and a shifting of attention to those aspects of the task which provided the most substantial performance feedback. There was also evidence that in a group situation feedback from others may have been the more readily available source of information and resulted in a tendency for individuals to pace themselves with the group mean. The present experiment is concerned with both of these possibilities i.e. shifting of attention and pacing.

The basic manipulation involves making one task parameter more public and accessible to evaluation than the others. Specifically, half of the subjects write their RDS answers while the other half report them orally. It is expected that those who give their answers orally in a group situation will focus more attention on the RDS task, since it would be more public and would therefore be the task parameter most likely to be chosen for comparison purposes. A shift of attention in the oral condition would be indicated by either reduced group variance on the RDS measure (pacing) or improved RDS performance accompanied by a decrement in speed or accuracy on the cancellation task (as in Experiment II).

There were two problems with the design of the present experiment. Firstly, because one of the conditions requires oral reporting, only one subject in the group can work at a time, meaning that performers must work in turns. Secondly, since the tasks are different in one detail (the method of reporting answers) it was not possible to make the conditions of testing exactly equivalent: for the oral task someone other than the subject must record the answers. There is also the possibility that one of the methods of reporting RDS might facilitate recall relative to the other, although this was not found in pilot investigations with the task. A repeated measures design which varied both Method of Reporting and Social Conditions for each subject would have required four testing sessions or else two long sessions in which both tasks were done on each day. Both of these designs were unfeasible because of (1) the difficulties in getting subjects to volunteer for four sessions and the possibility of boredom or habituation in later sessions and (2) the length of time involved for group sessions if subjects worked in turns on both tasks on each day.

In the design chosen subjects do only one task (oral or written) first with only the experimenter present and then in a group condition. The experimenter only condition (Day I) serves as a baseline and any effects due to the presence of others interacting with the method of reporting answers is determined by the analysis of difference scores between Days I and II. The Day I condition allows for the detection of any outstanding differences which might occur due to the different reporting methods, while the difference score itself should minimize effects due to small advantages or disadvantages. Any increases in Output on Day II (in groups) can not of course be interpreted as social facilitation because improvement would be expected on the cancellation task due to practice.

4.3.2 Task

The task is identical to that used in Experiment II but includes only 20 digit lists. The performance period was shortened in order to counteract possible boredom in the more lengthy group sessions. In all, subjects worked for approximately 10 minutes, the group sessions totalling just over 30 minutes. A shorter test was considered reliable based on results from Experiment II, which indicated little fluctuation in RDS performance after a short practice period.

Task instructions were approximately the same as those in Experiment II, except that subjects were told not to guess at the digits if they were unable to remember them all. In this way failure would be obvious. Otherwise, those observing would have difficulty evaluating the others' performance unless they were themselves able to recall all the numbers perfectly.

In the written condition subjects wrote their RDS answers in the margin of their papers (as in Experiment II), while in the oral condition the experimenter recorded the subjects' spoken answers. The possibility of having individuals do the task in isolation and report their answers to a tape-recorder was considered. However, being recorded is stressful for many people (Putz, 1975; Droppleman and McNair, 1971) and is probably more unnatural and obtrusive than the experimenter's presence. No audience effect due to the experimenter's presence was expected, based on the results of Experiment II (which involved two observers) and on verbal feedback from subjects which indicated that the experimenter was not viewed as threatening. However, the possibility of such an effect was not ruled out, particularly since the task was more open to immediate evaluation, and this was considered in the analysis of Day I data.

4.3.3 Method

Subjects were 12 male and 12 female undergraduates who were engaged in the same way as in the previous two experiments. All reported to the experimenter's office and were taken to the experimental room, the children's play room, which was equipped with several individual desks and tables. Subjects were randomly assigned to either the written or oral condition and worked with only the experimenter present on Day I.

Before receiving the task instructions, subjects were told that they would be doing an unusual task which required them to do two things at once. The RDS task was described as a subtest of the Wechsler-Bellevue Intelligence Scale, and the cancellation task was presented as a proof-reading test which was related to reading and vocabulary skills. Therefore, the social conditions could be described as E-T (Day I) and Ex-Comp (Day II), interest being in the interaction of others' presence with public feedback (F) on Day II. Since some subjects in the first experiment reported that the ego-threatening instructions were too obvious and contrived, those used in the present study were designed to provide more subtle evaluative cues. Comparison with others was not directly stressed and no interest was shown in subjects' personal level of performance. It was considered that with a university population an attested relationship between the task and intelligence test measures would be sufficient to arouse interest in performance level.

After this brief explanation of the task subjects were given the task instructions followed by 2 minutes of practice and then the test period. Seven to ten days later subjects reported in groups of 3 to the experimental room for the retest. A brief review of the Day I instructions and description of the task was given, and it was explained that, because such tests are thought to be influenced by subtle environmental features such as the number of people present, several subjects had been asked to report together. This was in fact the experimental hypothesis, although no information was provided concerning the direction in which change was expected.

The experimenter sat in the same place in relation to the subject on both days, which was at another table about 5 feet away. She avoided looking at the subject and appeared to be absorbed in either recording the answers or attending to other business. Waiting subjects sat facing the performer and about 6 feet from him. When one individual finished, he and the next subject simply changed places. The order in which subjects were tested on Day II was prearranged to maximize the possibility of subjects pacing with the group mean. Based on Day I data the middle performer on RDS was tested first, setting a pace which was higher for one subject but lower for the other.[5] It was recognized that the first subject to work would be in a unique position, having no others with which to compare his performance, and also that waiting different lengths of time before the test may have differential effects. Therefore, a record was kept of the order in which individuals worked.

Subjects were assigned to individual groups largely on the basis of timetabling demands. Since it was felt that evaluation stress might be greater in this experiment than in the previous two, due to subjects observing one another and the publicness of the task in the oral condition, all groups were single sexed.

Summary of Conditions

Day I

E-T	- Experimenter only, ego-threatening instructions, written RDS.
E-T (F)	- Experimenter only, ego-threatening instructions, oral RDS.

Day II

Ex-Comp	- Group of 3, ego-threatening instructions, written RDS.
Ex-Comp (F)	- Group of 3, ego-threatening instructions, oral RDS, i.e. public feedback.

4.3.4 Results

The data from Days I and II are analysed separately in 2 X 2

---

5. Obviously, pacing could be investigated by beginning with the best performer (setting a high standard) or the worst (setting a low standard). But for the sake of consistency and comparability with Experiment II, pacing with the group average was encouraged.

analyses of variance, the two factors being Mode of Response and Sex. The Day I data are intended to reveal any differences in performance suggesting that either: (1) one method of reporting RDS answers was easier or (2) the experimenter's presence had a different effect when she was involved with recording the subjects' answers. In analysing Day II data difference (Day II - Day I) rather than absolute scores are used, permitting a direct comparison of change between the two conditions in response to the addition of other subjects, while minimizing any small differences which may exist between the conditions due to either sampling factors or the different modes of response.

Day I

The analysis of Day I data reveals no significant main effects or interactions. The failure to obtain any effects either on the cancellation or the RDS task suggests that the different methods of reporting are equivalent in difficulty. Also, although the experimenter might have been perceived as a greater evaluation threat when recording the subjects' answers, this did not affect actual performance on any of the measures considered.

Day II

There are no significant effects on any of the Day II measures due to the different modes of response. There is a significant main effect of Sex ($F = 5.0$; $df = 1, 20$; $p < .05$) on the Omissions measure due to females reducing their error rate more than males. This difference in relative accuracy is independent of the method of reporting RDS. Also, it is not paralleled by any differential change on any of the other performance measures. However, the Day I data show an initial tendency for males to make fewer errors of omission (mean for males = .64, mean for females = .88), suggesting that some sampling bias due to the small number of subjects in the experiment may have contributed to this finding.

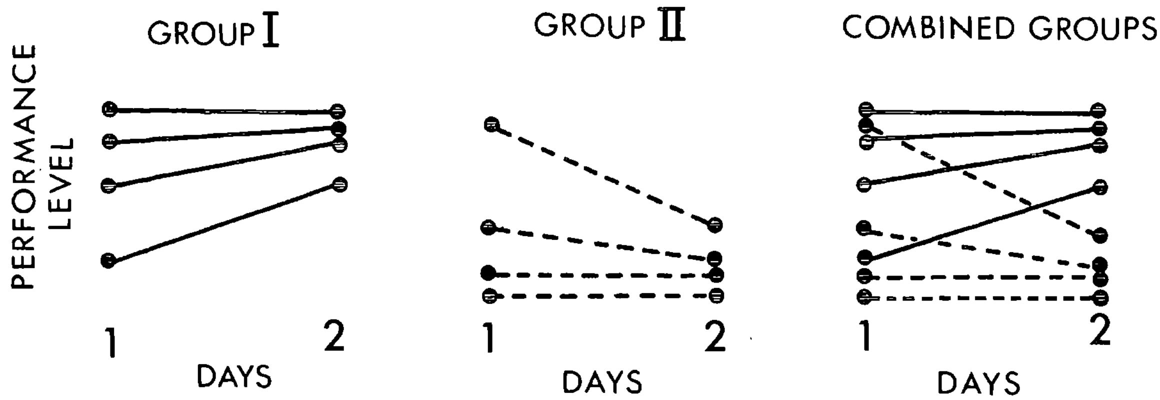
Since the analysis of variance showed no significant effects on any of the measures due to the type of reporting (publicness of the task), it can be assumed that no general facilitation or impairment occurred. The hypothesized shift in attention was not detected, at least by this method of analysis.

However, it was also suggested that pacing might occur, and if so, that this would be reflected in a reduced variance for the Ex-Comp(F) condition, rather than in mean performance differences between the two conditions. But, comparisons for Output, Hits and RDS on Day II show that the variance reductions in the oral and written conditions are of the same magnitude. Both showed some reduction in variance on Day II compared to Day I, but this would be expected due to a natural regression to the mean, as initially good performers would show progressively less improvement with practice than would initially poor performers.

However, a closer inspection of the raw data suggested that the variance computed between subjects might be masking the effect of within group pacing. Depending on the range of abilities in each separate group, a reduction of within group variance need not result in greater between subject variance for the whole condition. Such an effect is presented diagrammatically in Figure 4.9 below. Although the example is extreme, it can be seen that when the hypothetical groups are combined, between subject variance is not changed between Day I and II. However, when examined separately, within group variability is reduced.

FIGURE 4.9

DIAGRAMMATIC REPRESENTATION OF
DIFFERENT WITHIN GROUP PACING PATTERNS



Days I and II represent a test and a retest on the same task. Both Groups I and II show reduced intersubject variability on Day II, but when the groups are combined, no variance change is apparent.

Therefore, to further explore the possibility of pacing, performance scores for Days I and II are examined in a correlation matrix, which compares individuals' Day I scores with their change score based on Day II performance. Pacing would be indicated if initially good performers show either little improvement or even decrement on Day II, while poor performers demonstrate the opposite tendency. This would result in a negative correlation between Day I scores and the change scores. Naturally regression toward the mean would be reflected in this statistic also, but the magnitude of the correlations should allow comparison between the two different conditions of performance and would indicate if any effect beyond the normal regression is taking place. The correlational method, as used here, has some built-in imprecision since subjects in the middle performance ranges may set either a higher or lower performance goal, depending on the abilities of other group members. However, pacing should be more marked at the extremes of ability, and the correlation should reflect any general tendency for performance scores to gravitate toward the group mean. Also, in the present study, the medium ability subjects performed first, and therefore, should not have been able to adjust their pace.

Table 4.9 (a) and (b) gives the results of Pearson correlations on the four measures of performance. Day I refers to raw data scores, while Change (Day II - Day I) is based on the difference between scores on the two days. The comparisons involving change scores for Omissions are inevitably difficult to conceptualize since improvement is indicated by a negative value, while improvement on other measures results in positive difference scores. However, in the tables a negative correlation may, for all measures, be interpreted as a relatively greater improvement from Day I to Day II for subjects performing poorly on Day I. Parts 1 and 2 in the table show the internal consistency between measures i.e. the relationships between measures on Day I (Part 1) and between the different change scores on Day II (Part 2). Part 3 indicates the relationships between level of performance on Day I and change from Day I to Day II for all measures.

As can be seen, in the Ex-Comp condition there is a tendency for regression toward the mean between Days I and II (Part 3), which is reflected in the low negative correlations for Output, Hits and RDS. However, the trend is much stronger in the Ex-Comp(F) condition and is significant for RDS and, surprisingly, Output. Scatterplots of these four relationships are presented in Figures 4.10 (a),(b),(c) and (d).

In addition to pacing, there is some evidence of shifts in attention based on these correlational data. Firstly, there is a stronger positive relationship in the Ex-Comp(F) condition between change in Output and change in Omissions (Part 2), suggesting a more pronounced speed/accuracy trade-off than in the Ex-Comp condition. Furthermore, subjects in the

TABLE 4.9

Pearson Correlation Matrix for Day I and Change Data; Output, Omissions, Hits and RDS

(a) Ex-Comp

		Day I Output	Day I Omissions	Day I Hits	Day I RDS	Change (Day II - Day I) Output	Change (Day II - Day I) Omissions	Change (Day II - Day I) Hits	Change (Day II - Day I) RDS
Day I	Output								
	Omissions	12							
	Hits	91**	-22						
	RDS	-03	-35	-14					
Change (II - I)	Output	-36	-15	-43	-08				
	Omissions	-18	-63*	10	38	25			
	Hits	-26	-39	-44	-26	89**	-20		
	RDS	17	18	22	-37	35	37	24	

+

1	
3	2

+The three parts of the table are referred to in the text

(b) Ex-Comp (F)

		Day I Output	Day I Omissions	Day I Hits	Day I RDS	Change (Day II - Day I) Output	Change (Day II - Day I) Omissions	Change (Day II - Day I) Hits	Change (Day II - Day I) RDS
Day I	Output								
	Omissions	-02							
	Hits	92**	-41						
	RDS	34	-13	37					
Change (II - I)	Output	-74**	09	-64*	-11				
	Omissions	-45	-67*	-11	-27	50			
	Hits	-58*	-09	-58*	05	91**	14		
	RDS	-09	14	-04	-69*	-38	21	-56	

Correlations which are discussed in the text are underlined in the table.

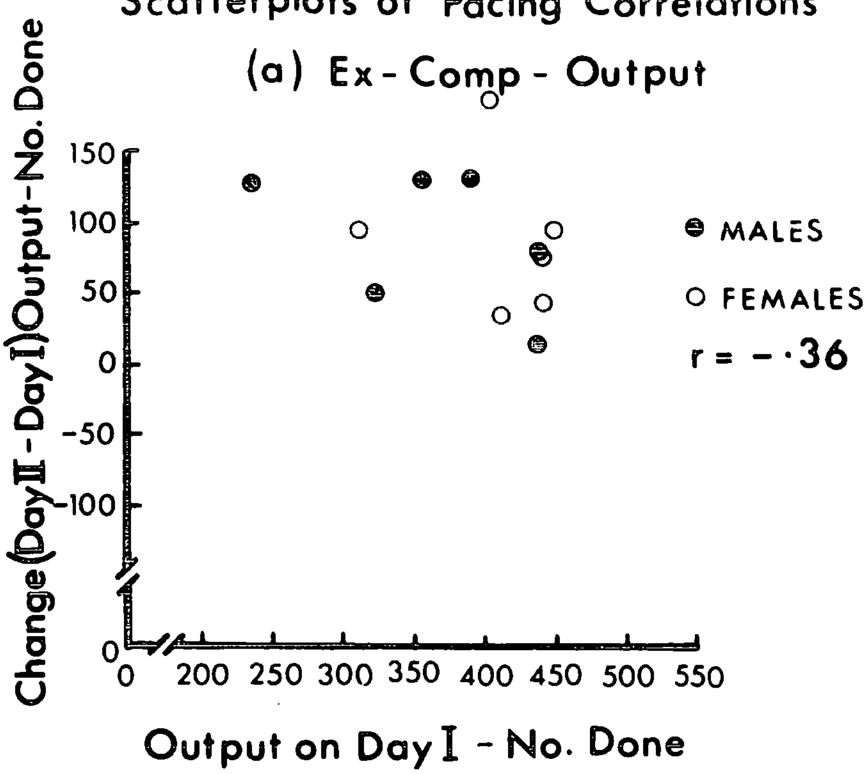
Levels of Significance, N = 12: *p<.05, **p <.01.

Ex-Comp condition show positive relationships between change in RDS and Output and change in RDS and Hits (Part 2). This pattern is reversed in the Ex-Comp (F) condition, suggesting that as one task assumed more importance, attention was directed away from the other task. Thus there is evidence here that making the task public leads to pacing and to shifts in attention. Logically, in order for pacing to occur, certain changes would need to take place in the manner in which the task was being executed, resulting in shifts in attention and changes in performance strategy.

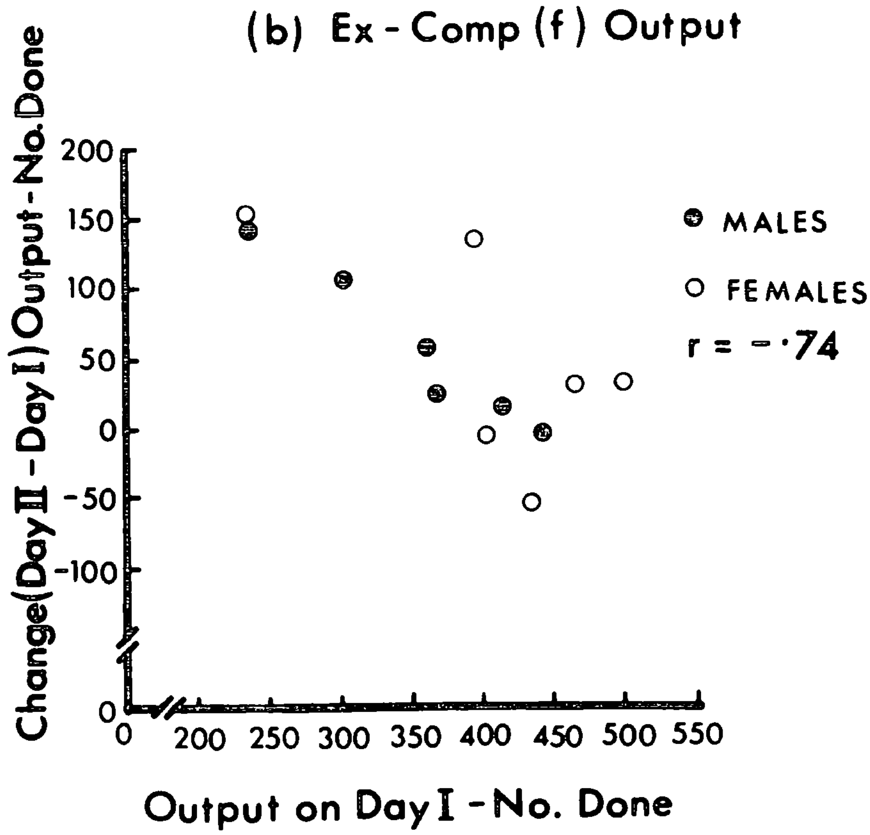
Figure 4.10.

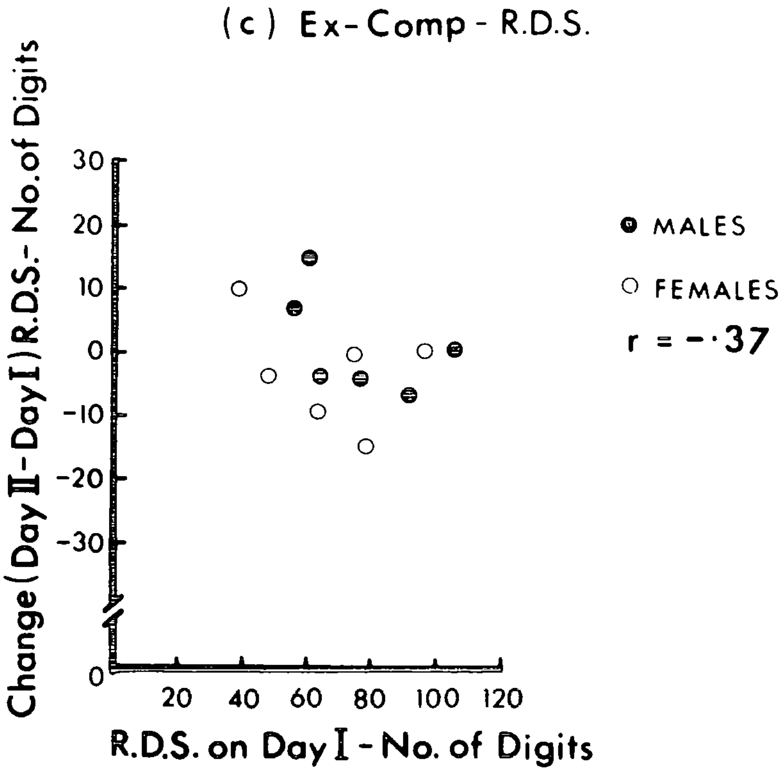
Scatterplots of Pacing Correlations

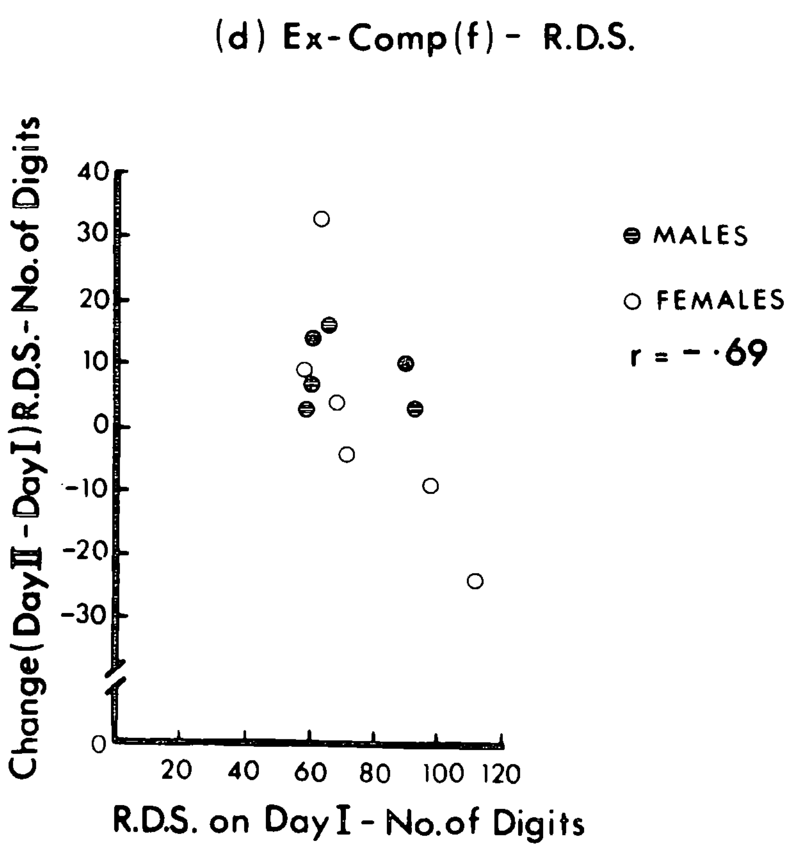
(a) Ex - Comp - Output



(b) Ex - Comp (f) Output




(c) Ex-Comp - R.D.S.
Change (Day II - Day I) R.D.S.- No. of Digits
30
20
10
0
-10
-20
-30
0
20 40 60 80 100 120
R.D.S. on Day I - No. of Digits
MALES
FEMALES
r = -·37


(d) Ex-Comp (f) - R.D.S.
Change (Day II - Day I) R.D.S.- No. of Digits
40
30
20
10
0
-10
-20
-30
0
20 40 60 80 100 120
R.D.S. on Day I - No. of Digits
MALES
FEMALES
r = -·69

4.3.5 Summary of Results and Discussion

Results from the present experiment can be summarized as follows:

1. A correlational analysis indicated that subjects in the Ex-Comp(F) condition of testing were pacing their performance with other group members such that good performers improved little and poorer performers improved a lot. This tendency was apparent on both RDS (as predicted) and Output.

2. There was some indirect evidence that pacing might be associated with attentional shifts from one aspect of the task to another. Shifts in attention were indicated by negative correlations between change scores for Output/RDS for subjects in the Ex-Comp(F) condition and a relatively stronger positive relationship between Output/Omissions change scores for subjects tested in the Ex-Comp(F) condition compared with those who wrote their answers.

3. Males in both conditions of testing improved their Omissions scores less than females (on Day II). This may reflect a general tendency for females to react less negatively to the presence of other subjects, although this finding does not fit well with the rest of the data and is in conflict with results from previous experiments. It is therefore considered likely that this result is due to a sampling artefact.

This experiment, unlike the previous two, did not look at overall facilitation or impairment of performance due to others' presence. In this case the task itself was manipulated so that in one condition the subjects' performance was more public than in the other. In some ways this might be considered to be a manipulation of the evaluation potential of the co-workers. Usually situations which are believed to possess higher degrees of potential evaluation are credited with having higher arousal value, and performance is expected to reflect this. If the Ex-Comp(F) condition used in this experiment was in fact higher in evaluation potential, this did not correspond to any mean performance changes.

The favoured explanation of these data is that the public situation made subjects more aware of their performance level and led them to use the performance of other group members as a standard for personal comparison. Again it is not clear whether the resultant pacing was intentional or not, although it is unlikely that subjects who were initially good performers would purposely suppress their performance level in response to this mild type of stress. The pacing interpretation is further supported by the data from Experiment II in which a smaller variance was noted in the more evaluative of the two group situations.

Although it was expected that the publicness of the RDS task in the Ex-Comp(F) condition would result in pacing on this measure, the effect was not expected to generalize to other measures. However, Output was also available for evaluation, at least in terms of the amount of work completed. It would appear that making the task public on one parameter need not produce effects specific to that task component. Once subjects have been made self-conscious of their performance level this may increase interest in performance level in general for all those task parameters for which some evaluation is possible.

The correlation approach used here seems a promising method for examining patterns of individual responding in social situations, which could be used in conjunction with analyses of mean differences. Of course, with sufficient subject numbers, initial level of performance could be examined as a separate factor in a factorial design. But even this more elegant statistical method would not enable comparisons of the degree of pacing between different conditions in the same way that the correlation statistic does.

The primary problem encountered in the present experiment was the manipulation of the publicness of the task. Although the task was felt to have many advantages in terms of sensitivity and the possibilities for examining attentional shifts, the publicness manipulation resulted in several methodological problems. Even though the two tasks proved to be effectively equivalent, the group testing conditions were somewhat unwieldy. The procedure of having subjects perform in turns was not judged to be ideal, considering the length of the session and the monotonous nature of the task. No systematic differences were observed due to the order in which individual subjects performed on Day II. Still, it is felt that further investigations of the 'publicness' variable might benefit from a more easily manipulated task.

4.4 OVERVIEW

The preceding experiments have looked at several varieties of non-interactive social situations in an attempt to identify the ways in which basic performance processes are affected by the social environment and to what degree the different situations under consideration are similar in their effects. In general, the results would seem to indicate that performance, especially in terms of speed, is facilitated when subjects work on the task in coacting groups. However, the more interesting finding concerns the ways in which group and individual performance can be modified by the introduction of evaluative stimuli. The effect of evaluative cues seems to be fairly general and not highly dependent on the type of manipulation employed. For example, in Experiment II subjects were given specific ego-threatening instructions, while in Experiment III evaluation was emphasized by making subjects' performance public.

The favoured interpretation of these data was that the presence of evaluative conditions induced subjects to become more interested in their own personal level of performance. It is doubtful that subjects actually felt stressed in these situations or that they were worried they might not perform as well as the other subjects. But the devices used in trying to create evaluation stress presumably had enough impact to alert subjects to the fact that their performance level was a matter of importance,leading them to seek information about the appropriate standard from whatever sources were available in the environment. When other subjects were present to supply this information, the result was an absence of extreme scores or, more specifically, a clustering of scores about the group mean.

It was suggested that this clustering tendency was the result of group members pacing their performance with others in the group.

Evidence for pacing came from two different sources: (1) the relatively small variance noted in Experiment II for the Ex-Comp subjects and (2) the correlational data from Experiment III. The latter present the stronger argument since the data show clearly that individuals were adjusting their performance styles so as to bring them nearer the group mean in level of performance. The fact that good performers in some cases showed a decrement suggests that subjects were not overly concerned about being the best performer; rather that they were willing to adopt that standard of performance which seemed appropriate or acceptable.

Although this interpretation may seem rather too well developed based on the limited data available, it is in keeping with other literature and theory. Theories of objective self-awareness stress the importance of any environmental stimulus which may lead the subject to focus on himself, and describe the effects that this increased self-awareness can have on performance (see Section 1.2.1). In addition, Festinger's (1954) theory of Social Comparison Processes predicts that, particularly in situations where the correct form or standard of behaviour is unknown, individuals will seek relevant information from available others. The task used in these experiments is conducive to such information seeking since it is more or less self-paced and, apart from the rather high target set for RDS, offers no cues which would indicate the appropriate standard of performance.

If the argument outlined above does adequately account for these data, then it is clear that individual differences (initial ability) determine to a large part whether overall facilitation or impairment of performance will occur. Mean effects may simply be a by-product of more subtle changes in social behaviour and performance strategies. Although in the specific situations used here, group presence tended to facilitate performance, it is easy to envisage cases in which the reverse might take place. For example, in Experiment II subjects tested in Explicit Competition showed more accurate performance on cancellation than those in Implicit Competition. This served to accentuate the mean difference between neutral and ego-threatening conditions on Output (due to a speed/accuracy trade-off), while differences were less striking on Hits. The drive theory of social facilitation would predict greater facilitation of speed on this routine task, not accuracy (e.g. Allport, 1920). Presumably the Explicit Competition condition possessed greater potential evaluation than Implicit Competition, but this did not lead to facilitation of either Output or Hits relative to the more neutral group situation. In this case it is likely that the different performance strategies adopted within conditions modified to some degree the pattern and size of mean performance differences between conditions.

It is likely that this explanation can account for at least some of the discrepant findings in the literature on social facilitation, particularly in light of the lack of attention which has been given to controlling evaluative stimuli such as task publicness and to examining performance strategies. In any event, it is clear that whenever there is a marked tendency for subjects to pace their performance with each other, this may alter or obscure overall facilitative or inhibitory effects due to the mere presence of others.

A pacing tendency would of course be of great importance in understanding and predicting the effects of others' presence on individual performance, provided these findings can be applied to other populations, settings and tasks. However, it should be borne in mind that the population used in these experiments is selective and that the experimental situation itself is atypical. The impact of the evaluative stimuli employed here may not be the same with a population less sensitive to performance situations. In addition, since subjects knew they were participating in an experiment, this may have encouraged pacing. In other words, in trying to comply with the experimental hypothesis, subjects may have been influenced to a greater degree by the behaviour of others than might have been the case in a natural situation.

Concerning the effects of audience presence on performance, there was little evidence from this series of experiments to indicate that a small audience of one or two observers had any impact on performance whatever, although there was some suggestion of a facilitative effect on Output when two observers were present. Post-experimental interviews with subjects supported the experimental findings. Subjects for the most part reported little awareness or discomfort in the presence of the audience. On the contrary, several subjects in Experiment III reported feeling reassured by the presence of the experimenter.

A few points from the above discussion have implications for further research. Firstly, the pacing tendency needs replication using the correlational method of analysis. The independent groups design used in Experiment II is unsatisfactory for two reasons: (1) the variance need not always reflect within group pacing, depending on the distribution of abilities in each separate group and (2) even when a relatively small variance is observed, such as in Experiment II, it can only be supposed that this is due to pacing and is not simply attributable to a sampling artefact. The correlation statistic as used in Experiment III, on the other hand, considers the individual subject's performance in the non-social situation in relation to changes in individual performance in the social situation. Thus, the pattern of performance changes is clear, and pacing can be more easily detected.

A second point needing clarification concerns the generality of results. Apart from the obvious problems centered around generalizing from a university sample and laboratory based findings, there are other more subtle issues. For example, would pacing be as likely to occur in a group composed of subjects who knew each other and had a firm idea of each member's relative ability? Festinger (1954) predicts that only individuals judged as comparable will be used for comparison purposes. Even when performing a novel task, comparative performance information may not be useful if the group members are already knowledgeable about differences in each other's general ability. Pacing may be a phenomenon largely confined to ad hoc groups of strangers.

A third issue needing further attention concerns the role of task variables in providing performance feedback. Experiment III used the Oral/Written manipulation, hypothesizing that oral reporting of RDS answers would accentuate performance comparison. As predicted there was evidence of pacing on this measure. However, a negative correlation of equal magnitude was obtained for Output in this condition. It was suggested that making the task public on one measure had the effect of increasing self-evaluation on all measures available for comparison. However, the fact that part of the task was public may actually have been irrelevant; the important factor may have been that subjects were speaking aloud. The act of speaking before others may have led to increased self-consciousness generally (objective self-awareness), with increased interest in performance level as a by-product. Therefore, the specific roles of performance information and performance publicness need more detailed investigation.

CHAPTER V

EXPERIMENTS IV AND V

5.0 INTRODUCTION

The two experiments reported in this chapter employ 5th and 6th form school children. The specific social conditions under investigation are Implicit and Explicit Competition. The effects of instructional manipulations and performance publicness on pacing tendencies are also a primary interest.

A sample of school children is used in order to test the generality of previous findings concerning the above social and task variables. Specifically, school children should be less restricted than university students in range of ability (which might influence pacing) and personality scores. The narrow range of scores on both subscales of the Argyle-Robinson test reinforced earlier concerns that a university population might be a self-selected group of people and also made the interpretation of personality data difficult. Sex differences as well, which did not emerge in the previous three experiments, might have been minimized due to the natural selection in a university sample of females who are good competitors.

The school children are between 16 to 18 years of age, and therefore reasonably comparable with first year university students in maturity. The tasks are similar to those used with the university sample in terms of difficulty. Experiment IV employs a version of the cancellation/RDS task, in order to facilitate comparisons between the two subject populations. However, subjects in Experiment V performed two cognitive tasks which varied in difficulty, one being routine transformation and the other a test of syllogistic reasoning.

The task was changed in this latter study partly to test the generality of effects across tasks and also because an activity was required on which publicness of performance could be more easily manipulated. In addition, since the main point of interest was the effect of public feedback on the most overt performance measure, there was little advantage in using the dual task which provides several different performance measures and, as a result, requires quite complicated data analysis. Finally, since all subjects worked on both tasks, the effects of task difficulty could be investigated within the same design.

To further enable comparison between these two and the first series of experiments, three of the same personality tests were used (EPI, TAS and ASI), with modifications appropriate for a school population.

Testing took place in several different comprehensive schools in County Durham. The subjects were volunteers who had been addressed in large groups in a request for their participation. Although the response was nearly 100 percent under these conditions, at times subjects had to be randomly omitted because the school had set a limit on the number of pupils that could be tested. In addition, subjects who were absent on the day of the test were deleted from the sample, as school staff were not receptive to the rescheduling of absentees.

5.1 EXPERIMENT IV

5.1.1 Aims

This experiment is concerned specifically with the effects of Experimenter Only (E), Implicit Competition (Im-Comp) and Explicit Competition (Ex-Comp) on cancellation/RDS performance. In addition to examining the overall effects of these social conditions with a different subject population, patterns of performance change within conditions are investigated using a correlational analysis (as in Experiment III). The specific effects of neutral and ego-threatening instructions on pacing tendencies is of central interest.

In keeping with previous findings, it was expected that pacing would occur on the Output measure in the Ex-Comp situation as opposed to either E or Im-Comp, although the magnitude of the effect with a different population was less predictable. No effects were expected for RDS performance since previously this measure had only shown differences when subjects were tested either alone and with ego-threatening instructions or when RDS performance was made public; neither of these situations was employed in the present experiment. For this reason the RDS task was simplified and incorporated mainly to make the primary task more difficult, although it also provided a check for any reduction in peripheral information processing. Based on Experiments I and II it was expected that both Im-Comp and Ex-Comp would result in facilitation of speed (Output) and overall performance (Hits) compared to E, but no predictions were made concerning mean performance differences between the two group conditions, apart from a more pronounced pacing tendency in Ex-Comp.

5.1.2 Task

The cancellation task was the same as that employed in previous experiments. However, for RDS subjects were instructed to listen to all the numbers but to write down only the third digit back in the series. This procedure reduced the variable amount of time subjects spent in actually writing the digits and deliberating over ones they could not recall. The third digit back was chosen as this was the point which marked the most difficult but still efficient task performance.[1] Although simplifying the task probably reduced its sensitivity, the third digit in the series should most clearly reflect any changes in the range of cue utilization; recalling the fourth digit in the sequence produces highly inefficient performance even under non-stressful conditions and recalling the second digit is not sufficiently taxing.

The test was composed of 40 lists, enabling a division of the performance session into 4 periods at the end of the 10th, 20th, 30th and 40th digit lists. The time interval between lists was reduced to 5 seconds since only one number was being recorded. Also, as it was anticipated that the pupils might find the task more difficult, a longer practice trial was allowed (10 digit lists). The entire session lasted approximately 15 minutes.

5.1.3 Method

Subjects were 45 male and 45 female 5th and 6th form pupils from three comprehensive schools in the vicinity of Durham City. Pupils were timetabled by the school staff and were tested over the entire school day. Each subject was given a Pretest with only the experimenter present, after which he was assigned to one of three experimental conditions; E, Im-Comp or Ex-Comp. The Pretest enabled a comparison of change in performance between the non-social and social conditions.

---

1. Initially subjects were asked to record the fourth digit back, but the error rate was unacceptably high (over 50%), and the first few subjects had to be discarded to allow for alterations in the task instructions.

One to two weeks after the recruitment meeting, pupils reported one at a time for the Pretest. This was identical to the experimental test, except for the particular pages of material used for the cancellation task. Instructions for the Pretest were designed to be neutral and emphasized that the experimenter was only interested in seeing how pupils learned to do the task. Following the instructions, subjects were given the practice trial and then the Pretest.

The experimental rooms differed between schools, but all were removed from the general school traffic and reasonably quiet. The experimenter gave the signal to "start" and "stop" and remained in the room during the testing session. As she was not involved with recording scores, the experimenter sat about 10 feet from the subjects, facing away from them, and was occupied in scoring previous tests. An isolated (A) condition could have been arranged, but since pupils were closely timetabled there was some danger of early arrivals interrupting the previous subject. Also, it was considered necessary that the experimenter be present to supervise the group sessions and ensure that subjects did not talk to each other during the test.

Approximately one week after the Pretest, subjects reported for the experimental session. The school staff was given a pre-arranged testing schedule, randomly varying the experimental conditions, and pupils were allocated to different conditions according to the times they were not in class. Although the procedure was not random in the usual sense, there should not have been any systematic bias in the assigning of subjects to different conditions. Groups were composed of either all male or all female members. One third of the subjects were tested under the same conditions as the Pretest (E). The other two-thirds were tested in coacting groups of three, half receiving neutral instructions (Im-Comp) and the other half ego-threatening instructions (Ex-Comp).

Summary of Conditions

Day I - Pretest (E)

Day II - E
Im-Comp
Ex-Comp

On Day II the task instructions were briefly reviewed with all the subjects, followed by the neutral or ego-threatening explanations of the task. The neutral instructions stated that because fatigue interfered with performance of the task subjects were being tested in two short sessions in order to minimize any such effects. The ego-threatening instructions stated that the test was related to reading skills and intelligence and also stressed comparison of the subjects' scores with those of other pupils: the Day I session was described as a practice session.

Following the experimental instructions subjects did the task and then completed the EPI, TAS and ASI. Pupils were not debriefed immediately after the experiment due to the risk of their communicating the experimental manipulations to other subjects. However, those in Ex-Comp were specifically asked not to tell the others that the test was an intelligence test, explaining that the preknowledge would give the others an unfair advantage. Letters describing the experimental manipulations were sent to each school after all the subjects had been tested. Because subjects were debriefed in this way it was not possible to obtain their reactions to the experimental manipulations. However, this population was experimentally naive, and, since the experiment was sanctioned by the Head Master/Mistress, it is unlikely that any suspected the deception during the actual experiment.

5.1.4 Results and Discussion

Separate 2 X 3 analyses of variance were computed on the difference scores for Output, Omissions, Hits and the number of Correct Digits, factors being Sex and Social Conditions. Difference scores were obtained by subtracting the mean score per performance period on Day I $(\frac{\text{Total score}}{4})$ from each score on Day II. As in Experiment II an initial examination of the data revealed no differences in performance over time which were related to the different conditions of testing. Therefore, a Time Periods factor was not included in the analysis. (Figures 5.1 - 5.4 show that the changes in performance over time are similar for all conditions.)

One obvious difference, however, between the data for the school children and the university students is that the latter show, if anything, improvement over time for Output and Hits; for the pupils there is a clear and constant degradation of performance on these measures. This may have been due to the younger subjects being more easily bored with the task, and therefore less persevering, or to the fact that the pupils had done the task before in the Pretest, thereby reducing on Day II any positive effects due to initial novelty.

Output

Table 5.1 shows the mean change in performance and standard deviations for Output under the three conditions of testing for males and females.

TABLE 5.1

Mean Change on Day II for Output According to Social Conditions and Sex

	E	Im-Comp	Ex-Comp
Males	-1.2	39.7	22.3
Females	50.1	15.7	11.7
Overall Mean	24.4	27.7	17.0
(S.D.)	(43.1)	(46.7)	(44.1)

Figure 5.1.

Mean Change over time on Day II for Output according to Social Conditions

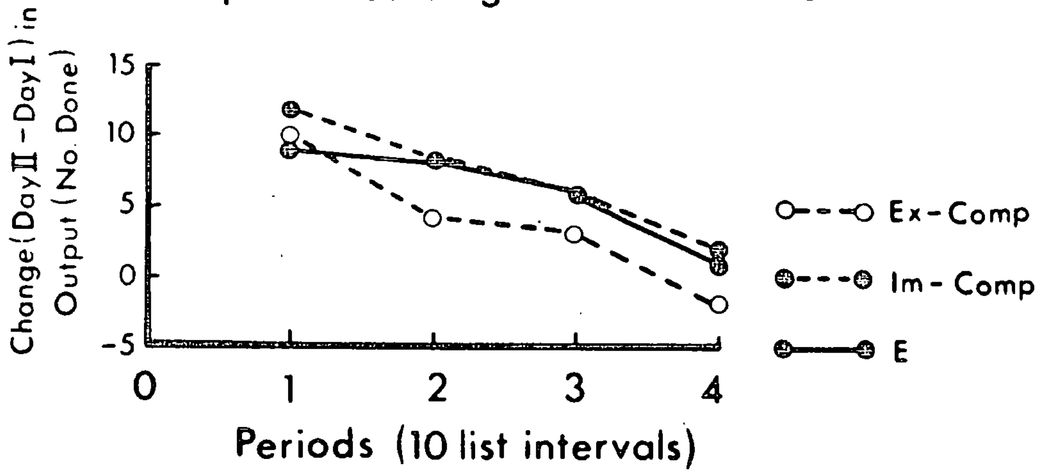


Figure 5.2.

Mean Change over time on Day II for Omissions according to Social Conditions

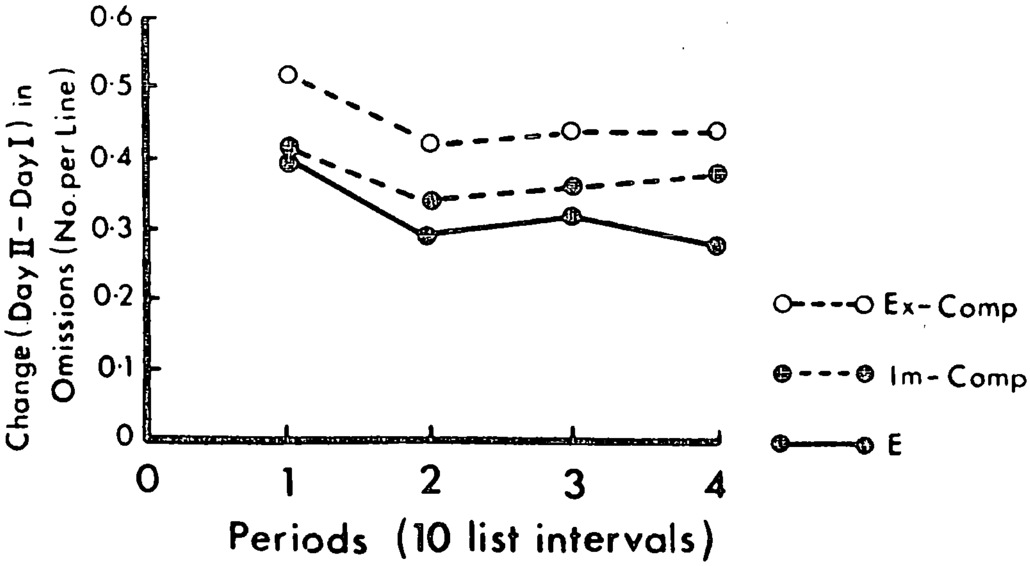


Figure 5.3.

Mean Change over time on Day II for Hits according to Social Conditions

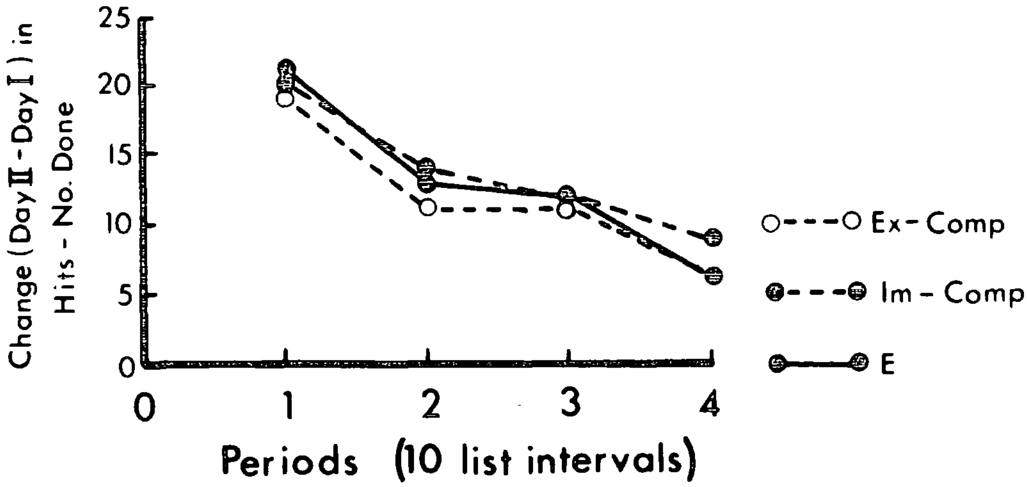
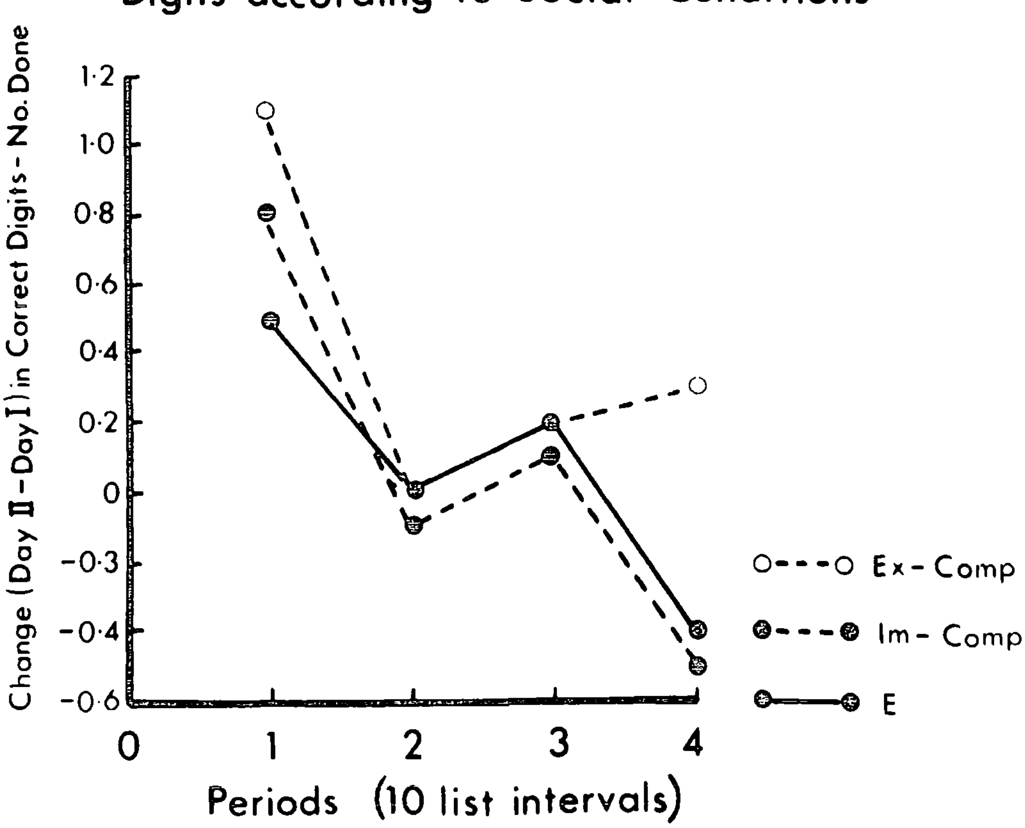


Figure 5.4.

Mean Change over time on Day II for Correct Digits according to Social Conditions



The most striking feature of the table is the difference in the pattern of performance change between males and females in social and non-social testing conditions. The analysis of variance revealed a significant Sex X Social Conditions interaction ($F = 6.53$; $df = 2, 84$; $p<.005$). This was the only significant effect on this measure. The sex difference is most clearly seen in the E condition, which is associated with the greatest relative improvement for females but is the worst condition for males. Simple main effects revealed a significant difference between males and females in the E condition ($F = 76.92$; $df = 1, 84$; $p<.001$). The pattern is reversed in the two social conditions with males showing relatively greater improvement than in E ($F = 3.66$, $df = 2, 84$; $p<.05$) and females showing less improvement ($F = 3.43$; $df = 2, 84$; $p<.05$). The two group conditions are undifferentiated in terms of degree of improvement. The level of performance improvement for males is higher than that for females in both social conditions, although these individual comparisons were not significantly different.

Omissions

There were no significant main effects or interactions on the Omissions measure. Subjects in all conditions of testing show increased accuracy on Day II, presumably due to practice (see Figure 5.2), but the differences obtained for Output (Sex X Social Conditions) were not paralleled by any marked changes in accuracy.

Hits

Table 5.2 shows the mean relative improvement for Hits on Day II in the three conditions of testing and separated by Sex.

TABLE 5.2

Mean Change on Day II for Hits
According to Social Conditions and Sex

	E	Im-Comp	Ex-Comp
Males	26.7	60.1	50.5
Females	64.9	48.7	49.5
Overall Mean	46.5	54.4	50.0
(S.D.)	(35.2)	(38.8)	(34.2)

The overall pattern is similar to that obtained for Output, although less pronounced. Once again the only significant effect was the interaction between Sex X Social Conditions ($F = 3.83$; $df = 2, 84$; $p<.025$). Females show the most improvement in the E condition while males show the least, and this difference between the sexes is reversed in the social conditions. In an analysis of simple main effects, only the comparison between males and females in E attained significance ($F = 8.06$; $df = 1, 84$; $p<.01$).

Correct Digits

There were no significant main effects or interactions obtained for this measure and, therefore, no evidence of any shifts in attention or restricted cue utilization. However, the comments made in Section 4.2.5 pertaining to the unusually high error rate on the digit task when performed simultaneously with cancellation are relevant here as well.

In general, for Output and Hits females show greatest relative improvement when tested with only the experimenter and least in the two group conditions, while males show the reverse pattern. This differential reaction for males and females cancels out any main effects for the social conditions. There is little difference between the Im-Comp and Ex-Comp conditions in terms of performance change; both males and females show a nonsignificant but lower mean performance in Ex-Comp (apart from females on Hits). This suggests that either competitive elements were operating in the two group conditions to about the same extent, regardless of instructions, or that the ego-threatening instructions were irrelevant and group presence was the variable responsible for the performance differences.

However, since evaluation stress was associated with pacing tendencies in the previous experiments, evidence of pacing in both Im-Comp and Ex-Comp in the present case would suggest that evaluative stimuli were present in both conditions. Table 5.3 shows the correlations between Output in the Pretest (raw scores) and change in Output on Day II (Day II - Day I). The rationale is the same as that in Experiment III i.e. pacing would be indicated by a negative correlation between the level of performance on the Pretest and the change score on Day II. Output was chosen as the appropriate measure for the correlational analysis since speed, or the amount of work completed, would be the only source of feedback available from others.

TABLE 5.3

Pearson Correlations between Output in Pretest and Change (Day II - Day I)

	E	Im-Comp	Ex-Comp
Males (n = 15)	-03	-52*	-64**
Females (n = 15)	19	-61*	-68**
Total Sample (n = 30)	10	-57**	-66**

(Decimal points are omitted from the table)

*p<.05; **p<.01

As can be seen, there is no correlation between the Pretest scores and change on Day II for subjects tested in E, not even the expected regression toward the mean. However, a significant negative correlation exists for both Im-Comp and Ex-Comp conditions, the magnitude being slightly greater in Ex-Comp as predicted. These relationships are represented in scatterplots for each condition of testing in Figure 5.5 (a),(b) and (c).

The different mean performance patterns shown by males and females is intriguing in terms of their possible relationship to pacing behaviour. Horner (1974) suggests that females react unfavourably to competitive stimuli and will avoid success rather than tolerate the stress associated with deviation from the traditional female role. If the better female performers were in fact avoiding success in the more competitive situations this might explain the relatively lower level of improvement obtained for females in the group conditions. Males on the other hand, would not be expected to avoid success but to follow the more usual pattern of avoiding failure. Therefore, those males demonstrating lower performance levels would be expected to be favourably stimulated by the group situation (the better performers already being above the group mean).

To investigate this possibility the 6 best and 6 poorest male and female performers on the Pretest were studied in detail. Table 5.4 displays the mean change scores on Day II for these individuals.

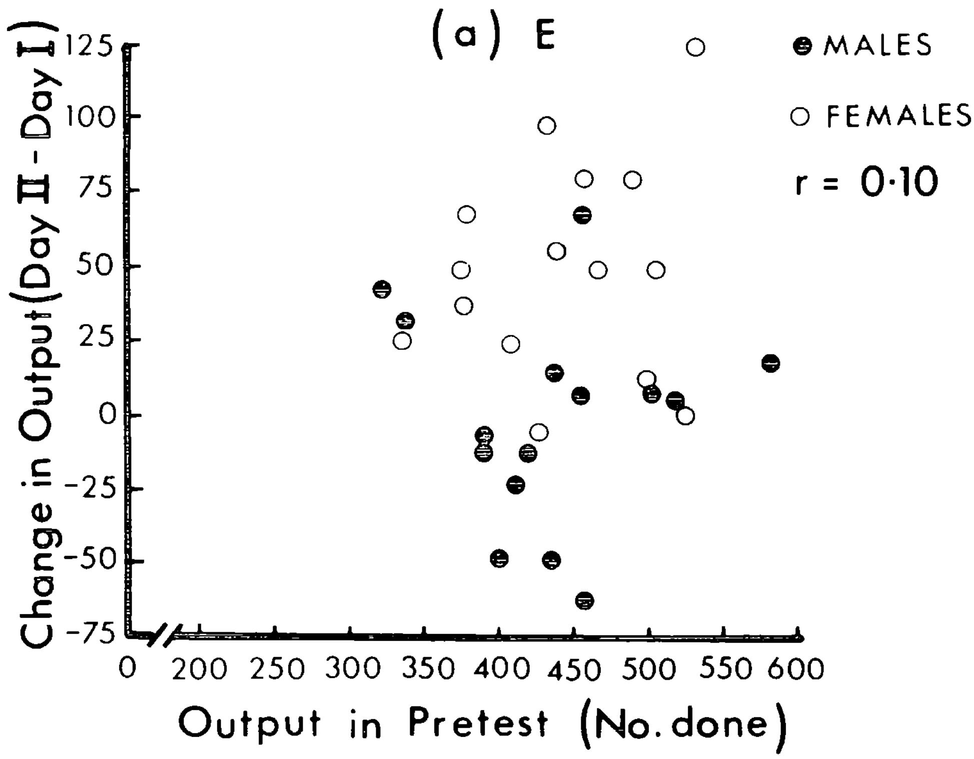
TABLE 5.4

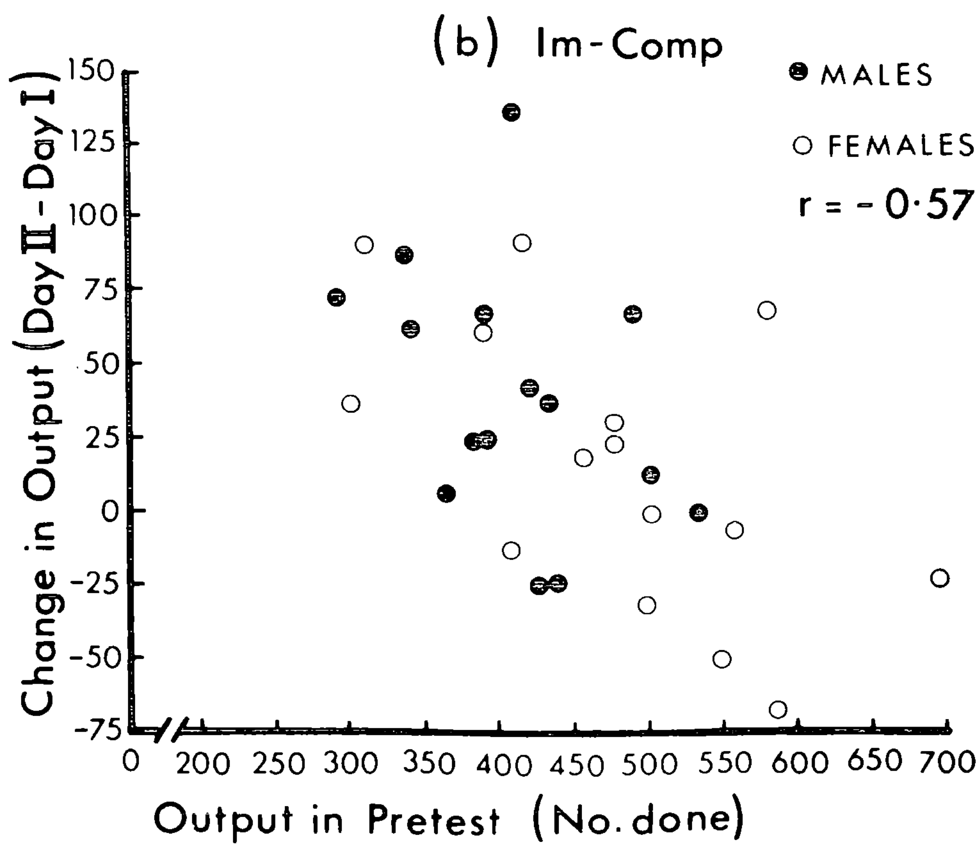
Mean Change for Males and Females
According to Initial Performance Level in the Pretest for Output

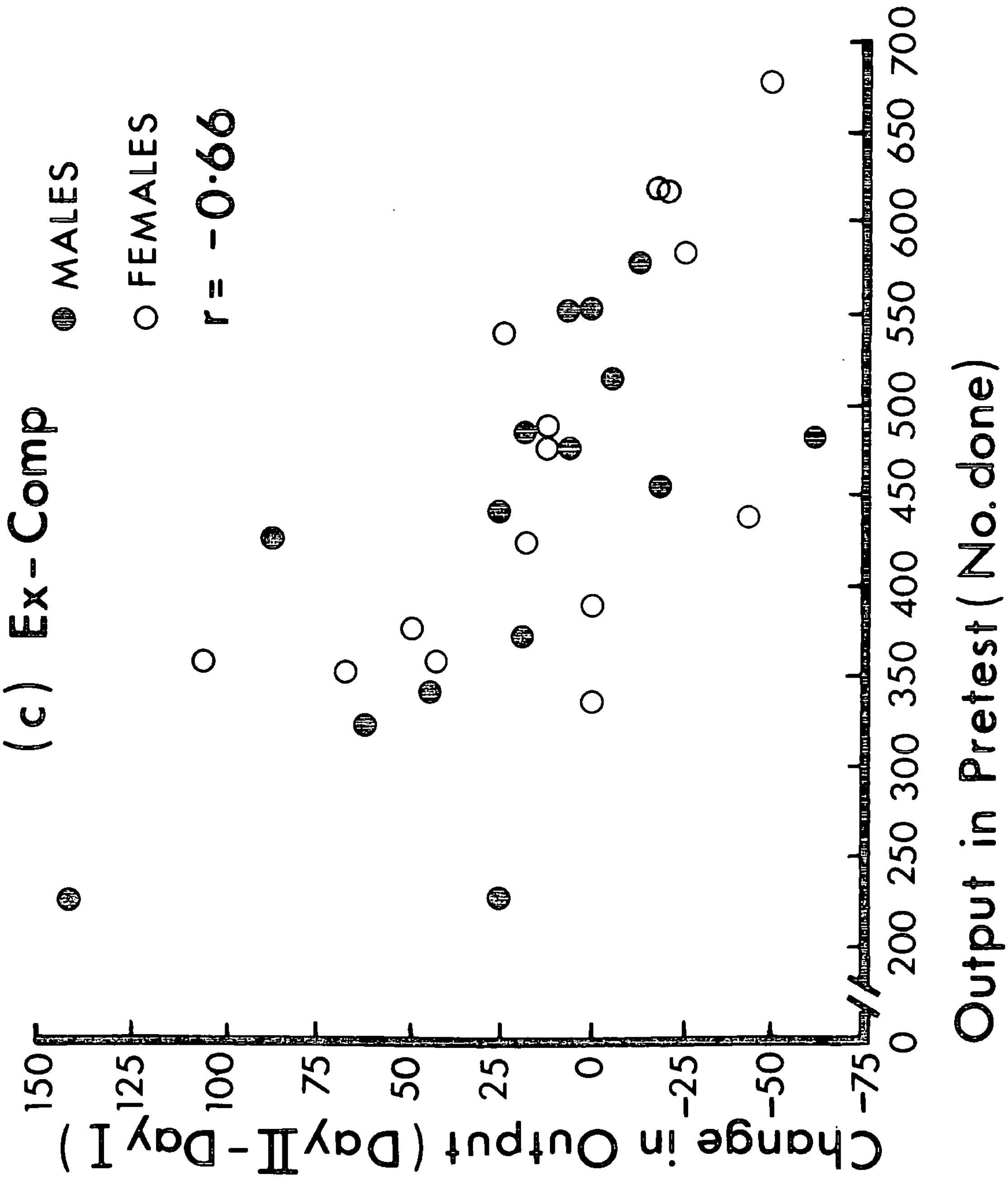
		E	Im-Comp	Ex-Comp
Males	Best 6	7.0	11.5	-8.3
	Poorest 6	-2.2	53.7	63.3
Females	Best 6	52.8	-13.2	-12.8
	Poorest 6	33.0	48.3	44.2

Figure 5.5.

Relationship between Pretest score and Change on Day II for Output.






(c) Ex-Comp
MALES
FEMALES
r = -0·66
Change in Output (Day II - Day I)
150
125
100
75
50
25
0
-25
-50
-75
0
200
250
300
350
400
450
500
550
600
650
700
Output in Pretest (No. done)

These figures suggest differential reactions between males and females in social-evaluative situations according to initial differences in level of performance. Males show little improvement from the Pretest to Day II when tested in E regardless of initial performance level. The better performers show relatively the same amount of improvement in all conditions. However, the poorer male performers show substantial improvement in both group conditions; this difference is significant for E versus Im-Comp ($U = 4$; $n = 6,6$; $p<.02$) and E versus Ex-Comp ($U = 4$; $n = 6,6$; $p<.02$). Clearly, the relative facilitation noted for males when tested in the social situations can be attributed largely to these initially poor performers.

The pattern of change for females is decidedly different. Females of both ability levels show substantial improvement on Day II in E. However, in both group conditions the initially better performers show performance impairment. These subjects show more improvement in E than in either group conditions; E versus Im-Comp ($U = 4$; $n = 6,6$; $p<.02$), E versus Ex-Comp ($U = 3$; $n = 6,6$; $p<.01$). Although the initially poorer females show greater relative improvement in the group conditions compared with E these differences are not significant. It can be seen from this analysis that the decrement in performance shown by the initially good female performers is responsible for the relatively lower mean improvement shown by females in general in the group conditions.

The analysis according to initial performance level suggests that not only did pacing occur in the social situations, but the pattern was different for males and females. The poorer male performers paced upward, raising the overall means, while the best female performers paced downwards, lowering the overall means.

5.1.5 Summary of Results and Discussion

The results from Experiment IV can be summarised as follows:

1. On both Output and Hits males and females showed a differential response to the social manipulations: females demonstrated greatest performance improvement when tested with only the experimenter present while males showed a slight decrement in this condition and relatively greater improvement when tested in groups.

2. Based on correlations between Pretest and Change scores, there was evidence of pacing in both group testing conditions, the magnitude of the effect being only slightly greater under explicit competition.

3. Although males and females demonstrated pacing to roughly the same degree, when separated into ability groups different patterns were observed. Initially good male performers showed no difference in relative improvement in any condition on Day II, while poor performers demonstrated relatively greater improvement in both group conditions. This pattern was reversed for females, initially good performers showing least improvement in social conditions, while poorer performers improved to the same degree on Day II regardless of conditions.

Although there was some anticipation that sex differences might feature more prominently with this subject population, the effect was more pronounced than expected. A differential response to the non-social condition was not predicted. Although pacing patterns in the group conditions can adequately account for the levels of mean improvement in these situations, the response to the E condition on Day II was uniform within sex groups. It is possible that female subjects were in general more conscientious and more eager to win the experimenter's approval; this would constitute a type of audience effect. Alternatively, males may have been stimulated by the novelty of the experiment on Day I, and the lower Day II mean might reflect habituation or boredom. In any event, the effect was quite marked and indicates some basic difference in the way the two sexes reacted to the most neutral experimental treatment.

The pacing data offer some support for Horner's (1974) theory postulating a motive to avoid success for females, at least for young female populations not already selected for academic or career interests. However, the independent groups design in Experiment II and the small number of subjects employed in Experiment III precluded a detailed analysis of male and female performance patterns according to initial performance level. As discussed previously, the analysis of mean performance changes can obscure different strategies and performance patterns within separate conditions.

The present data highlight the inadequacies of analyses based on mean effects only. It is noteworthy that, although in every condition on Day II there was mean improvement in performance level (with the exception of males in E who showed no change), the best performers in the group situations showed either little improvement or notable decrement. It is not clear of course to what degree the tendencies observed in these experimental conditions generalize to natural situations, but if they do, the implications may be far reaching.

The results from this experiment provide further indirect support for the suggestion made in previous chapters that evaluative cues can be nearly as prevalent in mere presence as in explicitly evaluative situations. That the pacing correlations were of substantial magnitude even in Im-Comp suggests that the pupils were at least as sensitive to evaluative cues as were the university students and were using the information available from others to gauge their own performance. On the other hand, it is well known that adolescents are more prone to conform to peer group norms than are most other age groups (Erikson, 1965), and the pacing which occurred may have reflected a general conforming tendency which was independent of any evaluative stimuli associated with the experimental procedure. However, had subjects simply been conforming to a group norm there is no obvious reason why poor and good performers should not have conformed in the same manner, regardless of their sex, or why some subjects (better males and poorer females) were unaffected by the experimental manipulations.

The similarities in performance change between Im-Comp and Ex-Comp suggests that specific ego-threatening instructions have little effect in this type of situation. The pacing correlations indicate that instructions, rather than being ineffective, were simply redundant because evaluative cues were already operative in the group situations. In fact several pupils at various stages in the experiment asked if the experiment was a test. That the sessions took place during school time and were officially authorized by the Headmaster/Mistress probably invited this interpretation. It is unclear to what extent the responses of this sample were dependent on the actual physical environment i.e. its resemblance to a testing situation, administration by a tester (experimenter), a set time period to work, etc. However, these situational factors should not pose serious problems in terms of the generality of effects, since whenever performance is being measured it is usually in a situation where there is some official overseeing or control.

Finally, in terms of the components of performance, it would seem that speed of task execution was the measure most sensitive to the social manipulations. Hits, although showing the same basic trends as Output, is after all derived from Output and Omissions. Speed is of course fairly easily altered on a task of this nature. It is important to point out that with a task which is not self-paced, sensitivity to social stimulation may be reduced or may be expressed differently. Although the social environment may have basically the same effects on arousal level and social processes, the task itself may dictate the speed of performance, precluding pacing or minimizing the usefulness of others as sources of information. One possibility for future research would therefore be to investigate the generality of the effects noted here across different tasks.

5.2 EXPERIMENT V

5.2.1 Aims

The present experiment, like the preceding one, employs a sample of school children. Three questions are of interest: (1) the effects of mere presence in a non-evaluative situation, (2) the effects of performance publicness and (3) the generality of the pacing effect across different tasks.

Although pacing had been noted on both Output and RDS performance in previous experiments, these both involve quite simple operations. It is not clear that this same pattern would emerge with a more difficult task, such as abstract reasoning, for which an increase in arousal might be expected to impair rather than facilitate performance, or on a task for which speed (Output) was less under the subjects' control. Tasks requiring a greater degree of internal thought and 'working out' may necessarily demand that attention be focused inwardly, attenuating the influence of the physical environment. In addition, tasks which are challenging or intrinsically motivating, such as I.Q. tests, are known to be relatively impervious to mild environmental stress (Kahneman, 1973).

To investigate these possibilities, the present experiment employs two tasks which are virtually identical in their methods of execution and the performance sources available for comparison, but which are essentially different in level of difficulty and the type of mental skill required; one is a routine transformation task and the other syllogistic reasoning. In addition to facilitating comparisons across qualitatively different tasks, these activities are more amenable to manipulations involving performance publicness than were either vowel cancellation or RDS.

Since it seemed likely that the experimental situation itself might suggest evaluation (based on Experiment IV) explicit competition was minimized. Instructions were designed to provide a legitimate explanation for the experiment for which individual performance level was irrelevant. The experimental design and methods for analysing mean effects and performance patterns within conditions are the same as those employed in Experiment IV.

It was expected that when coactors were prevented from seeing each other (by means of screening) mere presence would facilitate performance of the routine task but either impair or fail to affect performance on the more difficult task. Even though evaluation was not stressed, pacing, in terms of the amount of work completed, was expected to occur when others' performance level was accessible. There was some suspicion that this might be less pronounced on the more difficult task, as it not only demanded greater concentration but was less amenable to alterations of speed at will.

5.2.2 Tasks

Each subject in each experimental condition engaged in two tasks. The easier of the two was a transformation task (Transformation) for which a series of problems was presented in the form of a letter of the alphabet plus a number from 2 to 6. Subjects were required to count forward in the alphabet the designated number of steps and write down the letter reached.

Examples

A + 2 = __

S + 4 = __ (answers C, W, R)

L + 6 = __

The second activity, a syllogistic reasoning task (Syllogisms) based on grammatical transformations (Baddeley, 1968), involved statements about pairs of letters (AB or BA) which subjects had to identify as either true (T) or false (F).

Examples

A follows B - BA ___

B is not preceded by A - AB ___ (answers T, F, F)

A is followed by B - BA ___

The complete test can be found in Appendix IV.

Three performance measures were recorded for each task; Output (number of problems attempted), Errors (number of wrong answers) and Correct Solutions (Output minus Errors). As the Pretest and Day II test were of different lengths (2 and 6 minutes respectively), all data were transformed into comparable units i.e. Output, Errors and Correct Solutions per minute.

On Day I (Pretest) subjects were given printed test sheets with the problems followed by a space for the answer. However, on Day II they were each issued with a stack of cards on which the problems were printed and a separate sheet to record the answers. Each card contained 4 problems. Subjects took the cards from their stacks one at a time and put them in a separate pile when finished. As the cards were made of thick cardboard (about 1/8 inch), it was possible to see quite early in the testing session the relative standing of each group member (in terms of work completed) by the height of the finished or unfinished stack of cards. Subjects were seated around a small table with their cards in front of them, making it difficult to avoid noticing the others' cards. The problems were presented in this way in order to facilitate possibilities for comparison and reduce ambiguity regarding relative standing.

On Day II each subject was given 45 cards for the Transformation task (180 problems) and 32 cards for the Syllogisms task (128 problems). Based on the data from the Pretest this was regarded as considerably more than any subject was likely to complete in the time available.

Superficially the tasks were identical, as the test materials were the same and the problems consisted of only one line each and required only a single letter answer. Any differences which emerged could be assumed to be due to qualitative differences in the actual problems. That the Transformation task was found to be the easier of the two was supported by the Pretest data, as subjects solved these problems more quickly (greater Output) and made fewer errors (2.6%) than on the Syllogisms task (11.0%). Since subjects were instructed not to guess but to only put down answers they believed to be correct, this high error score for the Syllogisms task indicates that the problems were in fact quite difficult for these subjects.

5.2.3 Method

The design in the present experiment is identical to that of Experiment IV. All subjects were given a Pretest and were then assigned to one of three experimental conditions: Experimenter only (E), Implicit Competition with coactors visually screened (Im-Comp) or Implicit Competition with visual feedback available from coactors (Im-Comp (F)). Im-Comp in the present case is not directly comparable with the Implicit Competition condition in Experiment IV. Although in both cases public feedback was intended to be minimal, in this experiment deliberate care is taken to eliminate all sources of intersubject comparison in one coaction situation (Im-Comp) while maximizing possibilities for comparison in a second, otherwise identical, situation (Im-Comp (F)).

Subjects were 30 male and 30 female 5th and 6th form pupils from three comprehensive schools in County Durham. (None of the schools were the same as those used in Experiment IV). Subjects were addressed in large groups and asked to participate in much the same way as in Experiment IV. Those who agreed to participate (nearly 100%) were given the Pretest at that time in the large group. The personality measures (EPI, TAS and ASI) were also administered at this time in order to reduce the length of the individual experimental sessions. Before the Pretest, subjects were read the task instructions and told that with university students a relationship had been found between personality scores and different strategies for doing the tasks and that replication of these findings was desired with younger people. Evaluation was never mentioned. These instructions were followed by 2 minutes of work on each task.

Seven to fourteen days after the Pretest, subjects reported for Day II and were tested in one of the conditions described above, 10 subjects to each condition. In Im-Comp subjects were seated around a table as in the Im-Comp(F) condition but were prevented from seeing each other by a set of portable screens which ran diagonally across the table and intersected in the middle. However, subjects could still hear each other and were physically as close as in the Im-Comp(F) condition. Since there were 10 subjects per cell, one group in each of the social conditions had 4 members, the other two groups only 3. No effects were expected due to this difference in group size. All groups were single sexed.

Since initial level of ability had proved to be an important variable in the previous experiment, the precaution was taken of matching subjects on this variable in the experimental conditions, reducing any random effects due to dissimilar distributions of ability between conditions. Male and female subjects were ranked separately on the basis of their overall scores on the Pretest (total number of correct answers on both tasks) and assigned to different conditions according to a latin square rotation. The school staff was given a list of the different groupings of subjects and they assigned testing times accordingly. Therefore, the time of testing was not controlled for.

On Day II the task instructions were reviewed with all subjects followed by a more thorough explanation of the experiment, which again avoided any mention of evaluation. "Start" and "Stop" signals were given by the experimenter who remained in the room at all times, seated about 10 feet from the subjects. The experimenter faced away from the subjects and appeared to be busy with other work. Pupils performed both tasks for 6 minutes each with a one minute rest period between the first 3 and last 3 minutes of each task. During these breaks the experimenter indicated on each subject's answer sheet how far he had progressed, effectively dividing each test into two time periods. The interval between tasks was 2 minutes and allowed for the collection and distribution of the different test materials.

As in Experiment IV subjects were not debriefed immediately following the experiment but were sent letters explaining the experimental manipulations.

The design was not counterbalanced for the order of performance of the two tasks: all subjects worked on the Syllogisms task in the second half of the experiment. It was not expected that fatigue would affect performance at any point in the testing session due to its shortness and frequent rest periods. However, there was some concern that the effects of the social situations might reduce over time, as was suggested in Experiment I. Therefore, measures were taken to reduce the likelihood of this occurring. The complete break between the first and second half of the session and the issuing of fresh testing materials midway through was designed to counteract habituation by redirecting attention to the immediate environment i.e. the social situation. Furthermore, each performance period was divided into two shorter periods, allowing for the examination of performance over time. As suggested earlier, a decrement might be expected on the Syllogisms task under the social conditions of testing. However, this would be reflected in a stable but lower level of performance throughout the test rather than progressive degradation, which would indicate habituation or fatigue.

5.2.4 Results

A 3 X 2 X 2 analysis of variance was conducted on the data for each of the three performance measures on both tasks, factors being Social Conditions, Sex and Time Periods.[2]

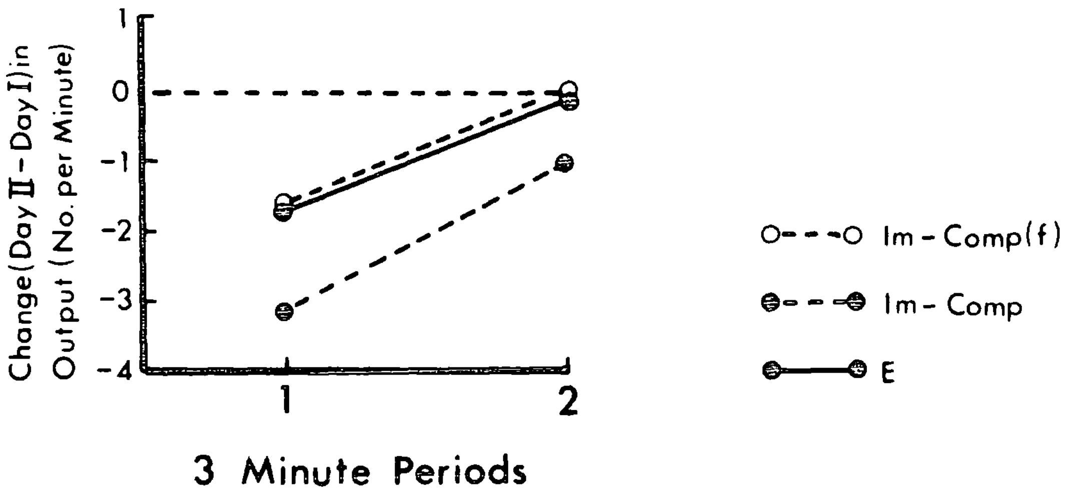
Transformation

Output

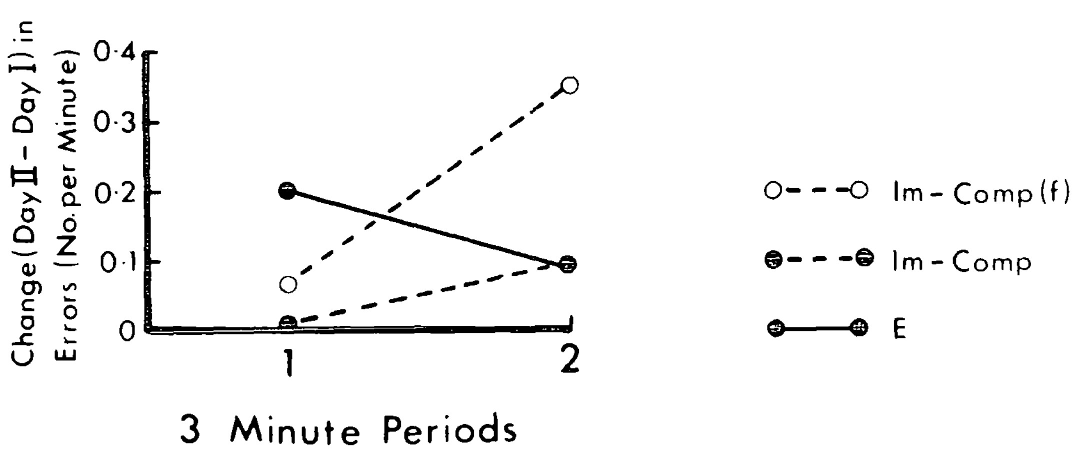
For the Transformation task the only significant effect was that of Time Periods. Figure 5.6 shows the mean change in performance for Output from the Pretest to Day II. As can be seen subjects in all conditions improve to about the same degree between Periods 1 and 2 ($F = 98.74$; $df = 1, 54$; $p<.001$). Unexpectedly, all conditions show a performance decrement on Day II (evidenced by the negative difference scores) which was significant in all three cases via a sign test: E ($N = 19$, $x = 2$; $p<.01$), Im-Comp ($N = 20$, $x = 1$; $p<.01$) and Ex-Comp ($N = 20$, $x = 5$; $p<.05$). The decrement can most likely be accounted for by differences in the testing procedure and materials. The cards used for the second test were somewhat less efficient than the printed test sheet employed for the Pretest, resulting in the loss of several seconds over the whole test period in turning up and disposing of cards. The

2. Unfortunately, one female subject in the E condition did not appear for the Day II test. As this was the final day of the school term there was no opportunity to procure an additional subject. Rather than use an analysis of variance for unequal cells, data for a 10th subject was estimated, after the method suggested by Winer (1962). However, all correlational analyses use only real subjects.

Figure 5.6.
Mean Change (DayII- DayI) in Output on Transformation according to Social Conditions and over time



Figures 5.7.
Mean Change(DayII-DayI) in Errors on Transformation according to Social Conditions and over time



decrement was remarkably uniform (52 out of 60 subjects), and, as reactions to social stress are usually less consistent between individuals, this suggests that task variables probably were responsible.

Errors

Figure 5.7 shows the mean improvement for Errors graphed over time for the three social conditions. The only noteworthy difference in the patterns of performance between the conditions is the marked increase over time in accuracy for subjects tested in Im-Comp(F). This difference is reflected in a significant Social Conditions X Time Periods interaction ($F = 5.10$; $df = 2, 54$; $p<.01$). Although none of the differences between conditions were significant in either time period, simple main effects indicated that the improvement for Im-Comp(F) subjects between Periods 1 and 2 is significant ($F = 10.20$; $df = 1, 54$; $p<.01$). Therefore, although no differences were obtained between conditions for the Output measure, there does seem to be some facilitation of performance in the Im-Comp(F) condition. This would appear to be an overall advantage since there is no evidence of any reduction in speed for these subjects.

Correct Solutions

The pattern of performance change for Correct Solutions is much the same as that for Output. Therefore, the means and standard deviations are shown in Table 5.5 rather than presented graphically.

TABLE 5.5

Mean Change (Day II - Day I) in Correct Solutions for Transformation according to Social Conditions

	Period 1	Period 2	Overall Mean	(S.D.)
E	-1.51	- .06	- .78	(1.37)
Im-Comp	-3.15	- .92	-2.04	(2.59)
Im-Comp(F)	-1.43	0.33	- .55	(2.16)

The only significant effect was that for Time Periods ($F = 93.01$; $df = 1, 54$; $p<.001$). Subjects in all conditions improved their performance between the first and second half of the testing session. The Social Conditions factor approached significance ($F = 2.77$; $df = 2, 54$; $p = .07$) due to the relatively greater decrement from the Pretest to Day II shown by the Im-Comp subjects.

Syllogisms

The only significant effect obtained on the Syllogisms task was that of Time Periods. As can be seen in Figure 5.8 all conditions improved performance for Output over time ($F = 12.75$; $df = 1, 54$; $p<.001$). The same trend is evident for Correct Solutions as shown in Table 5.6 ($F = 9.08$; $df = 1, 54$; $p< .005$). The initial decrement is less for all conditions than was the case for Transformation: less decrement would be expected as subjects became more proficient using the cards. Although subjects tested in Im-Comp do not show a performance change which is notably different from those tested in other conditions, they are the only individuals to show a significant decrement from the Pretest to Day II ($t = 2.56$; $df = 19$; $p<.02$). This is consistent with these subjects' marginally greater decrement for Correct Solutions on the Transformation task.

There is no evidence from the present data to suggest that fatigue or habituation were in operation to any extent. On the contrary, progressive improvement is the most outstanding feature of the data. Only E subjects fail to attain their Pretest level by Period 2 on Output and all three conditions show a positive gain on Correct Solutions by the second half of the testing session.

Pacing

Table 5.7 presents the correlations between the Pretest scores for

Figure 5.8.

Mean Change (Day II - Day I) in Output for Syllogisms according to Social Conditions

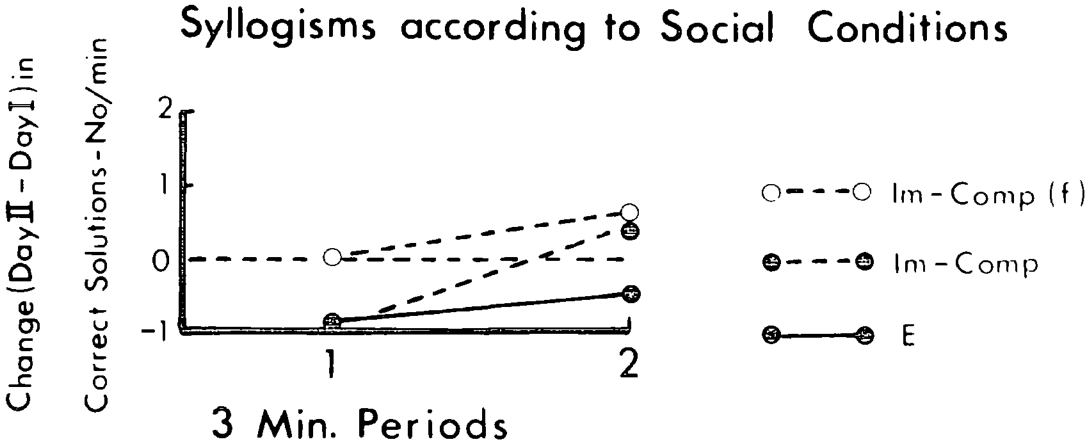


Table 5.6

Mean Change (Day II - Day I) in Correct Solutions for Syllogisms according to Social Conditions

	PERIOD 1	PERIOD 2	OVERALL MEAN (S D)
E	0·0	0·4	0·18 (1·88)
Im-Comp	-0·5	0·6	0·07 (1·52)
Im-Comp(f)	0·3	0·9	0·58 (1·80)

Output and the change scores (Day II - Day I) for both Transformation and Syllogisms. Only the scores from the first periods of each task are presented since pacing would be expected to be most pronounced in the early part of the session when uncertainty about performance level would be greatest: Period 2 data are confounded by whatever pacing took place in Period 1.

TABLE 5.7

Pearson Correlations between Pretest Output and Difference Scores

		E	Im-Comp	Im-Comp (F)
Transformation	Males (n = 10)	-32	-83**	-29
	Females (n = 10)	-62	-83**	-88**
	Total Sample (n = 20)	-42	-84**	-61**
Syllogisms	Males (n = 10)	13	-27	-57
	Females (n = 10)	-10+	-35	-22
	Total Sample (n = 20)	01+	-11	-45*

(Decimal points are omitted from the table)

+Based on n - 1 subjects

*$p<.05$; **$p<.01$

On the surface the table presents a less clear picture than has been the case in previous experiments. Firstly, for Transformation there is a more pronounced regression toward the mean in the E condition than was observed for vowel cancellation. The correlations for the Im-Comp(F) subjects are in the direction and of the magnitude expected. However, there are also significant correlations of even greater magnitude for

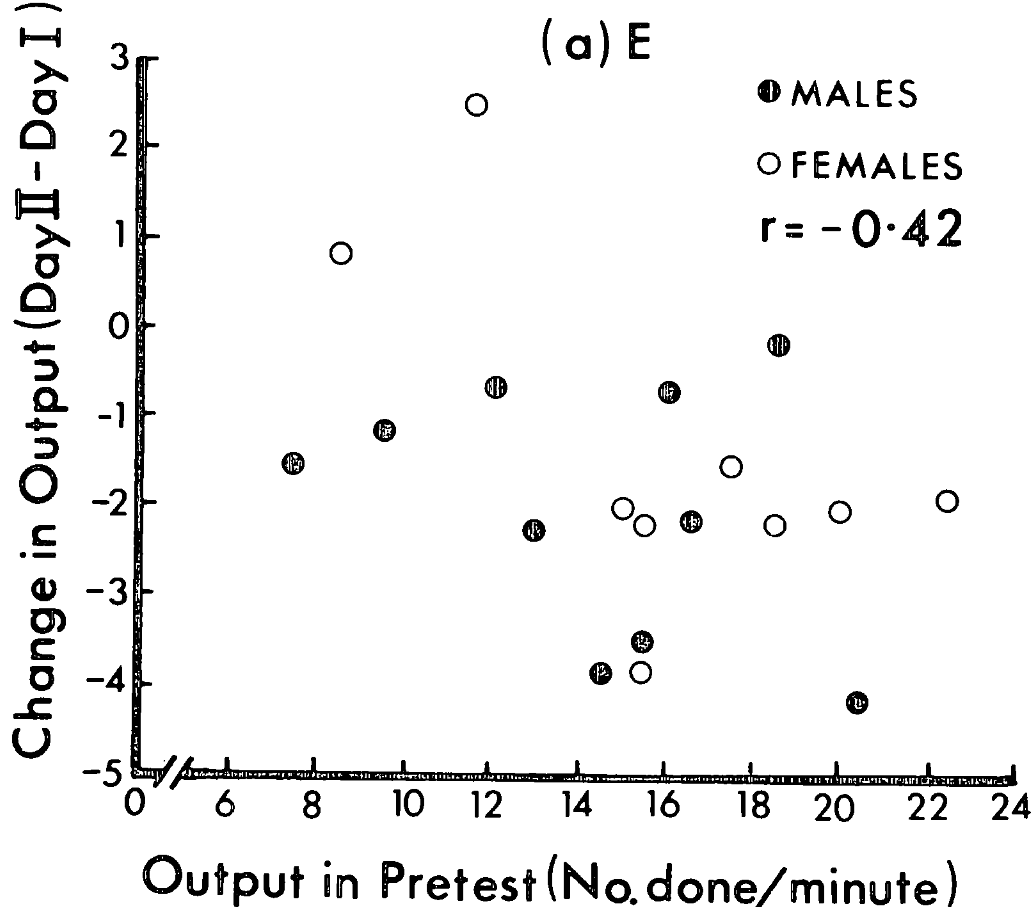
for Im-Comp. These relationships are shown in Figure 5.9. This latter result is puzzling since Im-Comp subjects could not possibly have paced themselves with one and other via visual cues. Auditory cues were available, but it is not likely that individuals would have been able to keep track of this information in such a way as to yield from it an accurate source of performance feedback.

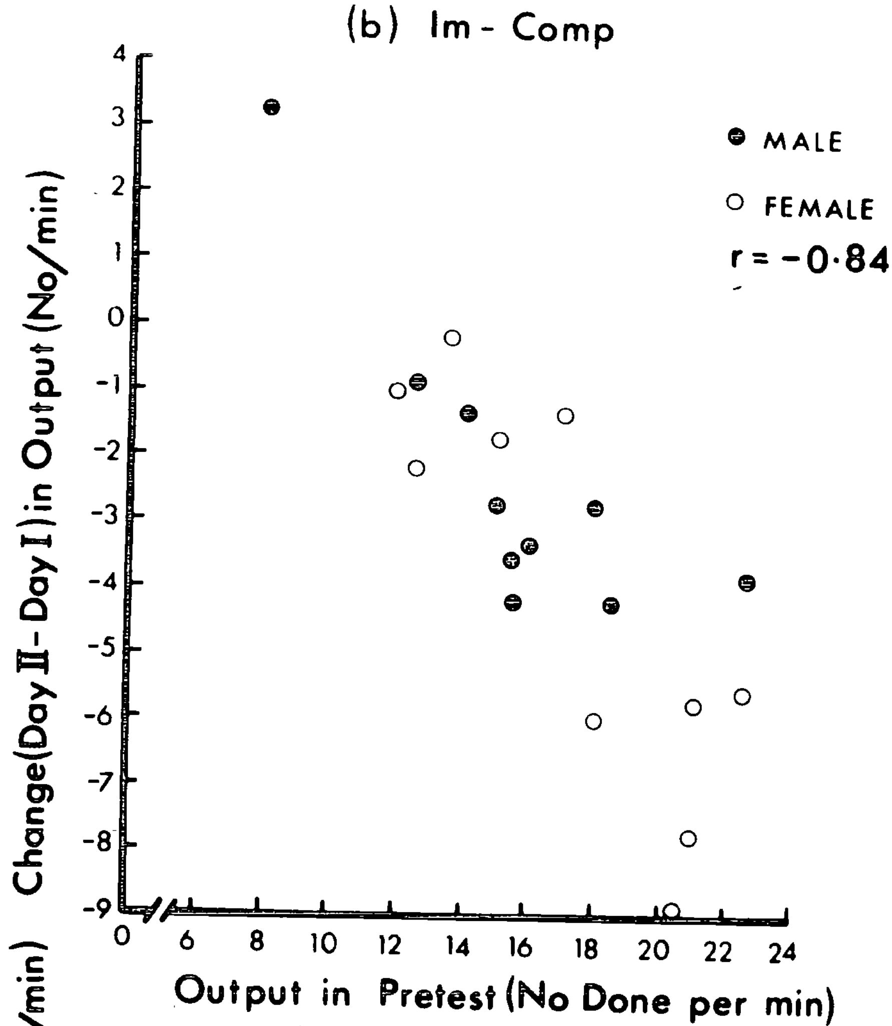
Correlations for the Syllogisms task are more in keeping with expectations. Here, as predicted, the only condition to demonstrate pacing to any marked degree is Im-Comp(F), and the magnitude of the values is consistent with the prediction that pacing would be less pronounced on this more difficult task. There is no relationship between the Pretest scores and Day II scores for E subjects and only low negative correlations for Im-Comp. Thus, subjects were either less able to pace on this task or the social stimuli were not powerful enough to elicit a pacing response when more intense concentration was required for task execution.

A more detailed look at the Transformation task data in Table 5.7 reveals that the pacing tendency is in almost all cases stronger for females than for males. Therefore an analysis similar to that undertaken in Experiment IV was completed, dividing males and females of each condition into groups of the four highest and lowest scorers (based on their Pretest performance). These data are presented in raw form in Table 5.8 as there are insufficient numbers in each ability group to merit statistical analysis.

Figure 5.9.

Scatterplots for Output in Pretest and Change in Output on Transformation.




(b) Im - Comp
MALE
FEMALE
r = -0·84
Change(Day II - Day I) in Output (No/min)
Output in Pretest (No Done per min)
4
3
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
-9
0
6
8
10
12
14
16
18
20
22
24

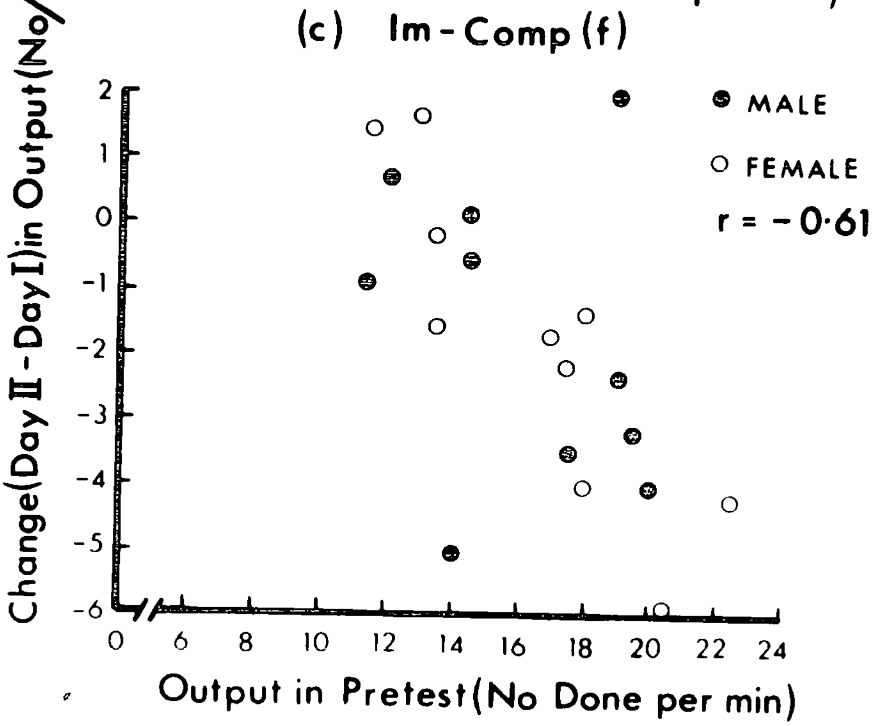

(c) Im - Comp (f)
MALE
FEMALE
r = -0·61
Change(Day II - Day I)in Output(No/min)
Output in Pretest(No Done per min)
2
1
0
-1
-2
-3
-4
-5
-6
0
6
8
10
12
14
16
18
20
22
24

TABLE 5.8

Change scores for Output on Transformation for Males and Females according to Initial Ability

	Males Best	Males Poorest	Females Best	Females Poorest
E	-4.17	-1.17	-1.50	2.50
	-0.17	-1.50	-2.00	-2.00
	-0.67	-2.33	-2.17	0.83
	-2.17	-0.67	-1.83	-1.83
Total	-7.18	-5.67	-7.50	-0.50
Im-Comp	-4.17	3.33	-8.83	-2.17
	-3.33	-2.67	-5.67	-0.17
	-2.67	-1.33	-5.50	-1.00
	-3.83	-0.83	-7.67	-1.67
Total	-14.00	-1.50	-27.67	-5.01
Im-Comp (F)	-2.33	-0.83	-4.17	-0.17
	-3.17	-5.00	-1.33	-1.50
	2.00	-3.17	-5.83	1.67
	-4.00	0.67	-4.00	1.50
Total	-7.50	-8.33	-15.33	1.50

These data, although not providing clear evidence of a sex difference, do follow a pattern which is roughly analogous to that found in Experiment IV for a comparable subject population. High scoring females show the most decrement in the social conditions and this is greater than that demonstrated by their male counterparts. Males are undifferentiated except in Im-Comp, where there was strong evidence of pacing. There is little evidence that initially low scoring males paced their performance upward as was noted in Experiment IV. However, Im-Comp(F) males showed a somewhat atypical pattern in general: as can be seen from Figure 5.9(c) two subjects responded in a manner which was inconsistent with the general pattern for this condition. This may have been due to a sampling artefact since each cell contained only 10 subjects. It is also clear, with the exception of Im-Comp(F) males that the initial decrement on Day II is largely attributable to the initially higher scoring subjects on the Pretest. This is especially true in the group conditions.

A similar post hoc analysis according to initial performance level was not undertaken on the Syllogisms task, since pacing was much less in evidence.

5.2.5 Summary of Results and Discussion

The results from Experiment V can be summarized as follows:

1. On Transformation subjects in Im-Comp(F) showed relatively greater improvement in accuracy over time than subjects in other conditions.

2. Im-Comp subjects showed some evidence of relatively greater impairment on Day II on Transformation (Correct Solutions) (marginally significant F ratio) and Syllogisms (Output) in Period I (significant decrement from the Pretest).

3. Pacing was evident in both group conditions on the Transformation task, based on correlations between Pretest and Day II - Day I change scores.

4. There was similar evidence of pacing in Im-Comp(F) for Syllogisms.

5. There was a tendency for females of initially high performance level on Transformation to respond to the social conditions with greater performance decrement than comparable male subjects.

6. An informal analysis showed that both males and females of initially higher performance level showed greater initial decrement in Transformation Output than males and females of initially lower performance level.

The most surprising result from the data in this experiment is the marked pacing tendency noted for subjects in the Im-Comp condition on the Transformation task. However, there was one important procedural difference between this experiment and the previous one; subjects were given the Pretest in large groups in the present case rather than in individual testing sessions. The Pretest conditions in this experiment probably encouraged immediate comparison moreso than in Experiment IV. After the Pretest the experimenter noticed several subjects asking each other about how much of the test they had completed. Therefore it is likely that on Day II individuals were already aware of their relative performance levels based on their experience in the Pretest. If the presence of coactors on Day II had the effect of increasing awareness of performance level, this may have led to a pacing tendency in the Im-Comp condition based on this prior knowledge. Preknowledge of relative ability would also account for the rather high negative values obtained in the E condition, since those pupils who knew they were above the norm may have felt less motivated to work hard during the second test, with the reverse tendency applying to the least able performers.

However, although both Im-Comp and Im-Comp(F) subjects showed similar patterns, their overall performance was not exactly the same. Im-Comp(F) subjects improved in accuracy over time on Transformation relative to both Im-Comp and E. In addition, there was some suggestion that the Im-Comp condition may have had a generally more disruptive effect on both Transformation and Syllogisms (see No. 3 above). These data cannot be explained by arousal theory, since subjects who had performed well in the group condition during the Pretest would be expected to show a similar tendency in group situations on Day II. However, in the present case the initially higher scoring pupils showed greater decrement on Day II.

There are reports which indicate that an invisible audience can be more debilitating than a visible one (Laughlin and Wong - McCarthy, 1975; Criddle, 1971; Ganzer, 1968). Also, Noble et al (1958) noted a decrement in choice reaction time when subjects were screened, relative to the same individuals' performance under normal conditions. Therefore, it is possible that the close proximity of fellow subjects without the customary visual cues might have constituted an unforeseen source of stress.

In this experiment, unlike the previous one, the only evidence of a sex difference came from the detailed examination of individual performance patterns for different ability groups: there were no main effects for Sex on any of the performance measures. However, procedural differences in the administration of the Pretest between the two experiments may have been responsible for this inconsistency.

The large group Pretest used in this experiment was judged in retrospect to be a rather poor procedure. In addition to the possibility that the Day II conditions would be less arousing than those of Day I, making comparisons between experiments difficult, it also afforded subjects a means of obtaining knowledge about their relative performance level prior to the experimental session. It was not anticipated that this would be a special problem, or rather it was considered that the pupils would more than likely talk about the Pretest amongst themselves regardless of how it was administered. However, the large group probably encouraged immediate comparison and permitted comparisons with a greater number of subjects than would perhaps have been the case with individual tests.

Although it was unexpected, the strong pacing tendency noted in all conditions on Day II for the Transformation task raises the interesting possibility that pacing (even in conditions of isolation) may take place solely on the basis of perceived peer group norms and independently of information available in the immediate performance situation. In natural settings performance is often assessed on an ongoing basis i.e. on a work site, in school or college, where individuals do have some notion of their relative standing in comparison with their larger peer group.

The failure to find any overall effects between conditions or any very marked differences in performance patterns within conditions on the Syllogisms task is disappointing, although not altogether surprising. As the task was quite difficult for the pupils, knowledge of Output may not have been judged a very useful measure of performance; the Error scores were rather high and individuals may have felt uncertain about their ability regardless of the amount of work they completed. Both of these factors would be expected to counteract pacing tendencies. It is also true that speed on this task was less under the subjects' control.

The problems required a certain amount of mental 'working out', and although performance could certainly be slowed down, conscious efforts to speed up may have interfered with efficiency. In any event this task proved to be fairly insensitive to the social conditions.

Although evaluation was specifically avoided, it seems likely that subjects were aware that their performance was being measured and were concerned about attaining an appropriate standard. Whether this was the result of evaluation stress suggested by the experiment or an adolescent aversion to departing from the group norm is unclear. However, the pacing correlations in this experiment, at least on Transformation, were equivalent in magnitude to those obtained in previous experiments in which evaluation was explicitly stressed.

Although one of the main aims of this experiment was to isolate the different effects of mere presence from those of public performance, no very strong differences were noted between the two group situations. The only measure which showed a clear difference between the two social conditions was Output on the Syllogisms task, on which Im-Comp(F) subjects showed a more marked pacing response. In Experiment III a question was raised concerning the different effects of performance publicness and performance information. On the Transformation task, information seems to have been the more important variable. Subjects in the private, Im-Comp, condition showed pacing to the same extent as did those in the public, Im-Comp(F) situation. In this case information from the Pretest would appear to be the only likely source from which such a result could have arisen. However, on Syllogisms, publicness presumably did matter, either as an additional source of performance information or a means of heightening objective self-awareness on this more involving task. Clearly, performance need not be public in order to stimulate pacing, but publicness may reinforce or sustain interest in performance level as well as provide an unambiguous means for comparison.

5.3 OVERVIEW

One of the primary purposes of the experiments presented in this chapter was to investigate the generality of findings from the first series of experiments (pacing, shifts in attention, performance publicness and explicit versus implicit evaluation) using a less select subject population. One result stands out quite clearly; younger and less self-selected subjects are at least as likely to show pacing tendencies when tested in group situations as are university students. Judging from the magnitude of the pacing correlations in even non-evaluative conditions, there is some reason to believe this population may be even more sensitive to social-evaluational stimulation.

The question was also raised as to whether subjects who knew each other and had some understanding of their relative ability in comparison with their peers would show pacing tendencies to any less degree than those who were complete strangers. The results indicate that in all situations where pacing was expected to occur, such an effect was found. It would seem then that, at least on a novel task, preknowledge of others' general academic ability does not reduce pacing tendencies. On the contrary, the results from Experiment V suggest that information about Pretest task performance was shared and pacing occurred on the basis of this, and independently of the immediate social situation, at least on the more routine Transformation task.

Another question of primary concern was the role of evaluative stimuli in affecting overall differences between conditions and particularly the patterns of performance change within conditions. A distinction has previously been made in this text between explicit and implicit competition in coaction situations. It would appear from the results of both Experiments IV and V that there is very little difference in patterns of performance between these two conditions, at least for this population. In Experiment IV specifically evaluative instructions produced performance patterns which were only slighly more exaggerated than those prevailing in neutral coaction situations, and in Experiment V where instructions were specifically non-evaluative, pacing on one task was as pronounced as in specifically evaluative conditions in previous experiments.

Analysis of individual differences data (initial performance level) suggests that the response to the social-evaluative stimuli used in these experiments is not competitive in the usual sense. Normally, competition is expected to stimulate all participants to greater effort. However, in these experiments the only individuals who showed clear evidence of positive stimulation were those of initially low ability. More able pupils either benefited little from the group situations or showed a performance decrement. Of course, even initially good performers who showed impairment were often still the best performers in their respective groups; the point is that they did not give their best performance under these conditions.

There are several reasons why the response patterns obtained may have differed from those which would be indicative of competition. Firstly, the tasks were novel with no clear goals, and no information was supplied as to the standard of performance expected. In such a situation group members may naturally look to others for information about 'correct' behaviour and may be influenced by it unknowingly. Secondly, the effect of the social situations and even the ego-threatening instructions may not have been to create evaluation stress or feelings of competitiveness. The situations may simply have alerted subjects to, or made them aware of, their own performance level. Attention to or interest in performance level need not be associated with any feelings of either competitiveness or evaluation anxiety. Even subjects who specifically asked whether the tasks were a 'test' need not have been particularly concerned about failure.

Finally, the situations used, as well as the tasks, were not very similar to those usually found in true competitive situations. Although individuals may choose to compete on almost any kind of activity, the important distinction is that they choose to do this. If the individual is not predisposed to compete or has no particular interest in his competence on the task, it is unlikely that instructions or social settings suggesting evaluation or competition will induce the same type of response as a free choice situation. Tasks which were superficially more like a contest or game might have been perceived differently by the subjects and have resulted in different performance outcomes.

In some respects the similarity of results (pacing, the absence of main effects and the effects of sex and initial level of ability) found between Experiments IV and V were rather surprising considering the differences between the tasks. In terms of the components of performance, it would appear that speed shows the most sensitivity to these social situations. It is the speed component on which pacing is evident, and pacing would appear to be the only fairly reliable response in these situations. The flexibility of the task, or the number of possible ways it can be executed, also seems an important variable. Generally, it was not felt that the tasks used in Experiment V were as sensitive or as likely to show differences between conditions as the cancellation/RDS task. This is because there was virtually only one way to do the Transformation and the Syllogisms tasks. Since subjects were specifically instructed not to guess at answers, speed could not be controlled by adjusting the criteria for accuracy. Also, there was no secondary task to which more or less attention could be allocated. The only choice available to subjects was to try to concentrate harder and hopefully accomplish more work.

It is likely that this inflexibility is in part responsible for the less pronounced differences between social conditions in Experiment V, particularly for the Syllogisms task. The experimenter noticed subjects at least attempting different strategies on Transformation by periodically using their fingers for counting, or other aids: even these simple strategies were not possible with Syllogisms. Previous research has relied heavily on either very routine or highly structured activities which are largely inflexible. The strategy that subjects adopt may in the end determine whatever differences are observed between different social conditions. When only one strategy is possible, performance change can only be reflected in either overall facilitation or impairment: even then differences may be small and go unnoticed unless the sample size is fairly large.

Corroborating the tentative evidence in both Experiments II and III, results from the present studies indicate that individual differences play a substantial part in determining the overall effects of the presence of others and potential evaluation on performance. In the present experiments it is clear that sex differences and differences in initial level of performance largely negated any overall facilitative or inhibitory effects due to the social manipulations. The most pronounced and reliable response was that of pacing, and even this did not follow a uniform pattern but was present in varying degrees, depending on the sex of the subject and initial level of performance. These results suggest that it is simplistic to expect much consistency in mean effects between experiments when these basic individual factors are not taken into account. Exclusive concentration on overall facilitation or impairment is likely to be disappointing and can complete obscure consistent differences between performance in social situations in terms of strategies and the patterns of performance change within conditions.

In terms of further research, several possibilities suggest themselves simply in regard to performance tasks: (1) would the effects found with the basically routine tasks employed in these experiments emerge also with tasks for which there is a clear dominant/subordinate response hierarchy, (2) would pacing be as evident on a task which was more clearly competitive or game-like, (3) what effects might be expected on a task which affords no degree of self-pacing or, alternatively, one for which several strategies are possible and performance is entirely self-paced? The effects of mere presence, instructions and performance publicness may be different for tasks which vary on these dimensions.

CHAPTER VI

EXPERIMENTS VI AND VII

6.0 INTRODUCTION

In some respects this chapter marks a major turning point in the style of the research reported in this thesis. The importance of task variables in determining the effects of social situations on performance has repeatedly been pointed out. In the previous chapter it was suggested that the tasks employed in the first five experiments were unusual in a number of ways; the absence of any clear performance goals, their novelty and dissimilarity to tasks which are normally present in group situations and the absence of any pre-existing or establishable response hierarchies for task performance. The question was raised as to whether the novelty of the tasks, coupled with uncertainty surrounding the experimental situation in general, might not have produced a highly ambiguous situation for subjects, making them more suggestable to and easily influenced by the behaviour and performance level of other subjects. Also, there was suspicion that the tasks did not elicit the type of motivation to do well which is a feature of most performance situations.

The two experiments reported here are alike in one major characteristic; both employ motor tasks which are superficially game-like and require skills which can be likened to those required in many competitive situations. As in previous experiments, the effects of mere presence, ego-threat and the degree of performance publicness are of major interest. The two tasks differ in that one (pursuit rotor) allows no degree of self-pacing and has a clear dominant/subordinate response hierarchy, while the other (darts) is entirely self-paced and permits unlimited variations in performance strategies.

In addition to examining the effects of social situations on tasks requiring an essentially different type of skill to those in previous experiments, it is hoped that the selection of these tasks will suggest a more realistic performance situation to the subjects and elicit behaviour typical of a natural situation. Although it is not assumed that the results found in the previous experiments were anomalous or confined only to experimental settings, it is felt that some attempt should be made to investigate the effects of social situations in more realistic settings and with tasks obviously related to real skills and abilities. Therefore a question of central interest is whether intrinsically involving tasks will in fact render the social manipulations ineffective or, rather, be more facilitative of greater effort under social stress. It is not presumed that the situations created in the following reports possess all or even most of the characteristics of natural social-competitive settings. However, they do provide some compromise between the sterile and rigorous laboratory performance situations used thus far and more realistic competitive social conditions. Literature concerned with attention and effort suggests that tasks which are highly involving are virtually impervious to environmental stressors such as noise (Kahneman, 1973). However, although mild social stress may be analogous to other types of environmental stress, social situations are open to variations in interpretation by subjects and a whole range of uncontrolled, unidentified variables which are not characteristic of physical environmental stressors.

Experiment VI (pursuit rotor) employs Open University students as the subject population, providing a sample of generally greater maturity and a wider range of abilities and backgrounds than either the university or school samples studied previously. Experiment VII (darts) once again involves university undergraduates. This experiment is exploratory, as the task is more complex than those typically featured in research on performance and it was unclear how efficiency should best be assessed. Because the procedure required repeated testing and fairly lengthy testing sessions, university students were the most feasible subject population.

Personality tests designed to measure sensitivity to audience and testing situations are once again employed, and these data are discussed in Appendix I.

6.1 EXPERIMENT VI - PURSUIT ROTOR

6.1.1 Aims

The present experiment examines the effects of the presence of other subjects, ego-threatening instructions and the public feedback of results on pursuit rotor performance. This task was chosen for two reasons. Firstly, it is intrinsically engaging and therefore likely to engender competitive efforts between subjects (Smode, 1958). Secondly, it essentially involves motor skill learning and therefore the testing session can be divided into learning and performance phases, enabling a test of arousal theory predictions using the type of task normally employed in drive theory research.

The pursuit rotor clearly falls into that task category for which there is a dominant/subordinate response hierarchy: the correct response initially is subordinate but is reinforced and strengthened with practice on the task. The use of such a task provides an opportunity to examine the generality of the pacing effect noted on more routine activities. For example, it is not clear whether pacing would feature in a performance situation in which subjects are occupied with simply learning the task: the degree of concentration required may limit the individual's ability to process and respond to information from others in the environment. Also, pacing tendencies may be relatively unimportant on tasks for which dominant responses can clearly be identified: the facilitation of dominant responses due to increased arousal in the presence of others may be the most striking performance outcome.

The present experiment also provides an opportunity to examine varying degrees of performance publicness and the possible interactions of task publicness with potential evaluation and mere presence. These three variables have received little systematic treatment in research testing arousal theory predictions, and the experiments reported in previous chapters of this thesis did not examine all three variables in a single design. It has been suggested (e.g. Section 2.2.2) that evaluative instructions may be more effective when the subjects' ongoing performance can actually be assessed (i.e. task publicness). Considering that potential evaluation in social situations has been deemed (in the literature) the necessary condition for raising the arousal level of performers, the interactive effects of these three variables merit investigation.

It is expected that feedback of results (publicness) will benefit both learning and performance on the pursuit rotor, regardless of the social situation (Bilodeau and Bilodeau, 1961; Annett, 1969). Most of this positive effect would presumably be due to the motivational aspects of knowledge of results. In the present case, an auditory feedback mechanism provides information whenever subjects are off-target, which might be an aid to correcting inaccuracies more quickly. However, this information does not actually offer help in developing more efficient strategies for task performance (Smode, 1958).

It is expected that the presence of others, particularly in the more stressful situations (evaluation and public feedback) will disrupt performance in the initial periods of the testing session (learning), but that performance in later stages of the session will be facilitated by these same conditions. The degree to which pacing can be anticipated is unclear, as it is questionable whether subjects would be able to pace themselves very effectively on this task. Firstly, the speed of performance is dictated entirely by the apparatus, and speed has been the task component for which pacing has been noted. Secondly, even though subjects in some conditions receive specific knowledge of results concerning their performance level relative to other group members, this information may not be useful for actually modifying performance. Individuals may not be able to learn the task better or improve their performance appreciably, regardless of their motivation to do so, if the information (feedback) does not suggest a better method for task execution (Gill and Martens, 1975). If this is the case then only downward pacing would be possible for those individuals who were initially good performers and who might become less motivated or even threatened by feedback showing their performance to be above average. However, it is predicted that if indications of pacing are found, the magnitude of the effect will be greatest in those situations in which performance is public and evaluation stressed.

No predictions are offered as to how ego-threatening instructions might affect the performance of subjects tested in non-social conditions. Unlike tasks used in previous experiments, the pursuit rotor engenders some degree of intrinsic interest: specific instructions stressing evaluation may be redundant or have no additional effect on subjects' basic interest in mastering the task. However, in social conditions these same instructions are expected to interact with the social manipulations in the manner suggested by arousal theory i.e. increasing the emission of dominant responses. It is also expected that the more stressful situations might be associated with greater individual variability in performance.

6.1.2 Task

The apparatus was a Forth Instruments pursuit rotor. Subjects were required to track an illuminated target which followed a symmetrical star-shaped pattern in a clockwise direction. Only that portion of the pattern momentarily illuminated by the light target (approximately $\frac{1}{2}$ inch square) was visible at any one time. The tracking device was a stylus with a photosensitive cell which registered the total time that the stylus point was directly over the light-target. The subjects' time-on-target (T.O.T.) was indicated to the nearest one-tenth second on a display panel in the front of the apparatus, out of the subject's view. Immediate feedback of performance could be manipulated via an auditory control which, if on, emitted a buzzing noise whenever the subject was off-target. This was a modification of the original apparatus, which was designed to give auditory feedback for time-on-target. It was felt that in the group situations auditory feedback which highlighted errors would be more effective in creating evaluation stress. Obviously the auditory signal was a form of supplementary feedback, since visual and proprioceptive feedback are intrinsic to the task. However, the auditory signal was the only source of immediate public feedback.

The target speed was set at 10 revolutions per minute. The task was not particularly difficult at this speed, and all subjects were on-target more than half the time by the end of the testing session. The session consisted of a one minute practice followed by a break of approximately three minutes. Between the first and second minute of the test period a 20 second rest was interpolated. The rest periods afforded an opportunity to give subjects knowledge of results and also were designed to help alleviate inhibitory or blocking effects, which are sometimes noted on motor learning tasks when trials are massed (Bilodeau and Bilodeau, 1961).

Data from the performance period was recorded in twelve 20 second trials. There was some concern that larger units might mask the point at which the initially dominant response to make errors became subordinate, signalling (in Zajonc's terms) the end of the learning and beginning of the performance phase. However, these units are collapsed into one minute periods for analysis, as the smaller recording units provided no additional information.

Two measures of performance were taken; T.O.T. totals for each period (60 seconds) and variability. The latter was simply the mean of individual variances calculated for the nine 20 second trials of the testing session.

6.1.3 Method

The experimental design was a 2 X 2 X 2 X 2 factorial, factors being Social Condition, Instructions, Feedback and Sex. Subjects were 48 male and 48 female Open University students attending a week's summer school at Durham University; the approximate age range was 22 to 55. Participation was elicited in two ways; either via notices asking for volunteers and or being personally approached by the experimenter. Of these two methods the latter was by far the most effective and most subjects were engaged in this way. All testing took place between 1700 and 2100 hours.

The experiment was conducted in one of the rooms in the college in which the students were staying. This was an ordinary student's bedroom with the usual furnishings. The pursuit rotor was placed on a table and subjects stood behind this and faced into the room during the test. The only other equipment present was a blackboard on which scores were written in the public feedback conditions. The room was relatively free from noise and interruption from the corridor, since most of the residents were attending other functions during these hours.

Subjects reported to the experimental room at designated times and were assigned to experimental conditions.[1] On arrival each subject was

1. The assigning of subjects to conditions was initially done randomly, but whether or not an individual was tested alone or in a group depended largely on the availability of subjects at any particular time. Also, half way through the experiment it became apparent that conditions were not likely to be equivalent in terms of initial level of performance by continued random assignment. Therefore, to avoid problems with the interpretation of difference scores when initial performance levels are discrepant (Schmidt, 1972), subjects remaining were assigned to conditions according to their scores in the Practice trial.

read basic instructions for performing the task and was given a short demonstration. Following this they had one minute of practice with the task. The Practice period scores were used as the baseline from which difference scores in the experimental period were derived. The Practice period was the same for all subjects regardless of the experimental condition to which he had been assigned: no one received either auditory feedback or knowledge of results. In the group situations the entire group received the task instructions together but practised individually while the others waited outside the room. After the Practice trial subjects rested for approximately 3 minutes (although this period was more variable for individuals in groups) and were then read the experimental instructions.

Instructions:

Neutral instructions stated that the experiment was concerned with motor skills learning in relation to aspects of the task such as target speed and size. It was stressed that individuals' performance was unimportant since all scores would be combined and the mean results analysed. The ego-threatening or competitive instructions emphasized that the task correlated highly with performance on various sports and with driving a car: the experiment was concerned with individual differences and success on the task. The point was made that the individual's performance score was of primary importance since this would later be correlated with measures of his personality to determine what type of people were most successful.

Feedback:

In the feedback condition the auditory signal was briefly demonstrated, and it was explained that the subjects' T.O.T. scores would be written on the blackboard so that they could see them during the rest periods and keep track of how they were doing. The blackboard was in full view of the observing subjects. In the no feedback condition, subjects received neither the auditory signal nor knowledge of their own or other subjects' scores.

Social Conditions:

The experimenter remained in the room at all times in order to record scores and operate the apparatus. She stood approximately 4 to 5 feet from the subjects and either wrote the scores on the blackboard (feedback) or on a private scoring sheet (no feedback). The experimenter's presence was fairly obtrusive since subjects were stopped and started frequently and this involved switching the apparatus on and off. Subjects were tested either with the experimenter only present or with the experimenter and 2 other subjects. In the group situations, subjects were tested in turn, order being determined by the order in which individuals had performed during Practice (this involved no particular system). In conditions where feedback was given, the scores from the Practice trial were posted before the test so that subjects could see their relative group standing. Therefore, unlike Experiment III, subjects were not handicapped by the order in which they performed in terms of knowledge of their relative ability on the task. The first subject could presumably pace his performance based on his knowledge of the Practice trial scores. Observer subjects sat about 5 feet from the performing subject and to his side. From this position they were not able to see the performer's T.O.T. score when it appeared on the display panel. However, in the public feedback condition, the apparatus and performer were angled such that the display panel was visible to the observers.

Sex:

Subjects were not segregated according to sex in the group conditions. It was felt that if sex differences were to emerge this would be more likely to happen in mixed group situations where any conflicts over male and female roles in response to competitive stimuli would be highlighted.

The entire experimental session (including Practice) lasted approximately 10 minutes for single subjects and 20 minutes for groups. Five seconds before each one minute trial subjects were given a verbal "ready" signal by the experimenter and the target light was switched on. After being told to "begin" subjects started tracking the light. At the end of each trial the target was turned off and subjects were either shown their results for that trial or were told to relax for 20 seconds.

Following the test all subjects were issued a self-report questionnaire concerning their reactions to the experimental manipulations. This questionnaire took the place of a post-experimental interview. Considering the number of experimental conditions and the large number of subjects tested, it was felt that this might be a more objective and systematic method for assessing whether the manipulations were effective in creating stress in the manner intended. The questions and mean responses according to the condition of testing are presented in Appendix V.

Subjects were given copies of the TAQ and ASI, accompanied by the usual instructions and were told to return them to the experimenter as soon as possible. The EPI was completed as part of the students' practical work at the summer school and subjects were asked to attach these scores to their completed personality questionnaires. All subjects were debriefed following the completion of the self-report questionnaire. Table 6.1 summarises the experimental conditions.

TABLE 6.1

Summary of Experimental Conditions

Social Condition	Instructions	Feedback	Abbreviation
Experimenter only	neutral	absent	E
" "	neutral	present	E(F)
" "	ego-threat	absent	E-T
" "	ego-threat	present	E-T(F)
E and two other subjects	neutral	absent	Im-Comp
" "	neutral	present	Im-Comp(F)
" "	ego-threat	absent	Ex-Comp
" "	ego-threat	present	Ex-Comp(F)

6.1.4 Results

Figure 6.1 shows the mean learning curves for all conditions plotted for Practice and over each one minute trial during the test period. (When means from the 20 second intervals were plotted, the curves followed the same general pattern over time. Therefore the smaller units were combined so that the data corresponded to the discrete performance units in which subjects actually worked i.e. 4 one minute trials.) As can be seen from the figure, learning is most marked between Practice and Period 1. There is further improvement after Period 1, but the curves begin to level out and there is little change between Periods 2 and 3. The data can conveniently be divided into an initial and later learning phase at Period 1, which is analogous to a learning and performance phase, since after Period 1 the correct response is dominant ie. is occurring more than half the time. However, to call this later learning period 'performance' is slightly misleading since improvement is still taking place. Hence, the initial and later learning stages will be referred to as performance Phase I and II respectively.

For the purpose of analysis, performance during Phase I is represented by the difference scores for Practice and Period 1 (Period 1 - Practice). Phase II data are derived from the raw scores in Period 1 and 3 (Period 3 - Period 1).[2] The difference scores for Phases I and II are analysed in two separate analyses of variance so that the two phases can be discussed separately: this was judged to be preferable to a within subjects design in which differences between conditions would be reflected in an interaction with time periods.

No differences were found between conditions for mean variability $\left(\frac{\text{Sum of individual variances}}{\text{N for each condition}}\right)$, and therefore data for this measure are not presented or discussed. However means can be found in Appendix III.

Phase I

Figure 6.2 shows the mean change for all eight conditions during Phase I. As predicted subjects tested in the non-social conditions show greater performance gains during this phase: only subjects in E-T do not show such an advantage, their mean improvement being equal to that of Ex-Comp. The analysis of variance reveals a significant main effect for Social Conditions ($F = 6.24$; $df = 1, 80$; $p<.025$). The means for the social and non-social conditions collapsed over other treatments can be found in Table 6.2. Thus, the prediction that initial

---

2. Period 2 data is omitted from analysis since it differs little from Period 3 in terms of level of performance or performance change. Also, during the final period the dominant, 'correct' response should be well established.

Figure 6.1.

Mean T.O.T. (secs) as a function of Social Conditions and Time Periods

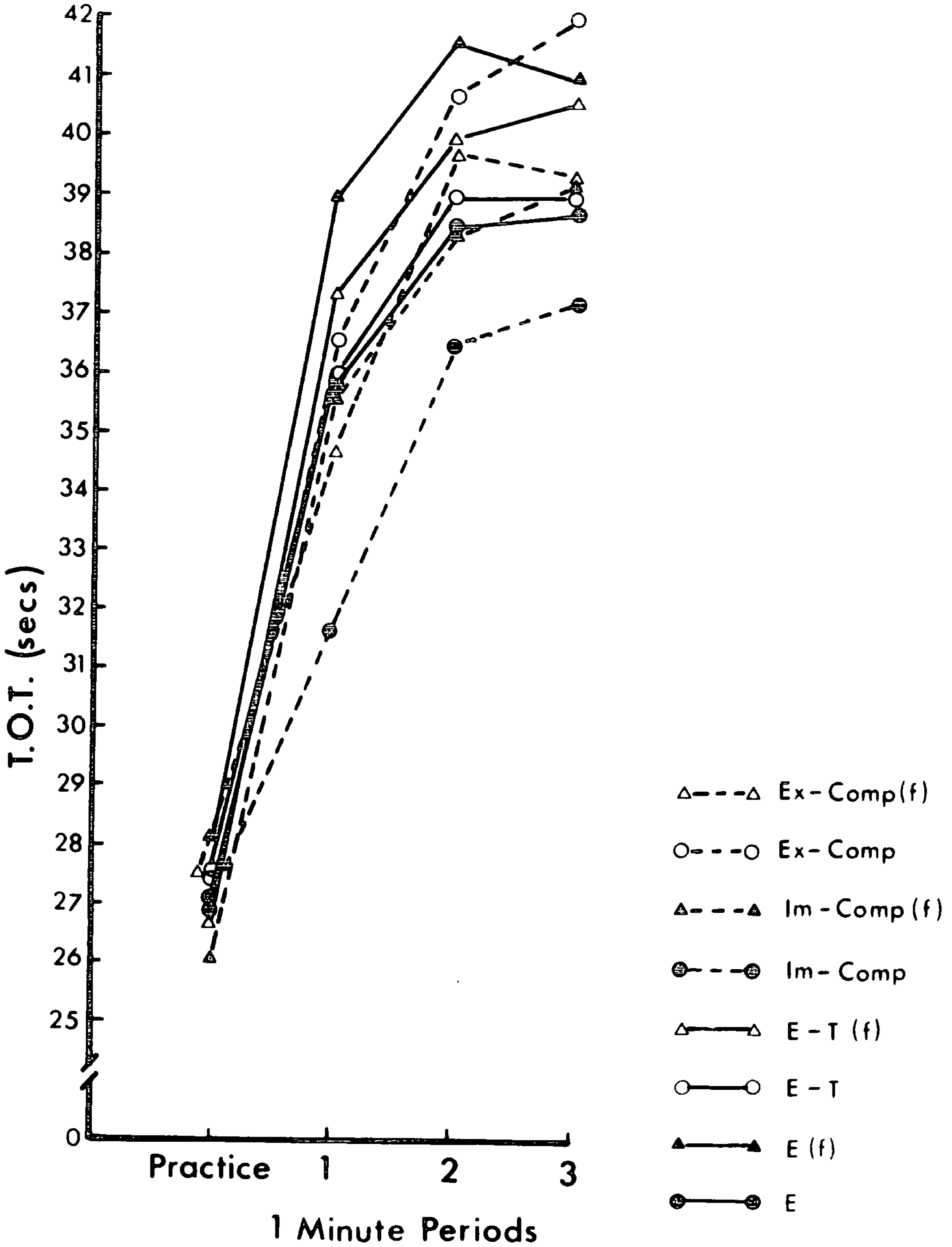
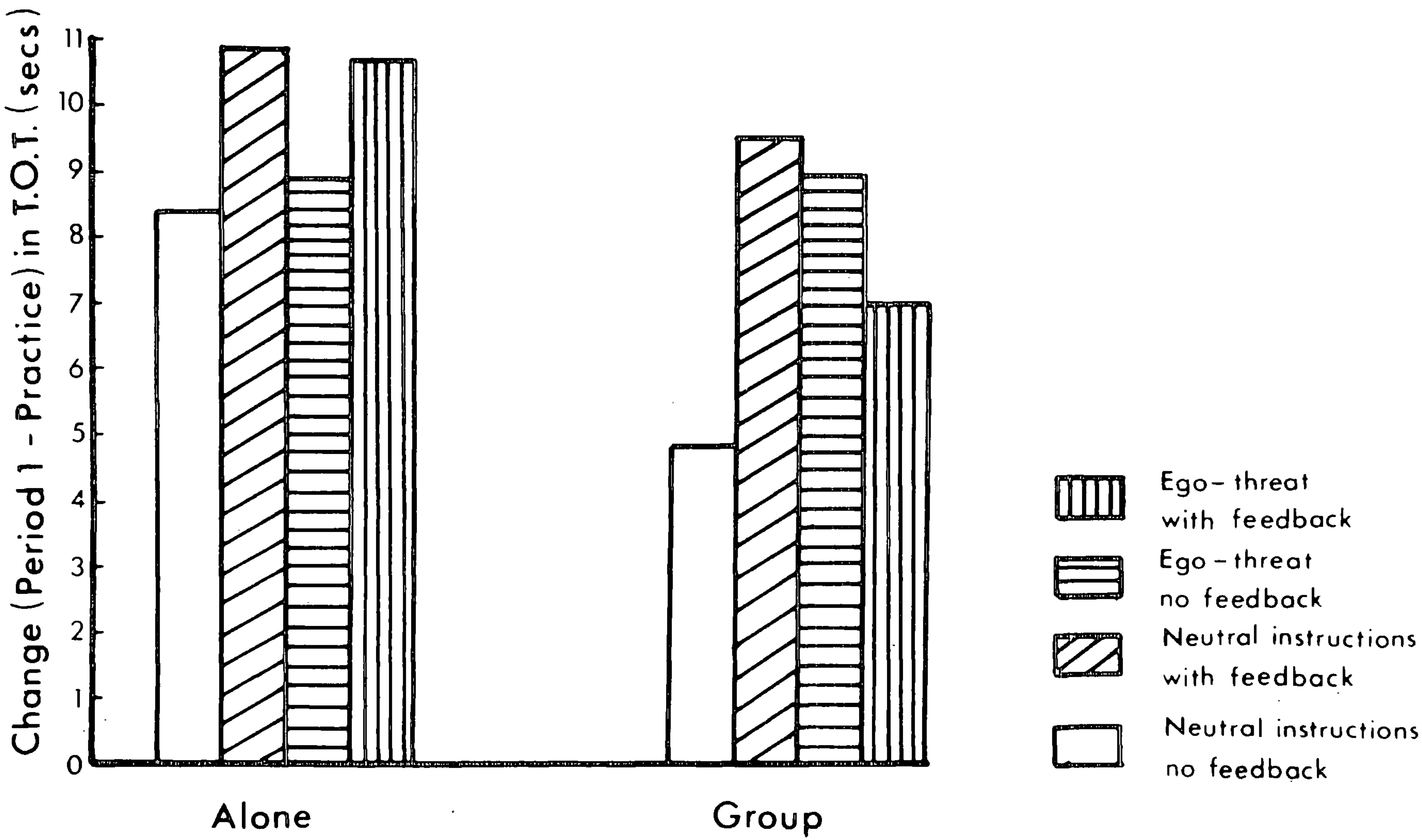


Figure 6.2.

Mean Change in T.O.T. (secs) for Phase I



learning would be impaired in the social conditions is supported.

Also, as predicted there is a significant main effect for Feedback ($F = 4.54$; df = 1, 80; $p<.05$), which is attributable to the relatively greater improvement shown by subjects in those conditions where auditory feedback and knowledge of results was given. (Means for Feedback and No Feedback collapsed over other treatments can be found in Table 6.2.)

TABLE 6.2

Single Treatment Means Collapsed Over Other Treatments (T.O.T.)

	Mean		Mean		Mean		Mean
E only	9.68	No Feedback	7.79	Neutral Inst.	8.41	Male	8.98
Group	7.64	Feedback	9.53	Ego-threat Inst.	8.91	Female	8.33

However, Table 6.3 below shows the means for the interaction of Feedback and Instructions, which also reaches significance ($F = 5.08$; df = 1, 80; $p < .025$).

TABLE 6.3

Mean Improvement According To Type of Feedback And Instructions (T.O.T.)

	No Feedback	Feedback
Neutral	6.6	10.2
Ego-threat	8.9	8.8
Mean	7.8	9.5

It can be seen from the table that although feedback has an overall beneficial effect, this effect is only present for subjects who are given neutral instructions ($F = 9.62$; df = 1,80; $p<.01$, for the simple main effect). Simple main effects also reveal that subjects with no feedback but ego-threatening instructions show greater relative improvement than those with no feedback and neutral instructions ($F = 4.11$; df = 1,80; $p<.05$). There is no difference between subjects receiving feedback, regardless of instructions.

The only other result meriting discussion is a marginally significant three-way interaction for Social Condition, Feedback and Instructions ($F = 3.41$; df = 1,80; $p < .065$). Reference to Figure 6.2 shows some differential effect of feedback and ego-threat for subjects tested alone and in groups. For subjects tested alone these variables are associated with relatively greater improvement compared to other alone conditions, but the same manipulations are only associated with moderate relative improvement when subjects are tested in groups. It is possible that the combined effects of the stress manipulations resulted in super-optimal arousal, leading to decrement. However, an alternative explanation is viable based on the data for pacing discussed below and presented in Table 6.4.

Table 6.4 lists the correlations between subjects' raw scores in Practice and their change scores for Phase I. The correlations for subjects in non-social conditions demonstrate regression to the mean, since pacing would not be possible in the absence of any sources providing relative performance information.

TABLE 6.4

Pearson Correlations for Individuals' Raw Scores in Practice and Phase I Change Scores

Condition	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Comp (F)
r =	-.26	-.26	-.35	-.11	-.51	-.32	-.18	-.63*

*p < .05

The only correlation in the table to reach significance is that for the Ex-Comp(F) subjects. The magnitude of the relationship is comparable to that of the pacing correlations obtained in Experiments III, IV and V. All correlations in the table are of course negative, but all (with the exception of Im-Comp) are substantially lower than that for Ex-Comp(F). The scatterplots for these values are shown in Figure 6.3. That pacing is indicated in Ex-Comp(F) suggests a partial explanation for the three-way interaction discussed above. Pacing would have the effect of suppressing the scores of more able performers while facilitating those of lesser ability. However, if the pacing tendency was more marked for the initially better performers than for the poorer ones (as was suggested in Section 6.1.1), this could suppress the mean for the whole condition. An examination of the improvement scores for the 5 best and 5 poorest performers (based on data from the Practice) indicates that this was the pattern that pacing followed in this Ex-Comp(F) condition. Table 6.5 presents the results from this analysis.

TABLE 6.5

Mean Change Scores for Phase I
for the Best and Poorest performers in Practice

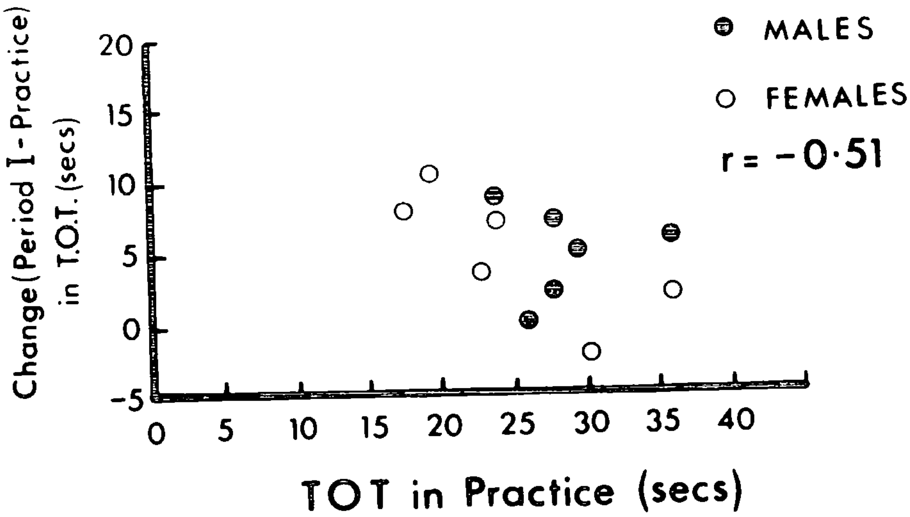
Condition	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Comp (F)
5 Best	6.6	9.3	9.1	10.2	3.5	8.0	7.2	4.0
5 Poorest	11.0	12.1	8.9	10.7	7.6	11.4	9.4	9.6

From the table it can be seen that the good performers in Ex-Comp(F) show the least improvement of any condition except Im-Comp. However, the poorer performers, although improving substantially are not particularly advantaged compared to other conditions. Clearly, pacing downward was more pronounced for good performers than was upward pacing for poorer

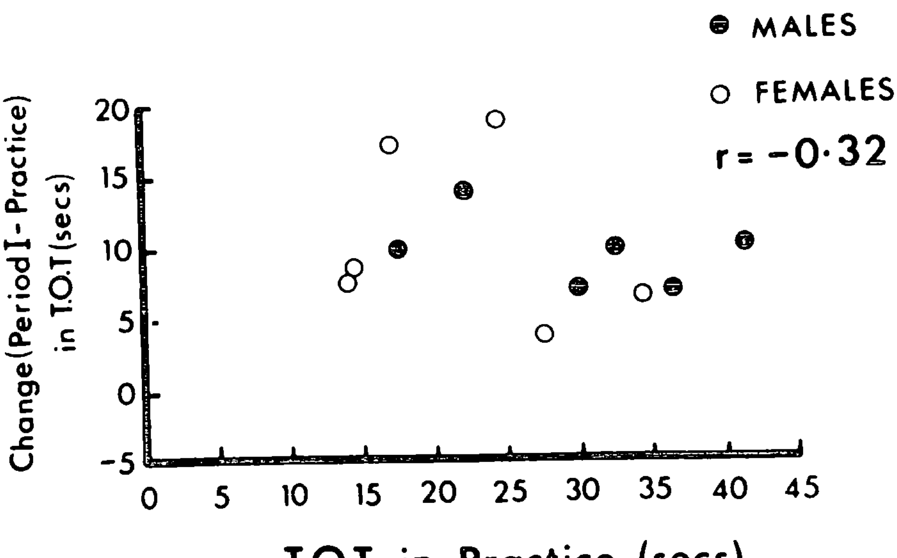
Figure 6.3.

Scatterplots for T.O.T. in Practice and Change in Phase I

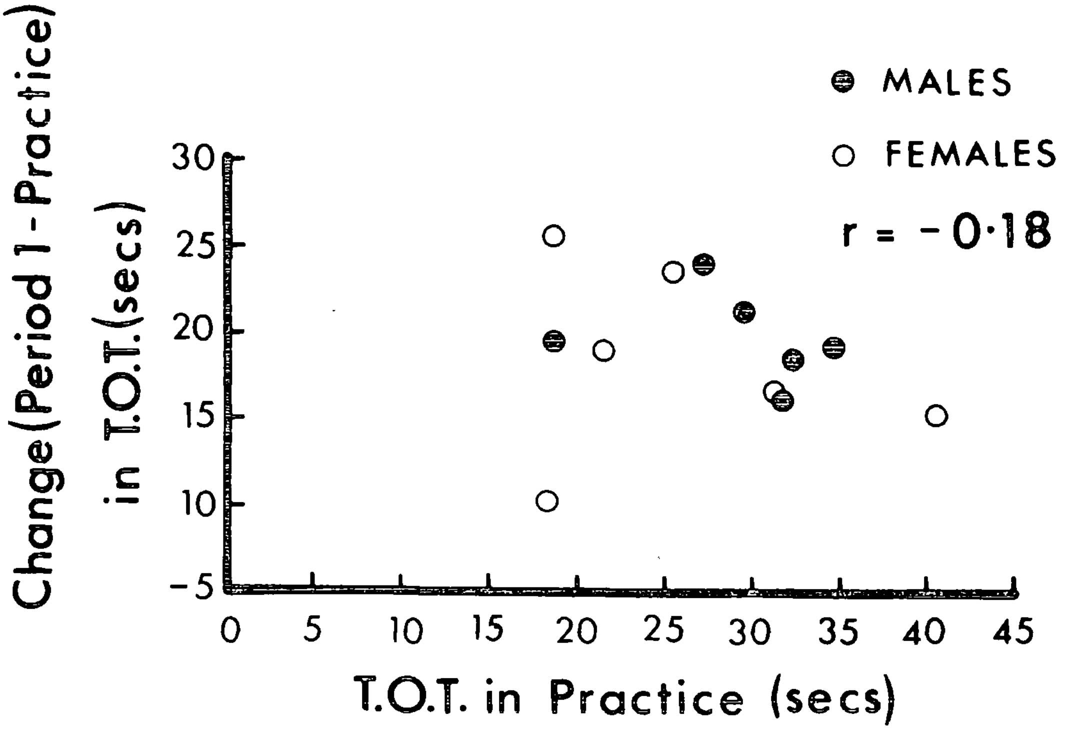
(a) Im-Comp



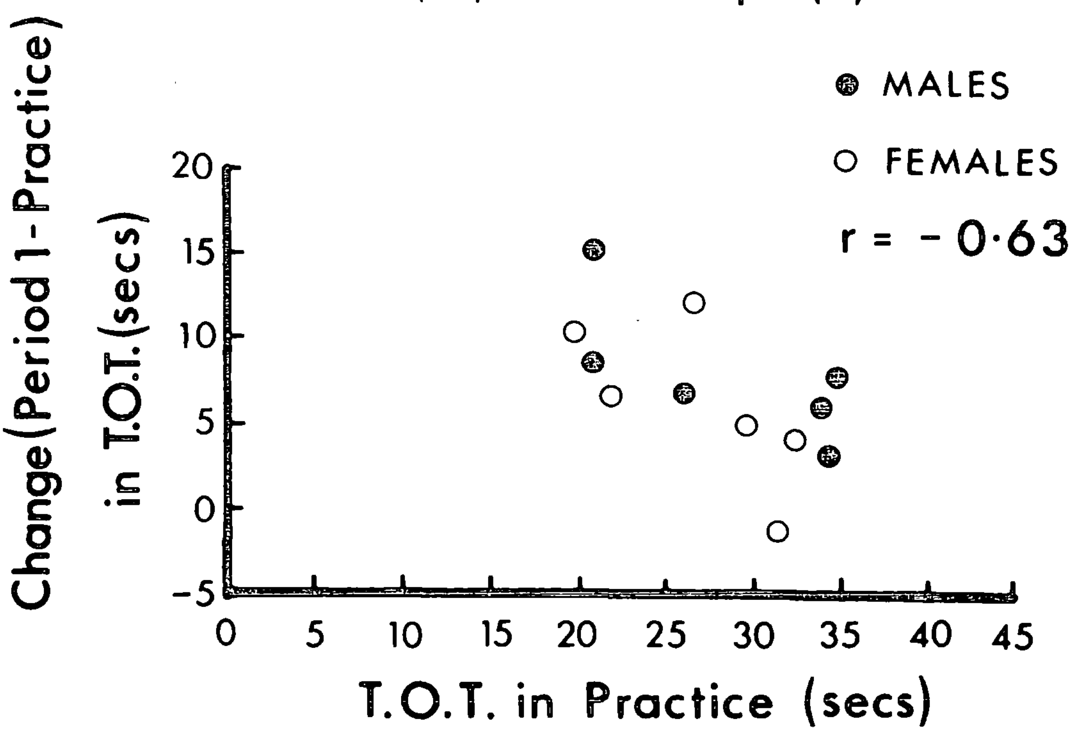
(b) Im-Comp(f)



(c) Ex - Comp



(d) Ex - Comp (f)



performers. This is in keeping with the expectation that subjects might not be able, because of the nature of the task, to pace themselves upward at will, meaning that what pacing did take place would necessarily be downward. It is interesting that no marked tendency to pace is indicated in Im-Comp(F), even though Feedback was public in this condition as well as in Ex-Comp(F).

Phase II

Figure 6.4 shows the mean change scores for the eight conditions of testing in Phase II. The figure suggests a consistent tendency for subjects tested in group conditions to demonstate greater relative improvement than those tested with just the experimenter present. However, the analysis of variance only just reaches significance for the Social Conditions factor ($F = 3.83$; $df = 1,80$; $p < .05$). Thus the results from Phases I and II are consistent with the predictions made in Section 6.1.1 (and with arousal theory) that in initial stages of learning the group situation would have an inhibitory effect, and that the reverse would be true in later stages of learning.

There are no other significant effects from the analysis of variance, though the interaction between Sex and Instructions approaches significance ($F = 3.19$; $df = 1,80$; $p = .08$). The effect can be seen in Table 6.6 below which shows the mean change scores for males and females in neutral and ego-threatening conditions.

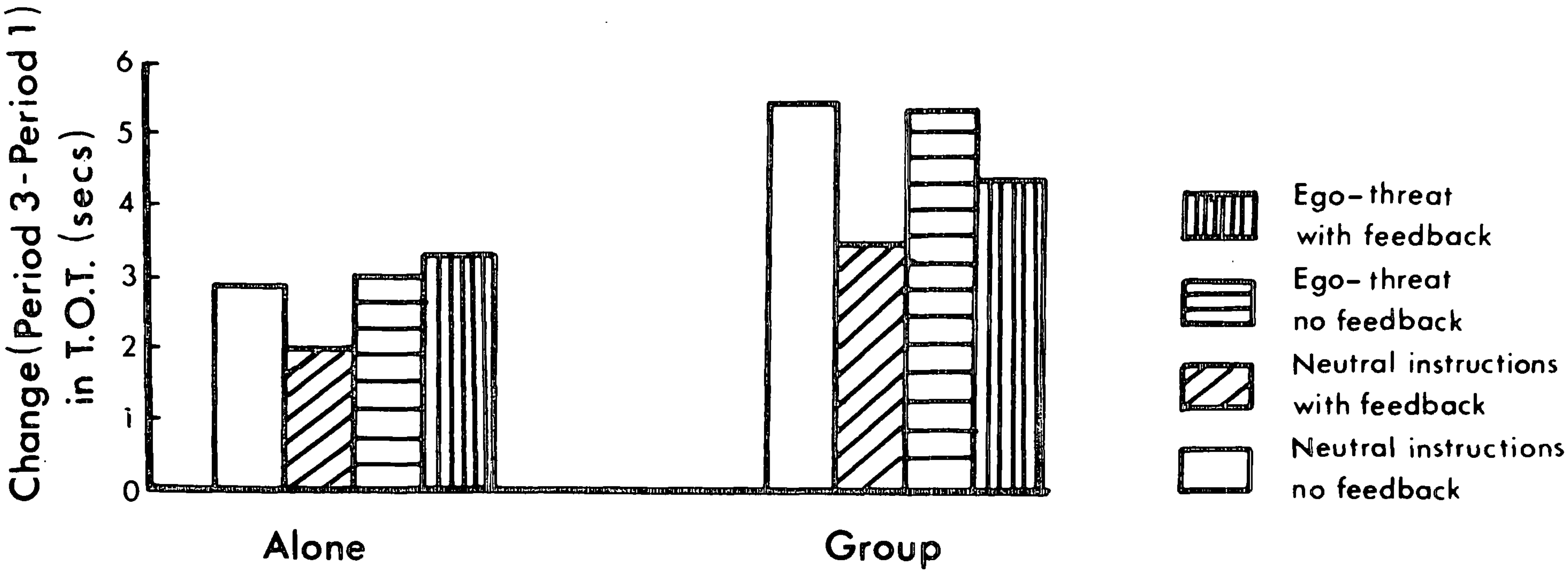
TABLE 6.6

Mean Change for Males and Females According to Instructions in Phase II (T.O.T.)

	Males	Females
Neutral	2.37	4.54
Ego-threat	4.71	3.38

Figure 6.4.

Mean Change in T.O.T. for Phase II



Females show greater relative improvement than males in the neutral conditions while males show an advantage and females a slight disadvantage with ego-threatening instructions. From previous experiments (IV and V) it would be expected that females would be impaired by the competitive instructions. However, there is no precedent to suggest that they would respond better than males in the neutral conditions. In any event the interaction does not reach conventional levels of significance and therefore any interpretation must be treated with caution.

No marked pacing effects were expected to occur in performance during Phase II since, in conditions where this pattern was expected, pacing in Phase I would already have substantially reduced the distance between individual performers. However, there was some interest in whether Im-Comp(F) subjects, who gave no evidence of pacing in Phase I, would show pacing later in the session. Therefore, correlations between the raw scores in Period 1 and difference scores for Period 3 (Period 3 - Period 1) are presented below in Table 6.7.

TABLE 6.7

Pearson Correlations for Raw Scores in Period 1 and Phase II Change Scores

Condition	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Comp (F)
r =	-.23	-.45	-.28	-.67*	-.02	-.17	.39	-.25

* p < .05

The correlations for this phase are less consistent than those found in Phase I. The range is greater and one of the values is positive. However, these values are confounded by the magnitude of correlations obtained in Phase I, since in cases where the correlations were high little further change in the same direction could be expected. In any event, in motor skill learning situations there is a marked tendency for a reduction in inter-subject variability over trials (Reynolds and Adams, 1954). It is interesting that in Table 6.4, on average, the non-social conditions demonstrate smaller negative correlations than the social conditions, while this pattern is reversed in Table 6.7. This further suggests that the values obtained for Phase II may be largely a function of the magnitude of the correlations obtained for Phase I. Thus, it appears that when pacing tendencies do occur in group situations and with this type of motor learning task, the effects are most pronounced, or at least more easily detectable, early in the performance period. Since Im-Comp(F) subjects show little evidence of pacing in either Phase I or II it is concluded that they did not use the feedback information to assess relative group standing, or, if they did, this was not accompanied by any tendency to match individual performance with the group average. Whether or not pacing occurs, at least on this task, would seem to depend on whether evaluation and/or competition are explicit.

<u>Self-report Data</u>:

The questions and mean responses for the post-experimental questionnaire are presented in detail in Appendix V. The questionnaire consisted of nine questions for single subjects and 10 for subjects tested in groups, which were all rated on a scale of 0 - 4 (not very stressing to very much so). The most striking feature of these data was that generally the mean ratings were rather low: only one question yielded means of more than 2.0 for most conditions, and this was concerned with subjects' interest in the task.

There is little evidence which would indicate that subjects were either defensive or reluctant to admit feelings of anxiety or distress, since defensiveness is considered to be apparent when an unusually low score is obtained for a fairly stressful situation (Boor and Schill, 1967): the pattern was rather one of marked consistency between conditions. It is felt that the ratings were probably fairly accurate representations of subjects' feelings. However, the questions were phrased in such a way as to emphasize the "worrying", "distracting" or otherwise distressing features of the experiment. Possibly subjects did not feel worried or distressed even though they may have found certain conditions more stimulating, interesting or slightly uncomfortable.

The most interesting and illustrative result from the self-report data was the uniformly high rating given by subjects to the question concerning basic interest in the task. It is possible that the task was involving to the point of rendering other environmental stimuli unimportant. Generally, subjects in groups gave higher ratings for those questions which specifically related to others' presence in the room, and subjects in Ex-Comp(F) gave the highest ratings on 4 out of 6 of these questions (nos. 2,3,5,7,8,10). However, it would appear that self-competition was the most important or at least the most recognized source of interest in the experiment, regardless of the conditions of testing.

6.1.5 Summary of Results and Discussion

Summary of experimental findings:

1. Subjects tested in group conditions showed less relative improvement in initial stages of learning (Phase I) than subjects tested with just the experimenter present.

2. Subjects tested in group conditions showed greater relative improvement in later stages of learning (Phase II) than subjects tested with just the experimenter present.

3. Feedback had a positive effect in Phase I, but this was confined to subjects receiving neutral instructions. When instructions were ego-threatening, feedback had no additional beneficial effects. However, of those subjects who received no feedback, individuals given ego-threatening instructions showed greater relative improvement than those given neutral instructions. Feedback had no differential effects on performance in Phase II of the testing session.

4. Pacing was evident in Phase I for Ex-Comp(F) subjects. Im-Comp subjects demonstrated a rather high negative correlation between initial level of performance and improvement, but this was not statistically reliable. Subjects in other group conditions did not show evidence of pacing which could be distinguished from regression to the mean.

The results from the present study provide support for the arousal theory prediction that the presence of others enhances the emission of dominant responses. However, initial learning proved to be more sensitive to the presence of others than did later learning. The results from Phase II show mean differences between non-social and social conditions which are actually quite small (means = 2.8 and 4.7 respectively). Furthermore, there was evidence that pacing was occurring in the Ex-Comp(F) condition. This was more pronounced for better performers, leading to some suppression of performance on the part of these subjects and resulting in a mean change score which was somewhat lower than that obtained in other conditions. This observation is rather important considering that the difference between grouped and alone subjects was greatest in this condition, therefore contributing substantially to the overall main effect for Social Conditions in Phase I.

Furthermore, Figure 6.1 reveals that 3 of the 4 non-social conditions show some decrement or no change between Periods 2 and 3, while 3 of the 4 group conditions show further improvement. That decrement in some conditions was occurring at this point in the testing session indicates that subjects may have been experiencing fatigue, loss of interest in the task or perhaps performance blocking, and suggests that the relatively greater improvement shown by subjects in group conditions during Phase II may have been due to greater perseverence rather than the facilitation of dominant responses.[3] Similar results were found by Kimble (1950) for pursuit rotor performance under competitive conditions. Therefore, although superficially the results support an arousal theory interpretation, it is obvious that alternative explanations are possible.

That performance blocking may have been a factor in this short testing session raises some interesting questions. For example, Eysenck (see Section 2.3.3) has argued that highly aroused subjects are able to tolerate inhibition (or blocking in this context) to a greater extent than those of lower arousal. If so, then the tendency for continued improvement noted for subjects in social conditions may suggest, indirectly, that these conditions were in fact more arousing. Generally, performance blocking can be directly assessed by providing a subsequent test on the apparatus and noting the magnitude of reminiscence (Feldman, 1964b; Reynolds and Adams, 1954). Such an experimental design provides a means of evaluating relative arousal levels which need not involve identifying dominant/subordinate response hierarchies.

---

3. There is some evidence from Figures 6.2 and 6.4 that the higher levels of performance shown by grouped subjects in Phase II may simply have been a function of their relatively lower improvement in Phase I because of ceiling effects. However, Figure 6.1 shows that none of the conditions should have been experiencing ceiling effects, since the maximum mean raw score obtained was only 42.0 out of a possible 60.0. Also, the decrement noted for some conditions in Period 3 does not bear any systematic relationship to the absolute level of performance in Period 2, implying that the decrements are not due to some conditions having reached asymptote.

The role of the task is felt to be of major importance in the present experiment. According to subjects self-report data, the pursuit rotor generated a high level of intrinsic motivation. Considering this, it is somewhat surprising that the expected pattern of performance change between social and non-social conditions emerged at all. As mentioned above, initial learning would seem to be most sensitive to the social manipulations. It was also during this phase that feedback and instructions had differential effects. However, with increased familiarity, task performance would necessarily become more stable and more resistant to influence from environmental stimuli. Thomas and Halliwell (1976) found that after five 20 second trials on the pursuit rotor, performance of their subjects continued to improve but the standard deviations levelled off, suggesting a stabilization of performance at this point. Therefore, the effects of others' presence, once performance had more or less stabilized, was perhaps minimal compared to subjects' general task-oriented motivation, and this may account in part for the less striking results obtained for Phase II. It is generally acknowledged that performance becomes progressively more resistant to influences in the environment with increasing task skill (Singer, 1968; Wilkinson, 1969).

It is noteworthy that pacing was apparent early in the learning period in that condition in which performance was public and evaluation was stressed (Ex-Comp(F)). Thus, pacing tendencies appear to be a source of performance variation on tasks which are classified as predominantly learning, which have a definite response hierarchy and which are not self-paced. However, unlike previous experiments, in the present case pacing only occurred when evaluation was made explicit. Although feedback was associated with better performance, it apparently did not lead to uncertainty about performance level and subsequent comparison in the same way as did explicitly evaluative instructions. Furthermore, pacing was evident only for those subjects who were of initially high ability, probably because the task effectively limited the individual's control over his performance. Many standard experimental tasks fall into this category (eg. serial and paired-associated learning), and are, therefore, potentially subject to similar effects, at least if inter-subject comparisons are in any way possible.

Unexpectedly, none of the evaluative manipulations (either public feedback or ego-threatening instructions) interacted with the social conditions. Although the combination of feedback and ego-threat did appear to facilitate pacing tendencies, subjects' relatively low ratings on the self-report questionnaire suggest that the evaluation potential involved in this experiment was not very great.[4] Still, if these situations were not construed as evaluative by subjects, it is difficult to explain how other reports have found significant interactions when manipulating only instructions for fairly meaningless tasks. The type of social situation is no doubt important. In most cases where such interactions have been found the social manipulation has been that of 'passive audience'. The social settings here were potentially more complex and the effects of evaluation may have been masked or overshadowed by other social processes.

Finally, some comment is needed concerning the absence of any sex differences in this experiment. There were no significant main effects or interactions due to sex, and a perusal of the data revealed no tendency whatever for females of initially high ability to show relatively less improvement than males in any of the testing conditions

4. The fact that variability was unaffected by the experimental manipulations further suggests that none of the situations were very stressful, or were certainly not stressful to the point of producing any performance disruption.

(as was found in Experiments IV and V). However, an examination of the raw data showed that females in every condition and in each of the 4 periods demonstrated a lower mean performance than corresponding males. Whether this is the result of innate differences or differential degrees of previous experience with perceptual-motor tasks is not an issue here: such sex differences in pursuit rotor performance have been previously observed (Costello et al, 1964). But, the fact that females did perform at a lower level throughout may have meant that the majority of them were never actually confronted with success, and avoidance behaviour would, therefore, not be expected.

It was felt both from the experimenter's observations during the testing period and subjects' self-report data that the pursuit rotor was in fact a good task for involving subjects and stimulating interest in performance level. However, the actual testing session was very short, necessarily so because of possible complications with fatigue and performance blocking with accummulated practice. As a result, there was little opportunity for subjects to develop much inter-individual competitiveness. Also, the task demanded virtually all the subjects' attention during performance and afforded no opportunity to observe others' reactions or to engage in any form of communication. A self-paced task and one on which performance proceeded over a longer time period may have been more effective in creating competitive/evaluative stress between subjects, which in the present experiment appeared to be minimal.

6.2 EXPERIMENT VII - DARTS

6.2.1 Aims

This final experiment examines the effects of inter-subject competition and performance in the presence of an audience using a realistic task and performance situation, while still enabling laboratory control to be exercised. Previous experiments (III, IV, V) indicated that social comparison and subsequent pacing with others can be important in determining performance outcomes in social situations when tasks are novel and self-paced. In Experiment VI the presence of others affected performance in a way consistent with arousal theory when performance could be analysed in terms of response hierarchies, although between subject pacing in the social conditions appeared to contribute to some degree to the overall results.

In these experiments the effects of social-evaluative stimuli on pacing tendencies was unexpected. Although various theories predict that emphasizing individual performance level or creating uncertainty about the acceptable standard can lead subjects to focus more directly on their own level of performance or seek comparative information from others, the pacing patterns which emerged were contrary to intuitive beliefs about the effects of social-evaluative stimuli. In some cases the introduction of evaluative cues led to a lower level of performance by the more able performers and an overall mean which was lower or not substantially different from that in non-evaluative situations. It could be argued that super-optimal stimulation might have led to performance disruption for the initially more highly motivated (and better performing) subjects. However, the pacing patterns suggested a systematic gravitation toward the group mean, not performance disruption as such. Furthermore, there was little evidence to suggest that motivation or arousal, as reported by subjects, was ever high enough to actually have a disruptive effect.

One consideration which comes to mind regarding these results is that the situations created thus far were not competitive in the same way as are real contests or games. The experimental situations themselves were somewhat ambiguous and the introduction of evaluative stimuli may only have made them moreso. Also, for the most part, the tasks only utilized a narrow range of responses, limiting the ways in which subjects could do the task or vary their performance strategy.

In light of these considerations, a final effort is made to examine the effects of mere presence and social-evaluative stimuli on performance in social situations by asking subjects to participate in an experiment involving darts throwing.

Darts throwing was chosen because it is an activity with which all subjects would have some previous experience and would associate with competitive situations. Also, the task itself is completely self-paced: there is no imposed time period and subjects are free to try to improve their performance via altering their throwing style in a variety of ways. Furthermore, dart throwing ability would be expected to be stable in that once subjects had had some minimal experience additional improvement would only become apparent over many repeated testing sessions. Thus, practice effects were not expected to interact with the experimental manipulations.

It is difficult to predict how the presence of others would affect performance on a task of this nature. Firstly, there is no clear response hierarchy, therefore predictions based on arousal theory are inappropriate: neither is the task of a routine nature such that speed/accuracy can be measured. A variety of performance outcomes are feasible, e.g. (1) mean performance may show overall disruption or facilitation, (2) performance may be unaffected because subjects have stable playing styles which are impervious to mild social stress or (3) the task may be too sensitive and individual reactions so variable that no overall patterns would be detectable.

In addition to making predictions about darts playing in conditions of social stress, there is further difficulty in deciding how best to measure the quality of performance. A few reports exist concerning the effects of audience presence on real performance tasks, but these are by and large unsatisfactory. For example, Elliott and Bartee (1976) found no differences in 10 pin bowling when subjects were exposed to stress via video-recording. However, only the final score was used as the measure of performance: this is a gross indicator of overall performance and is biased by the scoring procedure which gives bonus points for certain events such as knocking down all 10 pins. Gymnastic performance (Paulus et al, 1972) and gross balancing (Singer, 1965; Smith and Crabbe, 1976: Kozar, 1973; Fielding et al, 1976) have been examined in roughly the same way. None of these tasks (or methods of performance analysis) offer a very sensitive performance situation. Overall or total score measures may even give unrealistic accounts of a subject's performance throughout the entire game or test period.

On the other hand, darts throwing is a complex perceptual motor task involving a high level of skill and control. It provides measures of both variability and accuracy. High levels of stress would be expected to result in performance disruption, particularly in terms of variability (Fitts and Posner, 1967; Welford, 1976). In fact, one measure of a good darts player is his consistency, both between different situations and over prolonged competition. Therefore, in assessing the effects of social situations on darts throwing, consistency or variability, is judged to be the most sensitive indicator of good performance.

In order to examine the components of performance, separate measures for the horizontal and vertical dispersal of darts were developed. In this way not only can directional trade-offs and changes in performance strategy be detected, but should disruption occur, the specific nature in which skilled performance is affected can be explored. For example, since darts throwing mainly involves arm movement in the vertical plain, disruption should be reflected more in the dispersal of darts along the length rather than the width of the board. However, under high levels of stress, hand steadiness may be affected and dispersal in the horizontal plain may increase as well. Which of these measures best predicts overall accuracy (or nearness to the target) from the subjects' frame of reference is also a point of interest.

Subjects were tested in a variety of conditions; a pre-stress (baseline) period, experimenter plus video-recording (audience), a post-stress (identical to the baseline) and in inter-subject competition. Video-recording is considered to be an audience manipulation and has been used as such by several authors (e.g. Putz, 1975; Droppleman and McNair, 1971). Although the experimenter was also present, previous experience suggested that she alone would probably not constitute a very potent audience manipulation. The post-stress condition provides a check for any carry-over effects from the stress condition, as subjects can at times maintain stable performance under immediate stress, due to some compensatory process, but show disruption following its removal (Allport, 1924; Sanders and Baron, 1975). Should a post-stress decrement in performance occur, it would give some indication of the intensity of the stress manipulation.

Unlike previous experiments, in this case subjects are matched according to ability during the competitive session. This procedure is intended to minimize pacing tendencies, thereby increasing the likelihood of identifying disruption or facilitation. Also, 'real' competitive partners are usually somewhat matched, and equality of skill between opponents is generally considered to be the best circumstance for eliciting between subject competition (Caillois, 1961).

Another procedural difference in this experiment is that subjects report on three different occasions, twice being video recorded and the third time for the competition. A point of interest is whether stable patterns of performance change will emerge in response to the social manipulations when the conditions are repeated. The repeated testing also allows for some habituation to the novelty of the experimental situation before the final 'competition'. If marked differences are observed between the first two testing days some effect due to novelty on the first day might be inferred. In such a case the results from the second day session would be regarded as a more reliable baseline with which to compare the third day (competition) test.

6.2.2 Task

The equipment consisted of a standard competition dart board and 3 sets of darts, differing in weight (22, 29 and 36 gm). Subjects were allowed to bring their own darts if they wished, provided they used the same set throughout the experiment. A spotlight was mounted on the ceiling approximately 4 feet from the board and lines were marked on the floor at 7'6" for males and 7' for females. The room in which the experiment was conducted had been originally designed as a children's playroom and was equipped with several games and various play materials: a dart board was quite in keeping with the general furnishings and decor of the room.

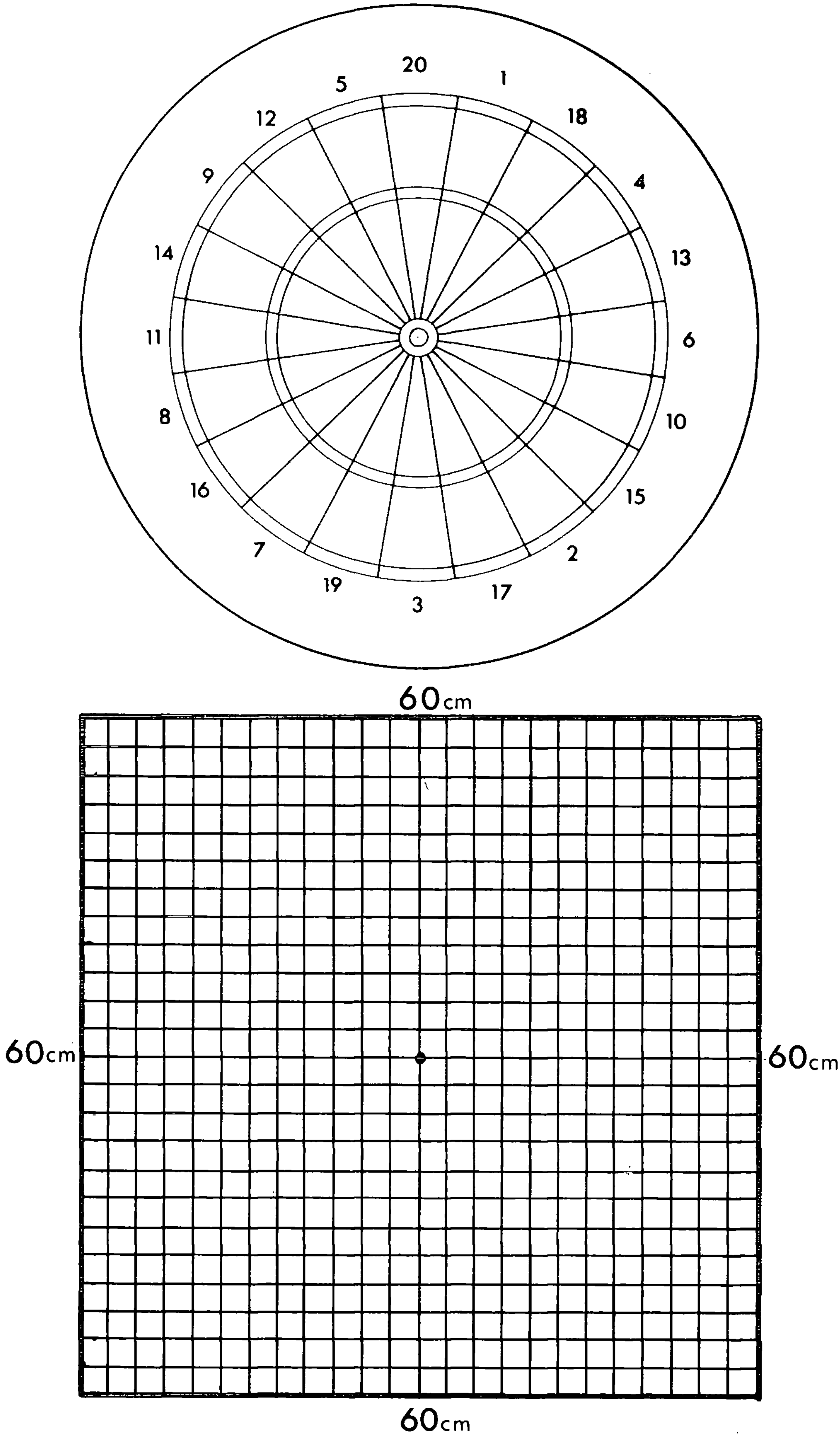
Scoring:

Unlike the usual variety of dart games, subjects were instructed to aim for the centre 'bull', and good performance was therefore defined (for the subjects) as proximity to the centre of the board. The bull was chosen as the target mainly to facilitate scoring, but also because it provided a standard target from which subjects could assess their accuracy in any direction. Although aiming for the centre is not generally the procedure in darts competitions, it is characteristic of most aiming tasks, e.g. archery and shooting, and is a realistic way to demonstate and evaluate dart throwing ability.

Since during the baseline and post-video conditions the experimenter was not in the room, subjects were responsible for recording their own scores, and to avoid confusion this same procedure was followed in all conditions. Each subject was issued with scoring sheets containing diagrams of dart boards drawn to scale (approx. 1:4) and was asked to plot the approximate position of each dart on the scoring minature. Only three points were plotted on each diagram, providing a separate record for each trial (three darts). This was a fairly easy procedure for the subjects because of the divisions on the diagram (which is reproduced in Figure 6.5). However, as there was some suspicion that reports might be less accurate when the experimenter was not in the room, each subject was observed from behind a one-way mirror in initial sessions of the experiment and scores were simultaneously recorded with the subject. Although there were some discrepancies due to the experimenter being at a different angle to the board than the subjects, there was no indication whatever that any subjects attempted to give an unjustly positive record of their performance. Subjects were observed during all sessions in which the experimenter was not actually in the room and all appeared to record their scores conscientiously, taking the scoring sheets up to the board and recording the points before removing any darts, as they had been instructed to do.

Figure 6.5.

Scoring Diagram and Perspex Grid for Recording and Measuring Performance (Darts)



Performance Measures:

Three measures are used to describe subjects' performance:

M - mean distance from the target

Var(x) - horizontal dispersal of darts

Var(y) - vertical dispersal of darts

The last two measures are estimates of variability. M represents the total distance (in millimetres) of all points from the centre of the board, irrespective of direction, divided by 24 (the total number of darts thrown in each condition).

The calculation of the different dispersal scores is somewhat more complicated. A perspex grid (Figure 6.5) with x and y axis intersecting at (0,0) was superimposed on the scoring diagrams, the zero point corresponding to the bull. Coordinates can then be recorded according to each dart's position in the corresponding quadrant. Since Var(x) and Var(y) are both measured this necessitates assigning each dart coordinates along the x and y axis of the grid. The mean coordinates for each set of three darts can then be determined and the dispersal scores are based on the separate variances of the x and y coordinates around the respective mean coordinates. The subject's mean dispersal in each dimension is the average of the variances from each 3 dart trial. The dispersal scores for each subject are calculated in this way and averaged over individuals to obtain the mean scores for each condition of testing. For all measures, a dart landing in the backboard was disregarded, since it was impossible for subjects to plot such a point, and the score for that trial was determined by only 2 darts.[1]

1. Some method for determining the position of these darts could probably have been devised had subjects not been scoring themselves. However, since all subjects were experienced and the target was the centre rather than a point on the edge of the board, darts landing in the backboard were expected to be (and proved to be) rare events.

6.2.3 Method

The experiment involved 20 subjects (10 males and 10 females) who reported on three different occasions. Subjects were undergraduates and were engaged mainly through personal contact by the experimenter. Only subjects who had had some experience playing darts were considered to be suitable, since novice players would show considerable initial improvement. A summary of the testing conditions appears below in Table 6.8.

TABLE 6.8

Summary of Experimental Conditions and Procedure

Day 1 (Aud 1)	Day 2 (Aud 2)	Day 3 (Ex-Comp)
Baseline 1	Baseline 2	Baseline 3
Video 1	Video 2	Video/Competition
Post-Video 1	Post-Video 2	

All subjects were tested between 1300 and 2000 hours. On Day 1 (Aud 1) subjects were tested under three consecutive treatments, as indicated in the table. The same procedure was repeated on Day 2 (Aud 2). However, Day 3 (Ex-Comp) involved only two conditions; Baseline 3 and Video/Competition.

Aud 1:

Subjects reported to the experimenter's office individually and, following the completion of a short subject information form, were taken to the experimental room. Initially they were told that the experiment was concerned with the relationship of throwing styles to personality factors. Subjects then chose a set of darts and were allowed a short warm-up (9 - 12 throws). The instructions for the task were then read by the experimenter; these included where to stand, what to do if a dart bounced off the board, etc. It was pointed out that no time limit was imposed and subjects could vary their throwing style in any way they pleased. The experimenter then demonstrated the scoring procedure and subsequently asked the subjects to score a sample diagram. When it was clear that the instructions and scoring procedure were understood, instructions for the Baseline condition were read.

Baseline 1 was described as a practice session, although it was emphasized that subjects should try to throw and score their darts as accurately as possible since it was likely that the routine they established during this period would carry over to some degree to the experimental session. The experimenter explained that she would be leaving the room but that subjects would be signalled when to begin by two flashes of a light mounted next to the door: this enabled the experimenter to get into position behind the one-way screen.[2] Subjects threw 24 darts during this period and in each subsequent condition on every testing day. When the subject had completed the final scoring diagram the experimenter waited an additional 2 minutes and then knocked on the door, asked if the subject had finished and re-entered the room.

At this point subjects were issued with a fresh set of scoring sheets and were told they were now ready to begin the experiment (Video 1). The instructions for this part of the experiment explained in more detail how personality might be related to throwing styles. It was emphasized that these styles were characteristic for individuals over many tasks involving eye-hand coordination and the experiment was

---

2. All conditions on Days 1, 2 and 3 were timed by the experimenter. Times were calculated from the moment a subject took aim behind the line to when he finished the final scoring diagram.

specifically concerned with which types of people and which styles of playing were most successful. However, in order to study these complex patterns of movement in detail subjects were going to be video-recorded. To further stress evaluation, it was pointed out that the tapes would later be analysed by a group of interested students who would classify the different individual styles of playing.

Following the reading of these instructions, the experimenter pointed out the video equipment, which had been standing in a corner of the room, pulled it forward and focused the camera on the subject. The television monitor was placed within the range of subjects' peripheral vision. The experimenter remained in the room both to operate the equipment and to provide an additional source of audience stress. Subjects were told to begin when they pleased after the recorder had been turned on. None of the subjects had been video-recorded at any prior time.

When the subjects had finished the final scoring diagram, a fresh set of scoring sheets were issued and subjects were asked to throw 24 more darts in order to check for any effects due to practice or fatigue over the experiment. The experimenter left the room during this period and subjects proceeded exactly as they had during the Baseline.

Aud 2:

One to two weeks later subjects were retested in the same conditions as in Aud 1. The only difference in procedure was that the instructions were simply reviewed and no demonstration was given of the scoring procedure. It was explained that the procedure was being repeated in order to ensure that sufficient data had been collected. The Baseline in this case was of course not a true baseline in the same sense as that of Day 1 since subjects suspected they were going to be recorded afterward.

Ex-Comp:

One to two weeks later subjects reported for the final experimental session. In this case two subjects, matched for ability, were scheduled at the same time. One was taken directly into the experimental room while the other was asked to wait in an adjacent room and complete two personality questionnaires (TAS and ASI). The former subject repeated the Baseline exactly as in Aud 1 and 2. Following this the subjects changed rooms and did the alternative activity.

When both subjects had finished the Baseline they were brought together and it was explained that they were to engage in a simplified darts competition. Once again they were video-recorded in order to see 'Whether individuals used any different strategies when faced with a real competitor". In order to emphasize competition, subjects were asked to record their opponents scores rather than their own and were thus forced to attend closely to the other's performance. In addition, a simple numerical scoring method was introduced to add interest to the game. Concentric rings were given numerical values (25, 15, 5, 4, 2, 1) and subjects were instructed to keep a running total of their opponent's score on the scoring sheet and to announce it aloud so that it would be recorded on the video-tape. The competition was described as the most important of all the taped sessions since it was most similar to a real game. Following this final session the subjects were debriefed and the experimental manipulations discussed. No Post-Stress session was conducted since both subjects could not perform at once and the data would have been confounded had a delay been interpolated for half the subjects.

Only single-sex pairs were used during competition in this experiment.

Although the effects of a male-female competition or some comparison between single and mixed-sexed pairs might have been interesting, it was decided that in this exploratory experiment the simplest combination was most advisable. In addition, matching would have proved difficult since most of the males were more skilled players than the females.

6.2.4 Results

In this experiment, rather than use the difference score to depict change in performance, ratios were calculated between the separate Baselines and corresponding experimental treatments on each day. The ratio indicates the subject's relative change and is judged to be a more meaningful index of change because of the rather great range in subjects' skill. Since each experimental treatment is divided by the Baseline for that day of testing, improvement in either M, Var(x) or Var(y) is reflected in a ratio of less than one, while a ratio of more than one signifies a decrement. Because the resulting distribution is skewed to the right, the ratios have been further transformed to logarithms: these transformed scores approximate a normal distribution with the mean at zero and are thus suitable for parametric statistical analysis.[3]

Initially only Aud 1 and 2 data are examined in order to identify any effects due to novelty or practice between days or between the Video and Post-Video conditions. Ex-Comp is omitted since it did not employ a Post-Video condition. The data are analysed in a 2 X 2 X 2 X 2 design with one 'between' factor, Sex, and three 'within' factors, Social Condition (Video, Post-Video), Days and Axis of Dispersal

3. The difference score may have biased the overall results due to the large changes in absolute performance shown by some of the less skilled players. The distribution of ratios necessarily is skewed to the right since improvement can only be reflected in values between 1.0 and 0.0, whereas decrement can range between 1.0 and 60.0 (the distance between the bull and the edge of the diagram in millimetres). This is analogous to the situation found in signal detection theory experiments where the criterion (Beta), also a ratio, is often transformed logarithmically to provide a more meaningful index of bias (McNicol, 1972).

(Var(x) and Var(y)).

This analysis reveals no significant differences for any of these factors. Therefore, it can be concluded that no practice or novelty effect was in evidence either within the experimental session (Video to Post-Video) or between Days. Neither were there any detectable carry-over effects from Video to the Post-Video. Furthermore, males and females did not respond differently to the experimental manipulations, and there was no discernable trade-off or change in performance for the two dispersal measures either in response to the stress treatments or between days.

A perusal of the data for all the performance measures shows great variation in the ways in which individuals respond to the social stressors.[4] Most of the mean differences are negligible because the number of subjects who improved in any one condition roughly equals the number who became worse. One sample t tests conducted on the means between conditions on the separate testing days reveal no significant differences.

Although there are no mean differences between the experimental conditions within days of testing, there are differences in the patterns of performance change between days. Namely, there is a tendency for subjects to respond differently to the Video and the Video/Competition on the variability measures: variability shows a decrease in Aud 1 and 2 under the Video treatments but increases in Ex-Comp under the Video/Competition. Var(y) reflects this pattern most markedly. In fact the difference between subjects' performance on this measure for Video 1 and Video/Competition is significant ($t = 2.12$; $df = 19$; $p < .05$) and nearly significant for Video 2 and Video/Competition

4. It should be pointed out that none of the performance measures are independent. M correlates with Var(x) and Var(y), .65 and .45 respectively. Similarly there is a low positive correlation between Var(x) and Var(y), $r = .25$.

(t = 1.72; df = 19; $p<.1$). However, these results are somewhat confounded by the different Baseline levels on each day of testing. Table 6.9 below shows the changes in absolute performance in these three conditions. Although subjects under Video/Competition become more variable (have greater Var(y) scores) than either of the Video conditions the significant difference obtained between Video 1 and Video/Competition may be partly due to the relatively larger Baseline 1 mean.

TABLE 6.9

Raw score means for Var(y) during Baseline and Video Conditions

	Baseline	Video	Difference
Aud 1	167.16	129.52	-37.64
Aud 2	121.00	103.85	-17.15
Ex-Comp	122.26	135.30	+13.04

The only other data available for analysis are that of the time scores for Aud 1 and 2. Ex-Comp is not comparable since subjects scored for each other, somewhat reducing the total time for each individual. The times for Aud 1 (shown in Table 6.10) indicate that subjects become progressively faster over the experimental period, although only the difference between Baseline 1 and Post-Video 1 is significant (t = 3.39; df = 19; $p<.01$). This is likely due to subjects becoming better at using the scoring sheets, the progressively shorter times reflecting increased ability to locate the corresponding points on the scoring diagrams more efficiently. However, the pattern in Aud 2 is somewhat different: subjects become faster in Video 2 but slow up when the stress is removed (Post-Video 2). The difference between Baseline 2 and Post-Video 2 is not significant (unlike Day 1), indicating that no progressive change in speed occurred over the experimental session.

However, the difference between Baseline 2 and Video 2 is significant ($t = 2.90$; $df = 19$; $p<.01$), suggesting that the faster time shown in Video 2 may be a response to the video recording rather than representative of further learning. The absolute difference between Video 2 and Baseline 2 is greater than that for Video 1 and Post-Video 1 which argues against a progressive learning effect; so does the slower pace in Post-Video 2 (compared to Video 2) although the times for these two conditions are not significantly different ($t = 1.63$; $df = 19$; $p>.05$).

TABLE 6.10

Mean Times for Baseline, Video and Post-Video for Aud 1 and Aud 2

	Baseline (seconds)	Video (seconds)	Post-Video (seconds)
Aud 1	390.21	370.21	356.21
Aud 2	355.45	336.85	349.30

As mentioned briefly above, individual differences in the response to the stress manipulations in most cases negate any overall effects on the three testing days. However, it is unclear whether these individual reactions are systematic within individuals or are simply due to random variation in task performance. Therefore, the relationships between the change ratios for all measures, both between and within testing days, are examined. Var(x) and Var(y) are chosen as the measures most likely to show inconsistency in individual playing styles.

On all three days there is a low positive correlation between the two dispersal measures ($r = .22$ to $.34$), suggesting that, rather than a directional trade-off, subjects show some tendency to become either better or worse on both measures. However, there is virtually no correspondence between the individual patterns of change in the stress conditions on different days.[5] The correlations for Video 1 and 2 (r = .1 for Var(x) and -.1 for Var(y)) indicate that the individual's response to the video recorder in Aud 1 bears no relationship to his response in Aud 2. (Correlations between Video/Competition and Aud 1 and 2 are r = -.2 and -.03 for Var(x) and -.38 and .17 for Var(y)). However, the pattern of change within days is more consistent for individuals. Correlations between Video 1/Post Video 1 and Video 2/ Post Video 2 are .66 and .59 respectively for Var(x) and .48 and .62 for Var(y).

The inconsistency between days could easily be explained if individuals showed markedly different baseline levels of performance on different days. If for example, a subject demonstrated much better performance during Baseline 2 than Baseline 1, for whatever reason, then his performance during Video 2 relative to Baseline 2 might reasonably be expected to be different as well. However, correlations between individuals' baseline raw scores from each day of testing for Var(x) and Var(y) show that this was not the case. Table 6.11 presents these data.

TABLE 6.11

Correlations between Baseline Raw Scores in Aud 1, Aud 2 and Ex-Comp for Var(x) and Var(y)

	Var(x)				Var(y)		
	Aud 1	Aud 2	Ex-Comp		Aud 1	Aud 2	Ex-Comp
Aud 1				Aud 1			
Aud 2	.77			Aud 2	.71		
Ex-Comp	.60	.80		Ex-Comp	.57	.71	
Raw Score Mean	105.5	104.0	99.0	Raw Score Mean	167.2	121.0	122.3

5. Of course Ex-Comp was a different situation and there is no reason to suppose that individuals would react to competition in the same way as they had to the presence of only the video recorder and the experimenter.

Although the Baseline conditions were different on these different days, in terms of the subjects' degree of prior experience in the situation, there is no evidence from Table 6.11 that this had much effect on either the relative standing of subjects (correlational data) or (except for Var(y) in Aud 1) the absolute level of performance (group mean). It would appear that subjects' relative standing was similar between the three Baseline conditions, that they adopted particular performance strategies at some point between the Baseline and the end of the stress condition on each day and were either unwilling or unable to alter this strategy for the remainder of that particular session.

A final attempt to clarify the nature of performance changes between individuals is an examination of subjects' information questionnaire data, completed on Day 1 of the experiment. On the basis of this information subjects are classified according to how often they play darts. It was thought that there might be some differences between players who are highly skilled and accustomed to playing in public situations and those who have less such experience. Unfortunately, the groups cannot be divided evenly and, as there are only 4, 8 and 8 subjects in each skill group, the data do not lend themselves to statistical analysis. (Appendix VI contains the data for the three skill groups and the number of subjects in each who showed either disruption, facilitation or no change in performance for all conditions of testing.)

To summarize these data, there is some tendency for highly skilled players to improve in mild stress (Video) on M but show disruption under greater stress (Video/Competition) for both M and Var(y). The least skilled group also show some tendency to become less accurate under Video/Competition for M and Var(x). It might be expected that highly skilled players would be most affected by the experimental conditions since they would have the most vested interest in good task performance. Similarly, relatively unskilled players have less control over their performance and might show disruption more easily even though experiencing little stress. The analysis suggests that planned examinations of performance according to level of skill might be fruitful in terms of understanding individual responses to the stress conditions, but the number of subjects in each skill group precludes any but tentative conclusions about the present data.

6.2.5 Summary of Results and Discussion

The results from Experiment VII can be summarized as follows:

1. There was some tendency for subjects to show increased Var(y) when in direct competition with another (and being video recorded) than when performing only in the presence of the video recorder.

2. There was some evidence that after subjects had become efficient with using the scoring sheets, speed of performance was greater when being video recorded than it was before the introduction of the audience stress.

3. The direction of change for individual subjects based on ratios for Horizontal and Vertical Dispersal was fairly consistent from the Video to the Post-Video conditions within the same experimental session. However, there was no consistency between days in the ways in which individuals responded to the experimental manipulations.

It should be clear from the foregoing analysis of results that assessing the effects of social situations on a realistic performance task such as darts playing requires rather a different approach to that usually seen in experiments concerned with variations of mere presence. In the first place, the task itself is considerably more complex than those employed in most laboratory research, and this means that fairly detailed analysis of task components must be undertaken if the exact nature of a performance change is to be identified. In addition, the effects of individual differences appear to be so variable and complex that they almost defy systematic analysis. The level of skill which subjects possess when entering the situation would seem to be a variable meriting further investigation, but there was considerable variation even within these subcategories.

The inconsistency between the changes in individual performance on one day and the changes in almost the same situation on another day raises some questions about the validity of research which tests subjects in the experimental situation on one occasion only. If in fact subjects do not display relatively the same reaction using the same task and in very similar social conditions, then the results obtained from most research into the effects of social situations on performance would apply only to those situations in which subjects have a single exposure to a novel task under some given social treatment. Such results may be valid enough if generalization of the findings is confined only to a very restricted range of performance situations.

In light of the failure to find many unequivocal findings in this experiment, despite the detailed method of data analysis, some discussion is warranted concerning the success of the manipulations in creating either audience or competitive stress. There was some evidence that subjects speeded up either their throwing and/or scoring during the Video 2 condition, and this was taken as an indication of either heightened arousal or avoidance behaviour. Either of these eventualities would indicate that the video recording was in fact an effective audience manipulation. However, in post-experimental interviews individual reactions varied greatly. Some subjects reported rather strong dislike for the presence of the camera, stating that they tried to avoid being in its range as much as possible. Others indicated that they were not disturbed at all about being recorded. Some were dubious that the tapes were actually going to be looked at by a group of students, as the instructions had claimed, and one subject reported that he never believed he was actually being recorded (even though all subjects were recorded). In all, the video manipulation can only be judged to have been moderately successful. A live audience might have produced more striking and consistent effects. However, the experimenter was present during the Video condition, and, as in previous experiments, subjects indicated that her presence went virtually unnoticed.

Generally, the Video/Competition was felt to be a more successful manipulation. Subjects could be seen to react in ways characteristic of more natural situations. They often expressed displeasure or embarrassment over a bad shot either by groaning, giggling or making self-denegrating remarks. Some of the subjects complained about the scoring method employed, pointing out that it did not take account of near and far misses and therefore did not discriminate adequately between performers. Of course, these data did not feature in the analysis of performance and therefore the scoring method was not very sophisticated. However, that subjects should complain about it would seem to indicate that they were interested in the competition and were concerned that they were not getting full credit in cases where a dart came very near a region of higher value.

Although the present study was exploratory, it is felt that darts playing as a task is not very satisfactory when used in the type of design usually employed in social stress research. But, this is probably more due to the research design than to the task or the method of performance analysis. Firstly, because darts is a 'real' task rather than one invented for the purposes of a particular experiment, individual differences can be expected to constitute a more important part of the variance. Individuals arrive for the experiment with differential degrees of experience and skill and have a variety of previous associations with the task in real performance situations. Whether the subjects has had previous experiences of failure, success or anxiety may lead to rather idiosyncratic interpretations of and responses to the experimental situation. This supposition received some support from the analysis of personality data, which will be summarized in Chapter VII.

Secondly, because the task is complex and the consistency of individuals between days appears to be minimal, it might be necessary to study the effects of social situations in a longitudinal design which permitted repeated testing of subjects over many sessions, both in order to allow for complete habituation to the experimental situation and also to discover stable response patterns. This again can hardly be cited as a flaw in the task, but points out the perhaps erroneous notion that stable patterns in the performance of a complex skill under social stress can be effectively studied in one or two relatively brief experimental sessions. It is possible to envisage an experimental design in which novice darts players are studied over a lengthy period, documenting the way in which skill develops and is affected by various social situations at various stages of practice. Team events could even be organized and, judging from the verbal responses of subjects in the present experiment, these could take on a fairly realistic countenance. Such a design would almost necessarily require on-line analysis of data and would demand a commitment from subjects for a period of several months. Still, the task is an engaging one and this latter difficulty may not be insurmountable.

As a method of investigating skill acquisition and performance under stress the present task and method of component performance analysis raises some interesting possibilities. For example, in the present study there was some evidence that the performance component most sensitive to changes in the social environment was Var(y). Subjects showed greater variability in the vertical plain when faced with an opponent rather than when being video-recorded only. Several subjects (all males) commented that they felt less in control and experienced a feeling of loss of physical power in the competitive situation. If subjects were in fact throwing with less force or, rather, variable degrees of force, this would be reflected in greater vertical dispersal, as the line of trajectory of the darts would be directly related to the force of the throw. This difference in variability was not reflected in M, but would perhaps be important in a real dart game where players often aim for the 20's section, and where slight differences in the accuracy of the horizontal line can mean the difference between a score of 20 or of 60.

In addition the task has considerable potential as an instrument for studying the specific nature of performance disruption under conditions of environmental stress such as variations in lighting, temperature or noise level or for investigating the effects of physiological stresses such as alcohol consumption or sleep deprivation. The former would seem particularly relevant for this particular task considering the usual association of alcohol consumption with darts playing in natural performance situations.

6.3 OVERVIEW

Writing an overview for the two preceding experiments poses some problems since, although both involved motor tasks and more realistic performance situations, the designs, hypotheses tested and situations employed were essentially different. There are however some general comments which can be made concerning investigations in 'realistic' situations and the use of 'real' performance tasks. Also, the results from these experiments provide some interesting contrasts to those described in previous chapters in this thesis.

The first point of interest and of contrast in the present experiments was the use of highly involving tasks. It will be recalled that the strategy used in the experiments reported in Chapters IV and V was to employ routine, boring tasks which would be expected to easily reflect changes in arousal or motivation. For both the Pursuit Rotor and Darts there was evidence, direct and indirect, that subjects found the tasks highly involving and were intrinsically interested in performing well. Therefore a question central to these experiments is whether the subjects' basic interest in the tasks rendered them insensitive to the social-evaluational stresses employed. The results suggest that this was not the case. The Pursuit Rotor showed small but reliable differences on both initial and later learning phases in response to the social situations, and darts performance, rather than being unaffected by the experimental manipulations, showed marked, sometimes extreme, individualized responses which negated any overall trends. There was evidence, however, in the case of the pursuit rotor that self-competition probably reduced the impact of the social situations, the latter merely accentuating what was already highly motivated task performance. Even so, it appears that even when the task is highly involving performance can be affected by mild social stress of the type employed in these experiments. Whether the mediating variable is arousal, social comparison or self-consciousness (objective self-awareness) is another issue and will be discussed in Chapter VII.

In contrast to previous experiments, the two studies reported here were not concerned in any central way with pacing tendencies. In darts performance this was not really an issue since subjects were matched for level of skill as nearly as possible. The pursuit rotor reduced the likelihood of subjects pacing because speed on the task was mechanically controlled and there was little opportunity for adjusting overall accuracy via alterations in performance strategy. Also, in this latter case, the goal of performance was rather more clearly specified than in previous experiments, and, although others could provide information about 'normative' performance, subjects could themselves evaluate their performance in terms of its absolute nearness to the specified goal. Even so, there was some evidence of pacing tendencies on the pursuit rotor, although this was apparent only in the most overtly stressful situation in which performance was made competitive and public. That pacing should occur under these conditions underlines the point made in previous chapters that emphasizing evaluation/competition has a deleterous effect on individuals of initially high performance level. The specific process responsible for this and the implications for performance in real-life situations will be discussed in Chapter VII.

Experiment VI (Pursuit Rotor) was essentially different from all other experiments conducted in this series because it was the only one in which arousal theory per se was tested using a task with a dominant/subordinate response hierarchy. With some reservations, it was concluded that performance on this task was consistent with predictions based on arousal theory. It has been argued previously that many tasks do not lend themselves to classification in terms of response hierarchies. Certainly it has not been the case in other experiments cited in this thesis that the presence of others has unequivocally led to better performance on simple tasks or worse performance on difficult ones. Interestingly, the situations employed in Pursuit Rotor were much the same as those employed in Experiments I to V. The fact that the predicted patterns emerged at different stages of practice suggests that the failure of previous experiments to show overall facilitation or impairment was not the product of some peculiarity in the situational manipulations but due to the tasks employed. This also suggests that the situations used have been largely comparable to those employed in the literature and, therefore, that the interpretations placed on the findings are in fact relevant to the larger body of experimental work which exists concerning the effects of social situations on performance.

The present two experiments also both point out the difficulties inherent in employing realistic performance tasks to measure the effects of social stress. Apart from the complicating effects of subjects' level of involvement with the task and individuals' previous associations with other performance situations, there are factors inherent in the learning and performance of skilled activity which in themselves may affect performance outcomes either directly or in interaction with situational variables. The possible complications of fatigue and/or performance blocking outlined in Section 6.1.5 for the pursuit rotor and the sensitivity of darts performance to individual variability between testing days are cases in point. It will be recalled that previous experiments have used sensitive tasks, but only ones which were routine and which could be mastered with a minimum of practice. To study the effects of social situations on skilled performance may require a different type of experimental design in which subjects can be tested on a number of occasions so that stable individual patterns of performance in the social situations of interest can be identified or at least thoroughly investigated. Furthermore, phenomena such as performance blocking and reminiscence are well documented features of performance on perceptual motor tasks. However, the ways in which these basic performance phenomena interact with social stress may only become apparent when studied longitudinally. Short samples of performance on these more complex tasks may yield results which are unrepresentative, transitory and which fail to detect important and interesting changes that take place over time. This of course is a criticism often levelled at 'one-off' experimental procedures but is perhaps even more justified in the case of motor skills where changes in performance over time have received much attention.

CHAPTER VII

CONCLUSIONS AND IMPLICATIONS FROM EXPERIMENTAL FINDINGS

7.0 INTRODUCTION

In the series of experiments described in the previous three chapters a variety of social stimuli were investigated, all of which were concerned with the effects of various audience and coaction situations on performance. The results have been complex, and, because of the variety of tasks and situations employed, an overview of the findings is presented here to help highlight and clarify general trends. In some instances previously discussed points are reiterated, but within the context of the whole experimental series. In other cases, new material, such as that dealing with personality differences, is presented. Sections 7.1 and 7.2 are intended to summarize the results concerning the role of social and individual factors in determining performance outcomes. Section 7.3 is an overview of general issues raised by the experimental results, an evaluation of different theoretical viewpoints concerning these issues and suggestions for further research.

7.1 THE EFFECTS OF SOCIAL SITUATIONS

7.1.1 An Overview of 'Mere Presence' Effects

An overview of the experimental findings indicates that the mere presence of others can have an effect on the performance of individuals. In all cases (except darts throwing) this effect was facilitative, either in terms of improved output or accuracy or the emission of dominant responses, the specific nature of facilitation depending on the particular task.

However, although some significant differences were found between alone and social conditions, this was not always the case, and in many instances the actual mean differences were rather small. It has been argued that a thorough analysis of experimental data should include not only tests of significance but estimates of the magnitude of the effect i.e. the proportion of variance accounted for by the comparison (Hays, 1963; Keppel, 1974). Without this latter information statistically reliable but trivial effects can be attributed an importance which is misleading. Hays (1963) suggests that only treatments producing results which are both significant and which account for a substantial proportion of the variance warrant further research. However, he leaves the question of what magnitude should be considered 'meaningful' to the individual experimenter. In most areas of psychological research 10 percent is arbitrarily designated as a substantial effect size.

By way of summary, Table 7.1 presents all significant main effects ($p<.05$) from the preceding experiments with the corresponding estimates of the variance accounted for (w^2).[1]

The table reveals that in only three cases does the magnitude of the effect exceed 10 percent. In one case (Expt I) the results were confounded by the counterbalanced design, and on the initial testing day the effects were restricted to only the first 5 minutes of performance. In Expt. V, subjects who were tested in groups and not visually screened improved their error scores over time. However, accuracy on the task (Transformation) was generally very high: this improvement did not alter either Output or the number of Correct Solutions and was probably not a very meaningful performance measure. Experiment II would appear to give the strongest evidence for a facilitative effect both for Output and Hits.

The effects of a passive audience (Expt. II), although marginally

---

1. The calculation of the size of effect is based on the general formula (eg. Hays, 1963): est. w^2 = (SS Treatment - MS Error)/(Total SS + MS Error).

TABLE 7.1

Summary of Main Effects and Estimates of Size of Effects

Expt. No.	Task & Measure	Nature of Effect	F & t values	P	Size of effect w^2 as proportion of total variance accounted for
I (N = 16)	cancellation (Output)	facilitation with group	F = 10.98 (df = 1,112)	.01	15.2
I (N = 16)	cancellation (Hits)	facilitation with group	F = 7.12 (df = 1,112)	.02	9.7
II (N = 100)	cancellation (Output)	facilitation with group	F = 10.24 (df = 1,90)	.001	8.4
II (N = 100)	cancellation (Hits)	facilitation with group	F = 12.59 (df = 1,90)	.001	10.3
II (N = 100)	cancellation (Output)	facilitation with audience	F = 3.98 (df = 1,90)	.05	2.7
II (N = 100)	cancellation (Output)	facilitation with group (vs audience)	F = 3.93 (df = 1,90)	.05	2.7
V (N = 60)	transformation (Errors)	facilitation in Im-Comp(F) (vs E and Im-Comp)	F = 5.10 (df = 2,54)	.025	11.8
VI (N = 96)	TOT (Phase I)	impairment with group	F = 6.24 (df = 1,80)	.02	4.9
VI (N = 96)	TOT (Phase II)	facilitation with group	F = 3.83 (df = 1,80)	.05	3.0
VII (N = 20)	Darts (Var(y))	variability greater with competition (vs video)	t = 2.12 (df = 19)	.05	8.0

significant account for only a small proportion of the variance. This is also true for group conditions and pursuit rotor performance (Expt. VI), even though the social treatments did in this case produce significant F ratios. Furthermore, although several results are listed in the table, these are relatively few compared to the number of situations and measures examined in these 7 experiments: there is some redundancy as well since in all cases Output and Hits are highly correlated ($r = .82$ to $.94$). Experiments III and IV do not appear in the table since there were no significant main effects for social conditions in these two experiments.

In conclusion, the mere presence of others would seem to be associated with occasional and relatively minor effects on performance.

7.1.2 <u>The Effects of Potential Evaluation</u>

All of the experiments reported in Chapters IV, V and VI employed some variety of evaluative manipulation; either specific ego-threatening instructions or public feedback of results with observation by other subjects. Having concluded that mere presence, defined in the widest sense, is a relatively minor factor in determining performance outcomes, the next obvious question is whether potential evaluation interacts with the social manipulations to produce more consistent and stronger effects.

The findings concerning this question are somewhat ambiguous. There is little evidence that evaluative manipulations interact with the social environment in the way suggested by Cottrell (1972). Experiment I (see Table 7.1) did yield evidence of performance facilitation in the presence of others. In this case ego-threatening instructions were given in the group situations, although in retrospect it was felt that the evaluation potential was minimal. An interaction between social conditions and instructions was found in Experiment II for both Output and Omissions, although this was largely due to an apparent change in performance strategy for subjects tested alone. Also, in this experiment a main effect for Instructions was obtained for RDS, but again the effect was strongest for subjects in non-social conditions. The only other instance in which evaluation appeared to be related to performance was in darts playing, where subjects tested in direct competition showed more variable performance than in less stressful audience conditions. However, as Table 7.1 indicates the size of this effect was small.

Although there is little evidence from these studies which suggests that evaluative stimuli interact with social environments in a manner consistent with drive theory, there is evidence that evaluation potential might be associated with changes in individual performance strategies. The direction of the change is determined by the individual's initial level of performance, and the effect appears to be related more to social than to task variables or performance processes. This conclusion is based on the correlational data on pacing. In the five experiments in which some investigation of pacing was possible there were indications that the social manipulations had the effect of bringing individual performances closer together. In four of these five experiments this tendency was most marked in the situations with the most explicit evaluative cues. Table 7.2 summarizes these results. The conditions are ordered, in terms of their presumed evaluation potential, from A and E to Ex-Comp(F). Passive Audience conditions are judged to be somewhere between these two extremes. A and E are presumed to be roughly equivalent as are Ex-Comp and Im-Comp(F), the latter involving one explicitly evaluative manipulation. For the present, these are considered to be of roughly equal salience and are placed next to each other along the continuum, public feedback being arbitrarily assigned the lower position.

TABLE 7.2

Summary of Pacing Results

Expt. No.	Measure	Conditions and Pacing Values								Statistic
II (N=100)	Cancellation (Output)	A 38360	A-T 37568	Aud 24030	Im-Comp 33476	Ex-Comp 15045				Var
III (N=24)	Cancellation (Output)	Ex-Comp -.36	Ex-Comp(F) -.74							r
III (N=24)	RDS	Ex-Comp -.37	Ex-Comp(F) -.69							r
IV (N=90)	Cancellation (Output)	E .10	Im-Comp -.57	Ex-Comp -.66						r
V (N=60)	Transformation (Output)	E -.42	Im-Comp -.84	Im-Comp(F) -.61						r
V (N=60)	Syllogisms (Output)	E .01	Im-Comp -.11	Im-Comp(F) -.45						r
VI (N=96)	TOT (Phase I)	E -.26	E(F) -.26	E-T -.35	E-T(F) -.11	Im-Comp -.51	Im-Comp(F) -.32	Ex-Comp -.18	Ex-Comp(F) -.63	r

Examination of the table reveals a fairly consistent trend in the magnitude of correlations, values being generally lowest in the Alone or Experimenter only conditions and highest under Explicit Competition. The pattern is not as clear in Experiment II where difference scores were not available, although the smaller variance for Ex-Comp suggests some suppression in the range of scores.

Furthermore, it does not appear that the particular form of the evaluative manipulation is especially important in producing this effect: evaluative instructions are associated with a pattern similar to that of public feedback, although combinations of evaluative cues seem to be the most powerful situation for eliciting pacing tendencies. Theories of Objective Self-Awareness point out the variety of stimuli which can lead individuals to focus on their performance level (mirrors, instructions, video recording). These results suggest that directly evaluative instructions did not differ in terms of either main effects or the magnitude of pacing correlations from other manipulations which were aimed at highlighting the subjects' performance level. A wide range of stimuli and situations could presumably fall into this category, many of which would not superficially be expected to have an effect analogous to evaluation stress. Whether or not evaluation apprehension is the common factor between these various manipulations will be considered in Section 7.3.1.

A final consideration, which was raised initially in Chapters I and II, is the role of implicit versus explicit evaluation. As pointed out above, with the situations and tasks used in these experiments, there appears to be little difference between manipulations which have some potential for evaluation (implicit) and those in which evaluation is explicitly stressed: neither appear to have much impact on mean performance differences. However, the data indicate a tendency for explicitly evaluative manipulations to lead to more pronounced pacing. Even so, the magnitude of the correlations were reasonably high in implicit competition situations and were always higher than those obtained for subjects tested alone. In fact, in most cases the differences between alone subjects and those tested under implicit competition were greater than those between implicit and explicit evaluation conditions. Considering this progression in the magnitude of the correlations it seems reasonable to conclude that influences were present in neutral groups which were similar but less pronounced than those in explicitly evaluative situations.

Although it was pointed out previously that most group situations would elicit competitive or comparative tendencies to some degree there are few examples in the literature which give estimates of the relative strength of these implicit and explicit factors. Most research does not permit such comparisons either because the experimental designs have employed only two conditions or because the methods of data analysis have been inappropriate for detecting patterns of change within different group conditions. The present data indicate that explicit evaluation is likely to lead to more pronounced pacing between individual group members, but implicit factors have nearly and sometimes equally powerful effects.

In some respects the role of explicitly evaluative cues is more straightforward in the case of subjects tested alone. For example, in Experiment II significant differences were found between alone subjects who were given either neutral or ego-threatening instructions. In this case Output, Omissions and RDS were affected by the instructions, and the results suggested that individuals had changed their performance strategy in attempting to achieve more efficient overall performance. The effects of these same manipulations in group situations is more complex; behaviour is more related to the individual as a member of a group rather than as a performer per se.

7.1.3 Coaction versus Audience Effects

There is little evidence from this series of experiments to suggest that the presence of a passive audience has an appreciable effect on individual performance. Neither the experimenter's presence coupled with another observer (Expt.II), video-recording (Expt.VII) nor manipulations of public and private performance (Experiments III and VI) resulted in any clear mean differences on any of the tasks or measures investigated. However, effects were obtained when the audience was comprised of other subjects. In one case this condition led to the facilitation of dominant responses (pursuit rotor) and in another to increased variability (darts). In two cases, when the audience manipulation was coupled with variations in the degree of public feedback, increased pacing in the public condition was evident (Experiments III and VI).

It is possible that an audience of subjects is more suggestive of evaluation than are uninvolved observers, since the performer might infer that the other subjects would be interested in comparing his task competence with their own. It also appears that an audience of subjects can be either a source of stress (increasing variability) or information (pacing). In the latter case, the subject spectators have an effect analogous to coactors. Whether the spectators or coactors are viewed as chiefly evaluative or informative is not discernible. However, one means of coping with the evaluation stress generated by the presence of other subjects may be to perform in a way which is indistinguishable from the others: the situation then becomes nonevaluative. Such a coping strategy may be an example of avoidance behaviour (reducing evaluation anxiety by making evaluation impossible) or affiliation (leading to increased group cohesiveness). Neither of these hypotheses can be explored on the basis of the present data, although this is perhaps a question meriting further research.

7.2 THE ROLE OF INDIVIDUAL DIFFERENCES

7.2.1 Initial Level of Performance

Perhaps the most interesting finding of the present work was the negative relationship between initial level of performance and change in performance in the social situations. Although this effect was not initially anticipated, it is not entirely novel. Allport (1924) found that subjects' judgments of odours and weights were less extreme when made in the presence of others and called this a 'levelling' effect, although attempts to replicate his results have not been very successful (Granberg et al., 1975). He also noted a negative relationship between the magnitude of social facilitation and initial ability. There are a few other isolated studies which have noted a similar pattern on performance measures. Paulus and Cornelius (1974) found that very skilled gymnasts demonstrated more disruption of performance when spectators were present (particularly if advanced warning was given) than did less skilled performers. Also, on a tracking task performance has shown decreased intersubject variability when individuals were given knowledge of other performers' scores (Foot and Lee, 1970). In a more complicated design, Clifford (1971) varied group composition, in terms of ability (heterogeneous versus homogeneous) and evaluation (reward versus no reward). Interestingly, in the heterogeneous groups subjects of high ability performed relatively worse under reward conditions than either subjects of low ability (who improved) or high ability subjects in non-reward conditions. Consistent with the correlational data presented in Table 7.2, Church (1968) obtained a value of -.66 between initial ability on reaction time and facilitation of performance under competitive conditions (subjects were given knowledge about which member of a dyad was faster).

Although interactions between initial performance level and social conditions have been a recurring observation, this has generated little systematic investigation or discussion of the mechanisms responsible. Paulus and Cornelius (1974) adopt an explanation based on Broen and Storm's (1967) discussion of the inverted U-shaped curve, i.e. that in the case of highly skilled performers increased arousal results in facilitation of the subordinate response since the dominant response has already reached an effective ceiling. However, Foot and Lee (1970) favour an interpretation based on social comparison processes. The present data offer some support for both points of view.

Regarding the former, care has been taken in the present discussion when referring to the relative differences in initial performance as differences in initial performance level rather than ability. Since no means were available to estimate the level of arousal or motivation in pretest sessions, it cannot be assumed that those subjects who showed relatively high initial performance were not more highly motivated to begin with. Kohfeld and Weitzel (1969) found that individuals who were achievement oriented, conscientious or concerned with making a good impression showed equally high levels of performance in both social (peers present) and nonsocial (experimenter only) conditions. In the present studies it should probably be assumed that certain individuals were more highly motivated than others during the pretest. The additional stressors in subsequent sessions may have reduced this initial discrepancy in performance motivation between subjects, leading to more uniform performance levels, or may even have produced super-optimal arousal in already highly motivated subjects and subsequent poorer performance.

However, although this interpretation is consistent with the mainstream of theorizing about others' presence and performance, it is too simplistic to give a complete explanation of the complex pacing patterns obtained. Firstly, although some positive relationship between motivation and level of performance would be expected, this would by no means be perfect. Individuals of high motivation may perform at their own maximum degree of efficiency but may still be relatively inept compared to other individuals. Also, an arousal interpretation cannot explain the differences between the magnitudes of pacing correlations in public and private performance conditions, unless it is further assumed that the former directly led to higher arousal.

An explanation based on social comparison processes (Festinger, 1954) provides a more convincing rationale. Firstly, in general, pacing was more pronounced in situations in which public feedback was given. In Experiment III for example, a negative relationship was obtained between change in Output for cancellation and change in RDS when RDS answers were given orally, whereas a positive correlation was found between these two measures in the Written condition; in this case there is direct evidence that subjects changed their performance strategy such that those who were initially good at RDS directed more attention to cancellation and vice versa. These complex patterns of change suggest that some cognitive factor, independent of arousal level, mediated between situational variables and performance. Similarly, Clifford's (1971) results showed that in homogeneous ability groups subjects of high ability did not show a decrement in performance, as was the case in heterogeneous groups, and those of moderately high skill showed facilitation. Presumably if arousal was the mediating factor its effects would be the same regardless of the relative skill of other group members.

Several authors writing about social facilitation effects have pointed out that the behaviour of others in the group can have a marked effect on the individual (Wheeler, 1968), and that although companions may be arousing and may lead to increased activation, they are also models and can guide the subject's responses (Berger, 1968). Supporting this notion, two experiments have shown that the presence of a 'worried' or 'calm' companion influenced subjects' performance on maze learning and altered the size of the predicted social facilitation effect (Shaver and Liebling, 1976, and Liebling and Shaver, 1973). Some 'levelling' in terms of arousal may of course have taken place in these situations and the ones investigated in the previous chapters. However, a social comparison interpretation seems more congruous with the present findings and does not involve untestable assumptions about the relationship between arousal level and initial performance or the relative arousing properties of the situations employed.

Considering the variety of tasks on which notable pacing effects were found, and the likelihood that these might have modified or negated mean performance differences between conditions, the initial level of subjects' performance would appear to be a crucial variable and one which should receive more systematic attention in future research.

7.2.2 Sex Differences

An overview of results from the preceding chapters indicates that the social situations investigated had little differential effect on male and female subjects. In only one case was a significant difference obtained involving the Sex factor: this was in Experiment IV in which female subjects showed significantly more improvement in performance than males when tested with only the experimenter present. This result was explained in terms of differential pacing patterns for males and females in group conditions. Females of initially high performance level tended to pace downward, lowering the mean for the whole condition, while males of low ability paced upward, raising the overall group mean. A similar pattern was found in Experiment V, although the results were less conclusive and were not reflected in mean performance differences.

It is somewhat surprising that subjects' sex did not interact more consistently with the experimental conditions, especially on tasks for which females tended to demonstrate less relative skill than males (pursuit rotor and darts). There is no evidence from any of these studies to suggest that mature women (university and Open University students) are more negatively affected by social-evaluative situations than are males. However, the literature is not unequivocal regarding the effects of competitive manipulations and subjects' sex (see Chapter II). Furthermore, attempts to demonstrate interactions between sex of subjects and audience sex have been largely unsuccessful, at least with mature subjects (Bird, 1975; Riki, 1974; Carment and Latchford, 1970; Singer and Llewellyn, 1973).

However, it is likely that sex differences are more striking with younger subjects (both Experiments IV and V employed high school pupils). Pacing has been noted for female elementary school children on marble sorting (McManis, 1966), and Morgan and Mausner (1973) found results similar to those obtained in Experiment IV: high school girls of high ability on a Hidden Figures Test depressed their scores when working in dyads with a low scoring partner, while low scoring males improved with a high scoring partner, low scoring females were unaffected by the experimental manipulations. In an extensive review of the literature on sex differences Maccoby and Jacklin (1975) conclude that boys are more favourably stimulated by competitive situations, and girls, although willing to work hard for grades, frequently avoid the implication of being better than boys, and this is particularly marked during adolescence.

It is possible that the conflict felt in competitive situations is highlighted for adolescent girls, who may still be experiencing some uncertainty about their sex role identities. Also, it should be pointed out that most of the literature which suggests that females are less achievement oriented and anticipate negative consequences following success are based on projective test measures. House (1974) found no relationship between fear of success imagery and actual performance, and in fact, it is questionable whether a direct relationship exists between projective test data and achievement behaviour in general (Lazarus et al, 1957). In a university population, females who do feel anxiety in competitive settings more than likely have developed means of coping with it. There may also be some cross-cultural differences in responding styles, since the bulk of studies concerned with fear of success have been conducted on American university students, who are generally less selected in terms of academic performance than are British students.

7.2.3 Personality Differences

In general, the relationships obtained between personality and performance measures were low and inconsistent between experiments. For this reason the data were not presented in the main text, although complete tables of all correlations can be found in Appendix I, along with a summary of the results. It was not expected that the magnitude of correlations would be consistently high. However, it was reasoned that over a whole series of experiments consistent trends might emerge in the magnitude and direction of correlations, and that these might suggest lines which future research might take up.

Several different tests were used, partly because some measures proved to be more suitable than others and also because it was not clear what individual factors might be most sensitive to the situations employed. For example, the Argyle-Robinson Achievement Motivation test (Experiments I and II) provided scales for both Achievement Motivation and Fear of Failure. However, with a sample of university students the range of scores was too narrow to permit a meaningful analysis of the data, particularly for the Achievement Motivation scale. Fear of Failure seemed a more reasonable dimension of individual differences and subsequent experiments employed the Test Anxiety Scale in place of the Argyle-Robinson.

Other measures of personality were the Eysenck Personality Inventory and the Audience Sensitivity Inventory. Surprisingly, in experiments in which both general trait (EPI) and specific anxiety measures (TAS and ASI) were used the inter-correlations were high. Test Anxiety correlated positively with Neuroticism (r = .60 to .81) and Audience Sensitivity was negatively related to Extraversion (r = -.38 to -.84). Although recent literature tends to favour the use of specific measures in situations which are oriented toward particular types of stress (audience, evaluation, etc.), these high correlations between the general and specific measures suggest that there is considerable overlap in the characteristics that these separate instruments are measuring. However Extraversion and Neuroticism appear to be independent (r = -.39 to .22), as do Test Anxiety and Audience Sensitivity (r = .12 to .40).

The most interesting results arising from the personality data were those from Experiment VII (darts). Negative correlations were obtained between Fear of Failure and the transformed ratios for all measures in both Video 1(r = -.43 to -.76) and Post-Video 1(r = -.02 to -.73). This indicates that FF is associated with improvement in performance under stress. However, this pattern was reversed on Day 2 (Aud 2) under the same testing conditions and FF was associated, although to a lesser degree, with performance decrement (r = .05 to .42 for Video 2; r = .22 to .32 for Post-Video 2). Also, there was some suggestion of a positive relationship between performance disruption and both FF (r = .10 to .40) and Audience Sensitivity (r = .26 to .51) on Day 3 when subjects were tested in direct competition. These findings are consistent with other reports about the effects of giving advanced warning of impending stress on anxiety responses (Paivio, 1963). In the case of darts playing FF was associated with better performance when the stress was sudden. However, on Days 2 and 3 subjects had some forewarning of the impending stress situation. This advanced warning may have allowed time for predisposed individuals to become anxious or worried and may be accountable for the different pattern in correlations obtained on these separate testing days.

There are no other data from the analysis of personality measures which are clear enough to warrant discussion. However, several qualifications need to be placed on the interpretation of the correlational data. Firstly, data from three of the experiments are based on sample sizes of less than 15 per cell. In these cases a full range of scores was not always obtained and may have led to somewhat spurious results. Secondly, several different performance measures as well as raw scores and difference scores for the same tasks were examined, and the effects of the personality factors on these different measures should perhaps not be expected to be the same. However, for purposes of analysis, tasks were categorized in terms of components of performance rather than overall measures and this precaution should have reduced the importance of some of the task specific factors.

Even bearing in mind the above qualifications, there are some comments which can be made about the results. Firstly, irrespective of the pattern of correlations between experiments, in few cases were individual r values of significant magnitude, nor were the patterns of correlations within experiments suggestive of any systematic effect due to the personality variables. Secondly, although the tasks used were in some cases rather dissimilar, the social situations were much the same. If the personality factors investigated in these studies do have any effect on the performance of individuals it appears they are relatively minor compared to task and situational variables.

However, before reaching a final conclusion a further precaution was taken in evaluating the results from the correlational data. The statistic 'r' assumes a monotonic relationship between the two variables being compared. If a curvilinear (or more specifically an inverted U-shaped relationship) exists, as has been suggested particularly in the literature on anxiety and performance, then the values obtained for 'r' would necessarily be low. To check for this possibility all performance and personality measures for all experiments involving 20 or more subjects per cell were computer analyzed and displayed in the form of scatterplots.[2] In none of these cases was there any suggestion of a curvilinear relationship between any of the variables correlated. On the basis of this final analysis it is concluded that the personality factors of Extraversion, Neuroticism, Fear of Failure and Audience Sensitivity have relatively little importance in accounting for the variability in performance in social situations on most of the tasks employed.

There are several reasons why these measures of individual differences may have proved ineffectual. Firstly, the personality tests used are only gross measures of individual characteristics and there are several other problems associated with instruments of this nature (see Chapter III). Secondly, Test Anxiety and Audience Sensitivity have been employed almost exclusively with American and Canadian university populations. As mentioned previously, British students are a more selective group and

---

2. This amounted to 167 scatterplots, which are not presented in Appendix I because of limitations of space, although the data are available for inspection. Experiments with less than 20 subjects per cell were omitted from this analysis since meaningful interpretations require a large number of data points.

there may be some differences in the ways in which these measures interact with performance with this sample.

Apart from these considerations about the personality instruments, the tendency for subjects to pace in group conditions probably confounded the effects of the personality factors. With correlations between initial level of performance and change in performance ranging from -.45 to -.84, the relative influence of other individual factors would be minor. It is noteworthy that the highest and most consistent r values were obtained in Experiment VII (Darts) in which pacing was minimized by matching subject pairs on level of ability. Whether the higher relationships were apparent because pacing was controlled or were based on the task being more life-like and perhaps more stress provoking is unknown. However, it would seem that future investigations involving personality factors should consider both these procedural features in their designs.

7.3 ISSUES AND SUGGESTIONS FOR FURTHER RESEARCH

7.3.1 Anticipated Evaluation versus Objective Self-awareness

Evaluation potential has been given considerable emphasis in the literature as a necessary condition for eliciting mere presence effects, at least when performance level has been the behaviour of interest. However, the results obtained in the evaluative situations investigated in the present work, pacing effects in particular, were uncharacteristic responses to evaluation stress. In many cases the initially better performers actually showed a decrement in performance when evaluative stimuli were present. This is inconsistent with some of the literature on anxiety which has shown that high ability test anxious subjects are not debilitated in real examination situations whereas highly anxious subjects of low ability are (Birney et al, 1969; Dubey, 1976; Spielberger, 1975).

Based on the findings from these experiments there are several reasons for concluding that the evaluative manipulations had the effect of creating self-awareness rather than evaluation stress. Firstly, in Experiment II subjects in the A-T condition showed a pattern of performance which was different from that of any other condition. These subjects made more errors of omission on the cancellation task (defined as the more important task measure) and showed better performance on RDS. It was suggested that in the absence of clear performance goals subjects shifted attention to that aspect of the task which provided the most objective performance feedback. Secondly, the results from Experiment III, as discussed above (Section 7.2.1), indicated that in the more evaluative situation attention was directed more toward evaluating level of performance than achieving high performance: this was evident from the changes in performance strategy in cases where individual performance was discrepant from the group average. Both of these studies suggest that the evaluative manipulations led subjects to focus on their level of performance (a product of increased self-awareness) but not necessarily to try to improve it, which would be the expected response to evaluation anxiety.

Furthermore, a variety of stimuli were used to create evaluation stress, and these seem to have had basically the same effect on performance patterns. It may be that all were suggestive of evaluation. Still it is not clear why evaluative instructions (which stress individual achievement) should have elicited pacing in the same way as did direct public feedback (which highlights inter-subject comparison). Finally, there was little evidence either from subjects' self-reports following the experiments or the experimenter's observations during the performance periods to suggest that subjects were aware of any feelings of real evaluation stress.[3]

Whether evaluative stimuli elicit evaluation stress or self-consciousness may seem an academic distinction. However, quite different predictions concerning performance outcomes can be made depending on these different orientations. Increased self-awareness does not imply increased arousal, therefore the effects of the former on tasks with distinct response hierarchies may be different. Also, for tasks on which several performance strategies are possible, predictions may depend on which task parameters are more accessible to self-evaluation of performance.

Of course, the possibility still exists that the situations used in this series of experiments were for some reason not as effective in creating evaluation stress as other experiments have been. It is difficult to imagine why this should be the case, but if it is, then there would seem to be even more reason for identifying just what features of experimental situations are salient in producing evaluation apprehension, how these are different from those which merely alert subjects to their level of performance and the specific ways in which these two orientations affect performance on sensitive tasks.

7.3.2 The Role of Ambiguity

That ambiguity, particularly regarding performance goals, may in part determine subjects' orientation to the task has already been mentioned. It was suggested that the tasks used in these experiments, being novel and for the most part self-paced, may have led individuals to seek information about the appropriate level of performance, particularly in those situations in which evaluation was stressed. Chapter I reviewed

---

3. The only exception to this general observation was during darts playing under competition and for two individuals during pursuit rotor performance who began showing hand tremor and loss of motor control in the public feedback conditions.

some of the literature concerned with social comparison and uncertainty and also conformity behaviour in situations in which subjects were given discrepant information. However, other areas of the social psychology literature point out that information seeking may be a common response in many situations and that the information received can modify the subjects' behaviour. This has been well demonstrated in studies on bystander intervention in which an accident is arranged so that subjects are uncertain whether or not they should intervene (Latane and Darley, 1968; Wolfson, 1977; Bickman, 1972; Milgram and Hollander, 1964; Smith et al., 1973; Clark and Word, 1974; Clark and Word, 1972). In these situations, the subjects' response was influenced by the reaction of others around him. The same effect has been found in more standard conformity studies, which show that individuals are more likely to give task responses which are consistent with the information given either by other subjects (confederates) or the experimenter when the task involves unclear performance guidelines or is extremely difficult or impossible (Walters, 1968; Nordholm 1975; Brightman and Raymond, 1975; Blake et al., 1956).

The present experiments did not manipulate task ambiguity systematically therefore it is not possible to claim with certainty that the same effects (pacing) would not have been obtained on more structured tasks with well defined performance goals. However, it was noted that on the pursuit rotor, which has a fairly clearly defined goal, pacing was restricted only to the most highly evaluative situation and only to the initial period of performance. Also, performance on Logical Syllogisms showed much less evidence of pacing, and it was suggested that information from others regarding only relative speed of performance (the only measure available for comparison) may have been irrelevant for this difficult cognitive task.

It is possible that the pacing effects noted in these experiments may simply be another example of a general human characteristic to comply with contextual cues when faced with an ambiguous situation in which appropriate behaviour is not clearly defined. Locke's work has inspired an impressive amount of evidence indicating that specific and particularly high performance goals are associated with high levels of performance for individuals both in experimental and applied settings (Locke and Bryan, 1967; 1966a; 1966b; Locke and Bryan, 1968; Latham and Baldes, 1975). It is suspected that providing subjects with clear performance standards may modify the tendency to pace, and experimental investigation of this variable may aid in better predictions about the effects of others' presence on individual performance.

7.3.3 The Role of Arousal

The present data have raised serious questions about the value of theories of arousal in predicting or explaining performance in social situations. Even when the predicted effects have been demonstrated their magnitude suggests that level of arousal (as discussed by Zajonc and Cottrell) may be of theoretical interest but little practical utility (see 7.1.1). It is perhaps unfortunate that most research has focused exclusively on arousal level in interaction with task variables. This conceptualization of performance behaviour may be too narrow: arousal level is after all a general state of the individual and would be expected to affect many aspects of his behaviour. Tolman (1968) has pointed out that the occurrence of facilitative effects depends very much on what behaviour is being measured. For example, animals who show increments in eating behaviour in the presence of social stimuli must show decrements in some other behaviour, even if it is only in "standing around" (p.47).

In social situations involving human subjects the task of correctly identifying the dominant response becomes even more difficult (Wheeler, 1968). Broadbent (1977) has described human behaviour in terms analogous to computer programmes, and has alluded to the problems of understanding performance processes in discrete situations (subroutines) when in fact the behaviour is embedded within an ongoing programme. In some respects an experimental situation can be viewed as a subroutine momentarily interpolated over the subjects' ongoing activity of studying or work, meeting social and domestic commitments and physical needs. Execution of the task is a further and quite specific subroutine embedded within the experimental situation. Conceived in this way it is not surprising that the effects of arousal level on task performance have often been inconsistent or weak.

Throughout this thesis the conditions investigated have been referred to as social-evaluative situations. The explicit assumption has been that the social context in which performance takes place <u>can</u> affect both the quality and quantity of work done. However, there is no reason to assume that even if arousal level is affected that it would necessarily be reflected directly in task responses. The individual's complete response hierarchy would include a vast range of behaviours and ones which are task specific may be subordinate to others of more relevance to the immediate situation, such as information seeking. This consideration does not deny that arousal can and does have an effect on performance, but raises the question as to whether focusing exclusively on performance measures has not limited the understanding of the effects of these basic social conditions.

Since the experimental variables under investigation are <u>social situations</u>, a more sensible way of approaching the topic may be via a more direct measurement of <u>social behaviour</u>. There are a few reports which have done this, looking at behaviours such as laughter (Chapman, 1973, 1974b; Chapman and Chapman, 1974), dominance responses (Meglino, 1976) and qualititative reactions to strangers or companions (Rajecki et al., 1976). The lattermost study (using chickens as the subject population) gave evidence of quite different response hierarchies in the animals depending on whether they were in the presence of strange birds (which elicited fear responses) or cage-mates (which elicited feeding). The present experiments suggest that social processes may be an important factor in determining performance outcomes in others' presence. Whether arousal level interacts with the occurrence of any particular social behaviour or whether stimuli present in the situation elicit responses which are independent of arousal is a question yet to be explored.

7.3.4 Task Variables

The importance of the particular task used and the manner in which performance is analysed has previously been discussed. The data presented in Chapters IV - VI do not suggest any modification of these initial concerns. However, some additional questions have been raised regarding specific features of tasks. One of these is the possible role of performance goals. Such goals may be evident to subjects either via specific information given to them or because the task itself provides fairly clear performance targets. Thus far this variable has virtually escaped notice and systematic manipulation in the literature. In most applied situations (particularly educational and industrial ones) individuals do have a reasonable idea of what standard of work they must produce, and it is not clear, except perhaps in the case of individuals whose ability diverges extremely from known norms, whether the presence of others in these situations will have much effect on either the way in which the task is approached or the quality and quantity of work done. In cases where information seeking is unnecessary, facilitative effects due to social stimuli (arousal) may be a more important and consistent source of performance variation.

The publicness of the task and the ease with which performers can evaluate their own and others' performance is another variable which has received relatively little attention. The present data strongly suggest that when performance is made public, subjects compare their level of task competence with each other and pace themselves with the group mean. Here again this may be a performance pattern unique to novel tasks, or may be counteracted by specific reward contingencies, offering monetary incentives for example (Scott and Cherrington, 1974). However, performance on some tasks (most motor tasks) is necessarily more public than on others (paper and pencil tests), and failure to give attention to this task variable, like that of performance goals, has perhaps been a source of error variance in previous literature.

A special feature of the present work has been the use of sensitive tasks and the analysis of performance in terms of task components. Not all tasks used were equally sensitive; for example the pursuit rotor only provides measures of accuracy and variability, although it demands a high degree of concentration and should be sensitive to disruption under stress. Darts throwing, on the other hand is probably too sensitive to be used to much advantage in short term studies: performance is highly variable within individuals and between different days and is probably affected by a number of unidentified and perhaps uncontrollable factors in the situation. The dual task of cancellation/RDS is judged to have been the most useful in the type of experimental design employed in this research and provides the most scope for further investigation of the variables mentioned above. For example, Hockey (1973) has shown that subjects' efficiency can be modified by instructions which suggest different strategies for RDS performance. Experiment III gave evidence that individuals could shift the focus of attention from one aspect of the task to another, depending on situational cues. Since the task can be controlled in this way by both the experimenter and the subject, it provides an effective tool for studying subtle and detailed aspects of performance. Although in many cases no changes in strategy or trade-offs between performance on different task components were noted, this may have been because the task was relatively unstructured for the subjects i.e. they were given few specific instructions regarding how or in what direction to focus attention. When changes in strategy were found (Experiments II and III) the results enabled interpretations about the effects of the social manipulations which could not be inferred quite so directly from less complex tasks.

In the situations investigated in the preceding chapters, simple speed or output would appear to be the performance component most sensitive to effects from social stimuli. This is probably because speed was the most easily accessible index from which comparisons with others could be made. Here again, measures of speed are not appropriate for all performance tasks (for example pursuit rotor and darts) and may be relatively insensitive when performance goals are well defined.

7.3.5 Generality of Findings

A question not considered in the design of the present research but important in terms of its practical implications is whether the effects (facilitative in some and pacing in most cases) are specific to ad hoc, short term groups working under ambiguous and novel situations or are also characteristic of real groups. Some of the early work on social facilitation employed experimental designs in which individuals were tested repeatedly either alone or in a group and found that facilitative effects endured at least for the length of the experimental period (Allport, 1920; Whittemore, 1924). However, current research seldom uses repeated testing, probably because most studies employ large numbers of subjects in factorial designs. Interestingly, Carment and Hodkin (1973) found a cross-cultural difference in the response to others' presence between Canadian and Indian subjects. The Indian sample was generally less influenced by the presence of another, both in terms of competitive and pacing responses. The authors suggest that living in crowded conditions and within an extended familial structure may habituate Indian subjects to others' presence, the implication being that mere presence effects may not endure with repeated exposure. At present there is little direct evidence pertaining to the longevity of facilitative effects due to these social situations.

There is, however, literature from several areas of applied research which highlight the conflict experinenced by individuals who depart radically from their peer group norm. Unfortunately most of this work has focused on the debilitating effects of performing at a level which is above the group norm. Studies of gifted or genius children describe the conflict and social isolation often experienced by highly talented children, which is sometimes accompanied by adjustment difficulties and low achievement (Zorbaugh et al., 1951; Torrance, 1962; Lewis, 1943). Kimball and Leahy (1976), looking at a cross section of age groups, report a marked tendency for both males and females to show greater fear of success imagery as a function of increasing age; the highest success anxiety appearing in the mid-adolescent years and then declining somewhat. Reports from industrial settings also point out the pressure placed on highly proficient workers to suppress performance, particularly where piece work or quotas are involved (Roethlisberger and Dickson, 1939;

Rommetveit, 1955). Whyte (1943) has also illustrated that even in games activities performance can be dependent on the individual's status in the group: in his sample players of high ability but low status usually performed poorly in group competitions.

Although these studies have dealt with the suppression of performance by competent group members, logically this system of peer group pressure should extend to those who are extremely incompetent, leading them to exert more effort in order to maintain their group membership. Superficially the patterns of performance change noted in the present studies are analogous to those described in these applied situations. However, there are obvious differences between ad hoc groups and real groups; in the latter case members are presumably motivated to modify their behaviour in order to remain in the group. Pacing in the former case may simply be a response of the individual to an ambiguous situation. In real groups the appropriate standards of behaviour should be understood by group members and any pacing which occurs would be a demonstration of compliance to that standard.

The puzzling feature of the pacing patterns in the present data is why they should occur at all in ad hoc groups and especially under conditions which specifically stress individual competence or achievement. Festinger (1954) in his outline of social comparison processes maintains that individuals have a need or desire to evaluate their performance level, which accounts for comparison. However, it is less clear why comparison should then lead to adjustments in the actual level of performance, particularly for competent individuals. Festinger also discusses the social pressures against deviation from the group norm but questions whether performance (as opposed to attitudes or opinions) is likely to suffer from such social pressure because of the strong overall cultural emphasis on achievement. Even if such pressure did encourage uniformity of performance in real groups, it seems unlikely that this pressure would feature prominently in a group in which members were strangers and would never be likely to meet again, or who would certainly not meet under the same circumstances.[4]

A possible explanation is that behaviour which is learned in real group situations may generalize to the ad hoc groups in the experimental setting. This suggests a model of performance in short term social situations in which initial uncertainty leads to information seeking, which, as a result of behaviour learned in other social situations, results in the matching of individual behaviour (in this case performance) to the perceived group standard. Such a model provides a rationale for the results obtained and suggests that, although the social situations studied in the experimental laboratory may be stripped of many features of natural situations, their superficial similarity may be sufficient for eliciting behaviour characteristic of natural situations.

An alternative explanation is that the pacing patterns observed in these groups may serve an adaptive function. It was suggested previously that the effective significance of pacing is that it makes evaluation impossible. Lazarus (1966) points out that stress reactions are reflections of coping processes and are aroused to combat anticipated harm. The reactions observed in these experimental situations may function to protect the individual and his peers from failure and/or social embarrassment. This may result in performance disruption, <u>from the experimenter's point of view</u>. The subject may be:

> "motivated less to perform the experimenter's task and more to cope with the appraised threat. The threat serves to mobilize highly organized ways of coping ..., even if these ways serve to undermine the experimentally defined task performance." (Lazarus, 1966, p.360)

---

4. Experiments IV and V used high school pupils who did know each other.

Chapman (1975) suggests that laughter may serve a similar adaptive function in coacting situations. Subjects in his experiment laughed more to homorous stimuli when in close proximity, presumably directing attention away from the social situation and thereby reducing social arousal. This is an attractive theoretical framework for these group effects for at least two reasons. Firstly, it does not restrict the examination of the subjects' behaviour to a few task related responses. Secondly, it takes into account both the structure of the situation imposed by the experiment (coaction, audience, ego-threat etc.) and the subjects' redefinition of the situation and attempts at modification. Viewed in this way the experimental situation can be regarded as a realistic social situation in which the subject not only reacts to his environment but exerts some degree of influence upon it.

7.4 MODEL FOR PERFORMANCE IN SOCIAL SITUATIONS

The previous sections suggested several situational factors, which may or may not be important determinants of performance in social-evaluative situations, and also discussed several theoretical points of view. Clearly the favoured interpretation of the data from the foregoing experiments is one which views performance within the context of the larger social situation. Basic biological variables such as arousal are considered to be significant when some stress additional to mere presence is applied (e.g. evaluative instructions, competition). Heightened arousal may, under certain conditions, affect performance directly (facilitating the emission of dominant responses). But more often its effects are clouded or indiscernible because of the various strategies subjects adopt in attempting to cope effectively with the stress situation. Coping responses are therefore considered to be of primary significance in studies of social-evaluative stress and performance. The effects of arousal level may be most evident in those situations in which subjects responses are severely limited, either because needed information is not supplied or because there are few or no alternative ways of carrying out the task.

Figure 7.1 is a flow diagram of the probable sequence which subjects follow in assessing and organizing responses to the social situation. It is not a categorization of situations (e.g. Foot, 1973), but a cognitive model of performance behaviour under social-evaluative stress. The focus is placed on the subjects' adaptive abilities, and the model indicates a number of choice points at which adaptive attempts may be effectively blocked. Many of the situational factors previously discussed are incorporated, and the choice points indicate where in the cognitive sequence their particular characteristics would be important.

The model is intended not only to provide a structure in which all of these divergent factors may be integrated, but also to suggest a format for further research. As already noted, previous research has not been highly systematic, and many important variables have escaped notice altogether. The choice points in the model represent those areas where more systematic attention is needed and also suggest what factors should be controlled in order to adequately assess the separate effects of whatever variables are being manipulated.

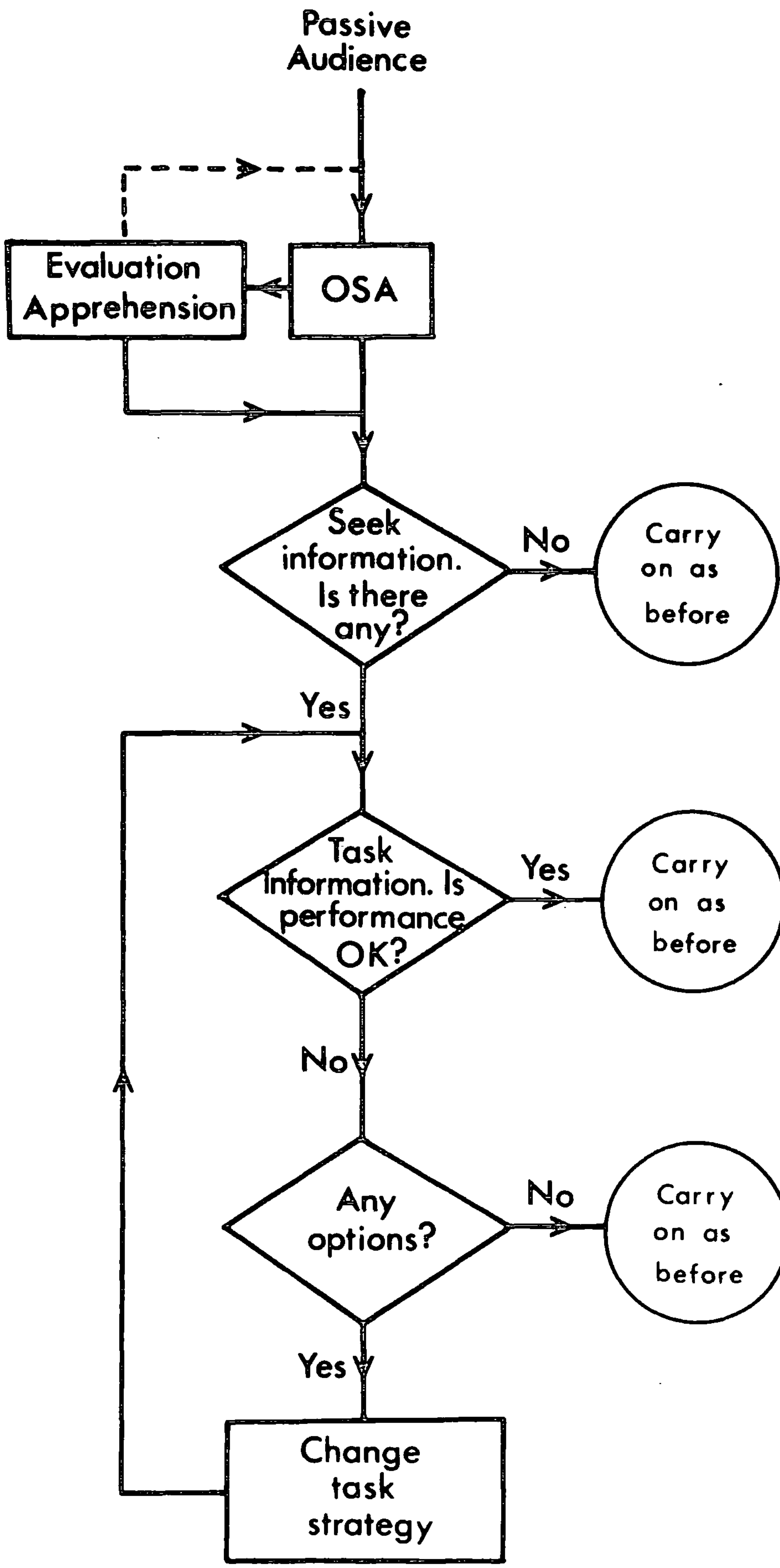
In Figure 7.1 two flow diagrams are presented which correspond to the two general social situations of interest, i.e. passive audience and coaction: the latter includes situations in which coactors work in turns and is termed 'participant audience'. It is immediately evident from the figure that of these two general situations coaction is the more complex. This is because, unlike audiences, coactors may be viewed as either evaluators or competitors. Also more sources of

FIGURE

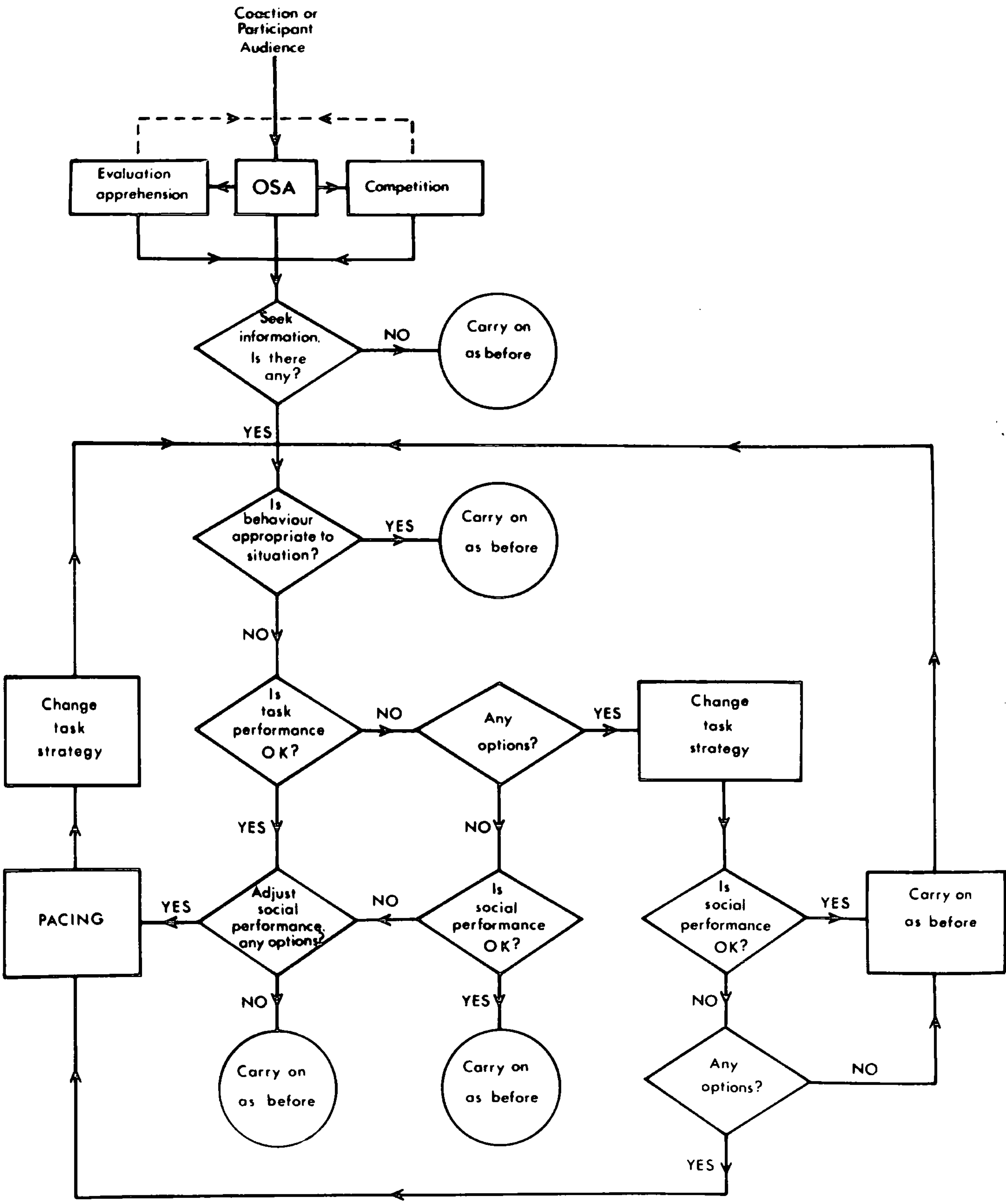
7.1 MODEL OF PERFORMANCE IN SOCIAL SITUATIONS

These two diagrams depict one way in which cognitive events in social/performance situations may be conceptualized. Only mere presence situations are considered. Several of the situational variables discussed in the preceding pages are incorporated, showing where and in what ways they may be influential. Certain symbols have been borrowed from computer language in order to highlight qualitatively different or similar stages in the sequence. For example, circles signify points at which the sequence terminates: diamonds indicate decision or choice points. However, the diagrams do not represent formal computer systems. Rather, the model outlines a general scheme by which individuals can attempt to arrive at behaviour which is appropriate to the particular situation or be blocked from so doing. The sequential placement of stages is at times arbitrary, e.g. attention to task information precedes that given to social information. This need not be the case and is a question for future research, although which comes first does not affect the final outcome. Important questions for research are located at the choice boxes (diamonds) e.g. what conditions determine whether or not social performance is judged acceptable? These questions do affect whether the sequence proceeds further or terminates and hence whether changes in performance strategies can be anticipated.

(a) Passive Audience



(b) Coaction or Participant Audience



information are available in the coaction situation. In any event, coaction presents more difficulties in terms of experimental study.

The first subject response in the sequence is considered to be objective self-awareness (OSA). This is perhaps a broader rendering of the term than is used by Duval and Wickland (1972). However, in order for audiences or coactors to have any effect at all, the subject must first perceive that others are present and that their activity does not compliment his own, i.e. they are either watching him or working non-cooperatively on the same task. Awareness of others as such also implies awareness of one's distinctiveness, and hence objective self-awareness. Evaluation apprehension or competition may follow as a consequence of this awareness and may induce even greater OSA. How the others are appraised is presumably related to the instructions given, whether or not the task permits evaluation and the general context of the performance situation (e.g. an academic test versus an experiment). Self-awareness, of course, need not lead to competition or evaluation apprehension, although, as discussed previously, a completely neutral interpretation of the others is probably rather rare, at least for short term groups. Cognitive appraisal is not represented at any distinct point in the flow diagram, rather it is considered that the whole sequence represents almost continuous evaluation, appraisal and reappraisal.

However the subject construes the situation (as competitive, etc.) his next action will be to seek information regarding appropriate behaviour, in this case performance behaviour, since it is the primary activity. Some tasks provide no such information (e.g. vigilance, in which subjects are unaware of undetected signals and no KR is provided). In audience situations this eventuality results in termination of the sequence, as subjects cannot evaluate their efficiency and hence are likely to maintain the same performance strategy. If, however, task information is available, either in the form of intrinsic feedback or KR, other courses of action are open. Performance may be judged acceptable, in which case no change in strategy should occur and the sequence terminates, or unacceptable, which should lead to a search for alternative strategies.

Information seeking is more complex in the coaction model since social comparison is possible. In some cases this information may be restricted to facial expressions or other non-performance behaviours. However, in most cases, social comparison is likely to also provide task information (e.g. pages turned, time to finish). In addition, when task information is available, this may highlight social comparison. Thus, these two information sources are presented as being inter-related.

Evaluation of performance is also more complex in the coaction situation. Subjects may judge their performance to be acceptable in terms of task criteria but unacceptable in terms of appropriate social behaviour (discrepancy from group mean) or vice versa. One source of information may therefore be rejected, depending on which is judged to provide the most appropriate information for the particular situation. For example, in an academic situation or athletic competition, unacceptable (discrepant) social comparison information may be irrelevant or may be used only as a guide for evaluating task efficiency. If performance is judged acceptable on both social and task criteria the sequence terminates, as no change in strategy is demanded. If, however, it is unacceptable on one or both, alternatives will be sought.

Here again the coaction model is more complex. The task may offer no alternative means of execution (e.g. the pursuit rotor), in which case the subject must either be content with inefficient performance or lower his efficiency criteria. In any event the sequence terminates.

This is true also of social performance criteria. However, in this latter case, there is an intermediate step between the perception of unacceptable performance and the adoption of an alternative strategy, this is the decision to pace. The adoption of a pacing strategy necessarily leads to some changes in the strategy for task execution.

Having adopted an alternative strategy, subjects then return to the information seeking stage and assess whether their action has in fact produced the intended effect. The most successful, or maximally adaptive, outcome is when the subject is satisfied that his performance is acceptable on all criteria and, having achieved a successful performance strategy, terminates the sequence. Obviously in this model, successful performance is not viewed as an end in itself. The subject is seen as attempting to cope with the imposed stress using the only device available to him, his performance. Successful performance has the effect of either satisfying competitive goals, alleviating evaluation apprehension or reducing the ambiguity engendered by heightened objective self-awareness.

In conclusion, a structure is provided for conceptualizing performance behaviour in social-evaluative situations. The model suggested is general and intended to be applicable to all mere presence situations and all tasks. However, little is said regarding the range of alternative strategies which may be adopted. The model suggests a positive approach to the stress situation which should lead to successful coping. However, it also points out that adaptive behaviour may be blocked at several points due to either lack of information or suitable alternatives. The implication is that performance will remain unchanged at these points or will be determined by non-cognitive factors, such as arousal. Still, it should be pointed out that at any stage in the sequence, subjects have the alternative of abandoning effort or psychologically 'leaving the field'. Although this may not be viewed as adaptive in any positive sense, it is a recognized means of coping with stress (Lazarus, 1966).

A defensive strategy is probably determined more by personal rather than situational factors and is therefore not incorporated in the present model. What the model does provide is a delineation of the important decision points in the performance sequence and the alternatives provided by the situation at each point. In general, research up to the present has provided little information about what should happen at these choice points. Most experimental work has been concerned exclusively with decisions only at the initial stage of appraisal (e.g. competition, evaluation apprehension). Clearly more attention needs to be directed at task variables such as the variety and quality of information provided by the task and also the alternative performance strategies available.

Secondly, in the case of coaction, there is little information to date about the conditions which predispose individuals to choose either task or social comparison information as the relevant performance criteria. Presumably this is related to decisions made earlier in the sequence about the role of the others and their relationship to the performer. However, it has always been assumed that which ever way the subject perceives the others, his reactions would be determined by the overriding importance of achieving good task performance. The present research indicates that this is not always the case, and that behaving appropriately in terms of the social situation may be more important than performing appropriately in terms of the task criteria.

Finally, more information is needed about alternative strategies. Pacing was specified in the present model as the most obvious alternative to unacceptable social comparison information. Other alternatives may be possible, depending on the situation and the degree of interaction possible. These include various non-verbal behaviours, laughter or other anxiety reducing devices or even overtly challenging the validity of the social stress situation.

7.5 FINAL REMARKS

Perhaps the most outstanding general observation arising from the present research is that the social situations investigated are far more dynamic than previous literature has suggested. Some definitions of 'group' specifically exclude ad hoc assemblies of individuals for which no interaction or commonality of interest is obvious (Cartwright and Zander, 1968). Yet, the data presented here clearly show that even the most simple social settings, where interaction is greatly restricted, can affect performance in a number of different ways; in simple facilitation (or impairment), social comparison and pacing and changes in the direction of attention and in strategies for task performance. This observation has several implications.

Firstly, apart from the literature specifically dealing with social facilitation, the mere presence of others has for the most part been uncontrolled in many areas of psychological investigation. Although 'mere presence' per se has not received much support as an important factor in determining performance outcomes, it should be clear that situations which use group testing procedures have the potential for eliciting a variety of responses from subjects which would not be possible in conditions of individual testing. That the simple presence of other subjects could affect performance outcomes in areas of research quite unrelated to social stress or environmental variables suggests that caution be exercised regarding the social environment in which performance takes place. The addition of one or two subjects working independently may have no effect or may have a dramatic effect.

A further question arising from these data is whether the current preoccupation with 'social facilitation' in the literature is warranted or likely to lead to fruitful theories which have practical applications. Although facilitation was noted on occasion in the present work and has been demonstrated many times in previous literature, there is uncertainty about the size of these effects as well as the specific conditions under which they can be expected to occur. Social facilitation has been given the status of a general and stable phenomenon, but the evidence given here suggests that its occurrence may be restricted to only a narrow range of situations. Exclusive focus on demonstrating facilitative effects using various tasks and degrees of evaluation stress will probably continue to produce some positive findings as well as many insignificant or equivocal ones. Not only is the search for facilitation perhaps sterile and wasteful, there is some danger that it may obstruct the investigation of interesting patterns of social behaviour within this type of restricted group setting.

Furthermore, performance measures have for the most part been used with little inspiration in the typical social facilitation experiment. Tasks have been either very simple (to show facilitation of output) or especially designed with specified response hierarchies (to show facilitation of dominant responses). Little of the enormous work on human performance and information processing i.e. manipulations of task strategies or variations in task stressfulness such as rate of stimulus presentation, has been applied to research on the effects of social situations. The presence of others may affect higher order performance processes in ways analogous to other types of physical environmental stress (noise) or may have quite different effects because of interaction with social processes. This possibility in itself merits attention.

Finally, the pacing patterns observed in the present experiments have practical implications. The results seem particularly relevant for educational settings, perhaps especially so for adolescent pupils. Although current feeling seems to be that individuals of high ability are generally not adversely affected by performing in groups where others are of less relative ability, the present studies indicate otherwise. Alternatively, individuals of low ability were quite often favourably influenced by exposure to others with a range of abilities. Even though there is some question about the generality and duration of the pacing effects observed, these experimental situations may provide a means for systematically studying an important social process. The effects on pacing of different instructional sets (cooperative versus competitive), individual differences in susceptibility to social pressure, age and sex differences and the degree of group cohesiveness all may be important variables in real as well as experimental situations.

The present work has perhaps raised more questions than it has answered, but in part it was intended to illustrate the complexity and dynamic nature of non-interactive social situations. Hopefully, future research will address some of the issues discussed in this chapter and will lead to a fuller understanding of social processes and their effects on human performance.

REFERENCES

REFERENCES

ABEL, T.M. (1938) The influence of social facilitation on motor performance at different levels of intelligence. American Journal of Psychology, 51, 379-89.

ADER, R. & TATUM, R. (1963) Free operant avoidance conditioning in individual and paired human subjects. Journal of Experimental Analysis of Behaviour, 6, 357-9.

ALLEN, J.L. & BOIVIN, M.R. (1976) "Women's will to fail" in a disjunctive reaction time competitive task. Bulletin of the Psychonomic Society, 8, 401-2.

ALLPORT, F.H. (1920) The influence of a group up on association and thought. Journal of Experimental Psychology, 3, 159-82.

ALLPORT, F. (1924) Social Psychology. Haughton Mufflin Co.

ALLSOPP, J.F. & EYSENCK, H.J. (1974) Personality as a determinant of paired-associates learning. Perceptual and Motor Skills, 39, 315-24.

ALPER, T.G. (1948) Task orientation and ego orientation as factors in reminiscence. Journal of Experimental Psychology, 38, 224-38.

ALPERT, R. & HABER, R.N. (1960) Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 61, 207-15.

ANASTASI, A. (1968) Psychological Testing (3rd Ed.), London: Macmillan.

ANDERSON, C.A. (1929) An experiment study of "Social Facilitation" as affected by intelligence. American Journal of Sociology, 134, 874-81.

ANGELINI, D. (1973) Immediate elder sibling as a factor in task performance. Journal of Individual Psychology, 29, 167-173.

ANNETT, J. (1969) Feedback and human behaviour, Penguin.

APPLEY, M.H. & TRUMBULL, R. (Eds.)(1967) Psychological Stress, New York: Appleton-Century-Crofts.

ARGYLE, M. (1969) Social Interaction, London: Methuen.

ARGYLE, M. (1972) The Social Psychology of Work, Allen Lane: Penguin.

ARGYLE, K. & LITTLE, B.R. (1972) Do personality traits apply to social behaviour? Journal for the Theory of Social Behaviour, 2, 1-35.

ARGYLE, M. & ROBINSON, P. (1962) Two origins of achievement motivation. British Journal of Social and Clinical Psychology, 1, 107-120.

ARGYLE, K. & WILLIAMS, M. (1969) Observer or observed? A reversible perspectic in person perception. Sociometry, 32, 396-492.

ARGYRIS, C. (1968) Some unintended consequences of rigorous research. Psychological Bulletin, 70, 185-97.

ATKINSON, J.W. & CARTWRIGHT, D. (1964) Some neglected variables in contemporary conceptions of decision and performance. Psychological Reports, 14, 575-90.

ATKINSON, J.W. & LITWIN, G.H. (1960) Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. Journal of Abnormal and Social Psychology, 60, 52-63.

ATKINSON, J.W. & REITMAN, W.R. (1956) Performance as a function of motive strength and expectancy of goal attainment. Journal of Abnormal and Social Psychology, 53, 361-6.

BADDELEY, A.D. (1968) A 3-minute reasoning test based on grammatical transformations. Psychonomic Science, 10, 341-2.

BALL, D.W. (1972) The definition of situation: some theoretical and methodological consequences of taking W.I. Thomas seriously. Journal for the Theory of Social Behaviour, 2(1), 61-83.

BAUM, M. (1968) Extinction of an avoidance behaviour through symbolic modelling. Journal of Personality and Social Psychology, 8, 99-108.

BAYER, E. (1929) Beitrage zur zweikomponententheorie des hungers. Zsch. f. Psychol, 112, 1-53. Cited in Harlow, H.F. Social facilitation of feeding in the albino rat. Journal of Genetic Psychology,(1932) 43, 211-21.

BEEZ, V.W. (1972) Influence of biased psychological reports on teacher behaviour and pupil performance. In Len Sperry (Ed.) Learning Performance and Individual Differences: Essays and Readings, Glenview, Ill. Scott, Foresman and Co. pp.86-92.

BELCHER, T. (1975) Effect of different test situations on creativity scores. Psychological Reports, 36(2), 511-15.

BEM, D.J. & ALLEN, A. (1974) On predicting some of the people of the time: the search for cross-situational consistencies in behavior. Psychological Review, 81(6), 506-21.

BERGER, S.M. (1968) Vicarious aspects of matched-dependent behavior, pp.169-185. In Simmel, E.C., Hoppe, R.A. and Milton, G.A. (Eds.) Social facilitation and imitative behavior, Boston: Allyn and Bacon.

BERGUM, B. & LEHR, D. (1962) Vigilance performance as a function of paired monitoring. Journal of Applied Psychology, 46, 341-3.

BERGUM,B.O. & LEHR, D.J. (1963) Effects of authoritarianism on vigilance performance. Journal of Applied Psychology, 47, 75-7.

BERKEY, A.S. & HOPPE, R.A. (1972) The combined effect of audience and anxiety on paired-associate learning. Psychonomic Science, 29 351-3.

BERKOWITZ, L. (1971) The "weapons effect", demand characteristics and myth of the compliant subject. Journal of Personality and Social Psychology, 20, 332-8.

BICKMAN, L. (1972) Social influence and diffusion of responsibility in an emergency. Journal of Experimental Social Psychology, 8, 438-45.

BILODEAU, E.A. & BILODEAU, I.McD. (1961) Motor skills learning. Annual Review of Psychology, 12, 243-80.

BIRD, A.M. (1975) Cross sex effects of subjects and audience during motor performance. Research Quarterly, 46, 379-84.

BIRNEY, R.C., BURDICK, H. & TEEVAN, R.C. (1969) Fear of failure. New York: Van Nostrand-Reinhold.

BLAKE, R.R., HELSON, H. & MOUTON, J.S. (1956) The generality of conformity behavior as a function of factual anchorage, difficulty of task, and amount of social pressure. Journal of Personality, 25, 294-305.

BOOR, M. & SCHILL, T. (1967) Digit symbol performance of subjects varying in anxiety and defensiveness. Journal of Consulting Psychology, 31, 600-3.

BORDEN, B.J., HENDRICK, C. & WALKER, J.W. (1976) Affective, physiological and attitudinal consequences of audience presence. Bulletin of the Psychonomic Society, 1, 33-7.

BOWERS, K.S. (1973) Situationism in psychology: an analysis and a critique. Psychological Review, 80, 307-36.

BRIGHTMAN, D. & RAYMOND, B. (1975) The effects of task ambiguity and expectancy control groups on the experimenter bias effect. Journal of Social Psychology, 96, 277-87.

BROADBENT, D.E. (1972) Individual differences in annoyance by noise. Sound, 6, 56-61.

BROADBENT, D.E.(1977) Levels, hierarchies and the locus of control. The Quarterly Journal of Experimental Psychology, 29, 181-203.

BROADHURST, P.L. (1959) The interaction of task difficulty and motivation: the Yerkes-Dodson law revived. Acta Psychology, 16, 321-38.

BROEN, W.E. & STORMS, L.H. (1961) A reaction potential ceiling and response decrements in complex situations. Psychological Review, 68, 405-15.

BROWN, B.R. & GARLAND, H. (1971) The effects of incompetency and acquaintanceship and anticipated evaluative feedback on face saving behavior. Journal of Experimental Social Psychology, 7, 490-502.

BROWN, M. (1974) Motivational correlates of academic performance. Psychological Reports, 34(3), 746-7.

BROWN, M., JENNINGS, J. & VANIK, V. (1974) The motive to avoid success: a further examination. Journal of Research in Personality, 8(2), 172-7.

BUCK, R.W. & PARKE, R.D. (1973) Behavioural and physiological response to the presence of a friendly or neutral person in two types of stressful situations. Journal of Personality and Social Psychology, 24, 143-53.

BURNHAM, W.H. (1910) The group as a stimulus to mental activity. Science, 31, 761-7.

BURRI, C. (1931) The influence of an audience upon recall. Journal of Educational Psychology, 22, 683-90.

BURSILL, A.E. (1958) The restriction of peripheral vision during exposure to hot and humid conditions. Quarterly Journal of Experimental Psychology, 10, 113-29.

BURWITZ, L. & NEWELL, K.M. (1972) The effects of mere presence of coactors on learning a motor skill. Journal of Motor Behaviour, 4, 99-102.

CALLOIS, R. (1961) Man, Play and Games, Glencoe: Free Press.

CARMENT, D.W. (1970) Rate of simple motor responding as a function of cooperation, competition and sex of participant. Psychonomic Science, 19, 342-3.

CARMENT, D.W. & HODKIN, B. (1973) Coaction and competition in India and Canada. Journal of Cross-Cultural Psychology, 4, 435-59.

CARMENT, D.W. & LATCHFORD, M. (1970) Rate of simple motor responding as a function of coaction, sex of the participants, and the presence or absence of the experimenter. Psychonomic Science, 20.

CARRON, A.V. & BENNETT, B. (1976) The effects of initial habit strength differences upon performance in a coaction situation. Journal of Motor Behaviour, 8, 297-304.

CARTWRIGHT, D. & ZANDER, A. (1968) Group Dynamics, New York: Harper Row.

CHAPMAN, A.J. (1973a) An electromyographic study of apprehension about evaluation. Psychological Reports, 33, 811-4.

CHAPMAN, A.J. (1973b) Social facilitation of laughter in children. Journal of Experimental Social Psychology, 9, 528-41.

CHAPMAN, A.J. (1974a) An electromygraphic study of social facilitation: a test of the 'mere presence' hypothesis. British Journal of Psychology, 65, 123-9.

CHAPMAN, A.J. (1974b) An experimental study of socially facilitated 'humorous laughter'. Psychological Reports, 35, 727.

CHAPMAN, A.J. (1975) Eye contact, physical proximity and laughter - re-examination of equilibrium model of social intimacy. Social Behaviour and Personality, 3, 143-57.

CHAPMAN, A.J. & CHAPMAN, W.A. (1974) Responsiveness to humor: its dependency upon a companion's humorous smiling and laughter. Journal of Psychology, 88, 245-52.

CHEN, S.C. (1937) Social modification of the activity of ants in nest-building. Physiological Zoology, 10, 420-436.

CHURCH, R.M. (1962) The effects of competition on reaction time and palmer skin conductance. Journal of Abnormal and Social Psychology, 65, 32-40.

CHURCH, R.M. (1968) Applications of behavior theory to social psychology: imitation and competition, 135-67. In Simmel, E.C., Hoppe, R.A. & Milton, G.A. (Eds.) Social facilitation and imitative behavior, Boston: Allyn and Bacon.

CLARIDGE, G.S. (1967) Personality and Arousal, London: Pergamon Press.

CLARK, R.D. III & WORD, L.E. (1972) Why don't bystanders help? Because of ambiguity? Journal of Personality and Social Psychology, 24, 392-400.

CLARK, R.D. III & WORD, L.E. (1974) Where is the apathetic bystander? Situational characteristics of the emergency. Journal of Personality and Social Psychology, 29, 279-87.

CLAYTON, D.A. (1976) Social facilitation of drinking of a partially satiated duckling. Animal Learning and Behavior, 4, 391-5.

CLIFFORD, M.M. (1971) Motivational effects of competition and goal setting in reward and non-reward conditions. Journal of Experimental Education, 39, 11-16.

COHEN, J.L. & DAVIS, J.H. (1973) Effects of audience status, evaluation, and time of action on performance. Journal of Personality and Social Psychology, 27, 74-85.

COHEN, P.A. & SHEPOSH, J.P. (1977) Audience and level of esteem as determinants of risk-taking. Personality and Social Psychology Bulletin, 3, 119-23.

COLQUHOUN, W.P. & CORCORAN, D.W.J. (1964) The effects of time of day and social isolation on the relationship between temperament and performance. British Journal of Social and Clinical Psychology, 3, 226-31.

CORCORAN, D.W.J. (1965) Personality and the inverted U relation. British Journal of Psychology, 56, 267-73.

COSTELLO, C.G. (1964) Ego-involvement, success and failure: a review of the literature, 161-208. In Eysenck, H.J. (Ed.), Experiments in Motivation, London: Pergamon Press.

COSTELLO, C.G., FELDMAN, M.P. & SLATER, P.A. (1964) The effect of success and failure reports on pursuit rotor performance and reminiscence. In Eysenck, H.J. (Ed.) Experiments in Motivation, London: Pergamon Press.

COTTRELL, N.B. (1968) Performance in the presence of other human beings: mere presence, audience and affiliation effects. In Simmel, E.C., Hoppe, R.A. and Milton, A.A. (Eds.) Social facilitation and imitative behaviour, Boston: Allyn & Bacon.

COTTRELL, N.B. (1972) Social facilitation. In C.B. McClintock (Ed.), Experimental Social Psychology, N.Y.: Holt, Rinehart + Winston.

COTTRELL, N.B., RITTLE, R.H. & WACK, D.L. (1967) Presence of audience and list type (competitional or noncompetitional) as joint determinants of performance in paired-associate learning. Journal of Personality, 35, 217-26.

COTTRELL, N.B., WACK, D.L., SEKERAK, G.J. & RITTLE, R.H. (1968) The social facilitation of dominant responses by the presence of an audience and the mere presence of others. Journal of Personality and Social Psychology, 9, 251-256.

CRATTY, B.J. (1967) Social Dimensions of Physical Activity, Englewood Cliffs, N. Jersey: Prentice-Hall Inc.

CRAWFORD, M.P. (1939) The social psychology of the vertebrates. Psychological Bulletin, 36, 407-46.

CRIDDLE, W.D., (1971) The physical presence of other individuals as a factor in social facilitation. Psychonomic Science, 22, 229-30.

CROWNE, D.P. & MARLOWE, O. (1964) The Approval Motive: Studies in Evaluation Dependence, New York: Wiley.

CRIDDLE, W.D. (1971) The physical presence of other individuals as a factor in social facilitation. Psychonomic Science, 22, 229-30.

CRUTCHFIELD, R.S. (1955) Conformity and Character. American Psychologist, 10, 191-8.

DASHIELL, J.F. (1930) An experimental analysis of some group effects. Journal of Abnormal and Social Psychology, 25, 190-199.

DASHIELL, J.F. (1935) Experimental studies of the influence of social situations on the behavior of individual human adults. In Murchison, C. (Ed.) A Handbook of Social Psychology, Worcester, Mass: Clack University Press.

DAVIES, D.R. & HOCKEY, G.R.J. (1966) The effects of noise and doubling the signal frequency on individual differences in visual vigilance performance. British Journal of Psychology, 57, 381-89.

DAVIES, D.R., HOCKEY, G.R.J. & TAYLOR, A. (1969) Varied auditory stimulation, temperament differences and vigilance performance. British Journal of Psychology, 60, 453-457.

DAVIS, J.H. (1969) Group Performance. Reading, Mass: Addison-Wesley.

DAVITZ, J.R. & MASON, D.J. (1955) Socially facilitated reduction of a fear response in rats. Journal of Comparative and Physiological Psychology, 48, 149-151.

de CHARMS, R. & DAVÉ, P.N. (1965) Hope of success, fear of failure, subjective probability and risk-taking behavior. Journal of Personality and Social Psychology, 1, 558-68.

DEFENBACHER, K.A., PLATT, G.J. & WILLIAMS, M.A. (1974) Differential recall as a function of socially induced arousal and retention interval. Journal of Experimental Psychology, 103, 809-12.

DOCTOR, R.M. & ALTMAN, F. (1969) Worry and emotionality as components of test anxiety: replication and further study. Psychological Reports, 24, 563-8.

DOHRENWEND, B.P. (1961) The social psychological nature of stress. Journal of Abnormal and Social Psychology, 62, 294-302.

DROPPLEMAN, L.F. & McNAIR, D.M. (1971) An experimental analog of public speaking. Journal of Consulting and Clinical Psychology, 36, 91-6.

DUBEY, R.S. (1976) Manifest anxiety and educational performance. Psychological Studies, 21, 52-9.

DUFFY, E. (1962) Activation and Behavior. New York: Wiley.

DUNN, J.A. (1968) Anxiety, stress and the performance of complex intellectual tasks: a new look at an old question. Journal of Consulting and Clinical Psychology, 32, 669-73.

DUVAL, S. & WICKLUND, R.A. (1972) A Theory of Objective Self-Awareness. London: Academic Press.

EASTERBROOK, J.A. (1959) The effect of emotion on the utilization and the organization of behavior. Psychological Review, 66, 183-201.

EDWARDS, A.L. (1968) Experimental Design in Psychological Research, 3rd Edition, Holt, Rinehart and Winston, Inc.

EKDAHL, A.G. (1929) Effect of attitude on free association-time. Genetic Psychology Monograph, 5, 253-338.

ELLIOTT, G. & BARTEE, H.H. (1976) Motivational effect of simulated videotaping on performance. Perceptual and Motor Skills, 42, 1193-4.

ELMS, A.C. (1972) Social Psychology and Social Relevance. Boston: Little, Brown and Company.

EMILEY, S.F. (1975) Effects of crowding and interpersonal attraction on affective responses, task performance and verbal behaviour. Journal of Social Psychology, 97, 267-79.

ENDLER, N.S. (1975) A person-situation interaction model for anxiety, pp.145-164. In Spielberger, C.D. and Sarason, I.G. (Eds.) Stress and Anxiety, Vol.1. Washington D.C: Hemisphere Publishing Corporation.

EPLEY, S. (1974) Reduction of the Behavioural effects of aversive stimulation by the presence of companions. Psychological Bulletin, 81, 271-83.

ERIKSEN, C.W. LAZARUS, R.S. & STRANGE, J.R. (1952) Psychological stress and its personality correlates. II The Rorschach test and other personality measures. Journal of Personality, 20, 277-86.

ERIKSON, E.H. (1965) Childhood and Society, Harmondsworth, Penguin.

EYSENCK, H.J. (1957) The Dynamics of Anxiety and Hysteria, London: Routledge and Kegan Paul.

EYSENCK, H.J. (1966) Personality and Experimental Psychology, Bulletin of the British Psychological Society, 19, 1-28.

EYSENCK, H.J. (1967) The Biological Basis of Personality, Springfield, Ill.: Charles C. Thomas.

EYSENCK, H.J. (1972) Personality and attainment: an application of psychological principles to educational objectives. Higher Education, 1, 39-52.

EYSENCK, H.J. & EYSENCK, S.B. (1964) Manual of the Eysenck Personality Inventory, University of London Press Ltd.

FARBER, I.E. (1955) The role of motivation on verbal learning performance. Psychological Bulletin, 52, 311-27.

FARBER, I.E. & SPENCE, K.W. (1953) Complex learning and conditioning as a function of anxiety. Journal of Experimental Psychology, 45, 120-5.

FARNSWORTH, P.R. (1938) Concerning so-called group effects. Journal of Genetic Psychology, 35, 587-94.

FEATHER, N.T. (1965) Performance at a difficult task in relation to initial expectancies of success, test anxiety and need achievement. Journal of Personality, 33, 200-17.

FELDMAN, M.P.(a)(1964) Motivation and task performance: a review of the literature. In Eysenck, H.J. (Ed.), Experiments in Motivation, Pergamon Press, London.

FELDMAN, M.P.(b)(1964) Pursuit rotor performance and reminiscence as a function of drive level. In Eysenck, H.J. (Ed.), Experiments in Motivation, Pergamon Press, London.

FESTINGER, L.A. (1954) A theory of social comparison processes. Human Relations, 7, 117-140.

FESTINGER, L.A. (1957) Theory of Cognitive Dissonance, Evanston: Row, Peterson.

FIELDING, L., MURRAY, J. & STEEL, D.H. (1976) The influence of spectator reaction and presence during training on performance. International Journal of Sport Psychology, 7, 73-82.

FISHER, S. (1975) The microstructure of dual task interaction. 2. The effect of task instructions on attentional allocation and a model of attention-switching. Perception, 4, 459-75.

FITTS, P.M. & POSNER, M.D. (1967) Human Performance, Wadsworth Publishing Co. Inc. Brooks/Cole Publishing Co., Belmont, Calif.

FLEISHMAN, E.A. (1958) A relationship between incentive motivation and ability level in psychomotor performance. Journal of Experimental Psychology, 56, 78-81.

FLEISHMAN, E.A. (1975) Toward a taxonomy of human performance. American Psychologist, 30, 1127-49.

FOOT, H.C. (1973) Group learning and performance: a reclassification. British Journal of Social and Clinical Psychology, February 1973, 7-17.

FOOT, H.C. & LEE, T.R. (1970) Social feedback in the learning of a motor skill. British Journal of Social and Clinical Psychology, 9, 309-19.

FRANKENHAEUSER, M. & PATKAI, P. (1964) Catecholamine excretion and performance during stress. Perceptual Motor Skills, 19, 13-4

FREEDMAN, J.L., KLEVANSKY, S. & EHRLICH, P. (1971) The effect of crowding on human task performance. Journal of Applied Social Psychology, 1, 7-25.

FREEDMAN, J.L., LEVY, A.S., BUCHANAN, R.W. & PRICE, J. (1972) Crowding and human aggressiveness. Journal of Experimental Social Psychology, 8, 528-48.

FRIEDMAN, S., MASON, J. & HAMBURG, D. (1963) Urinary 17 - Hydroxycortico - Steriod levels in parents of children with neoplastic disease. Psychosomatic Medicine, 25, 364-76.

FREISCHLAG, J. (1974) Stress in pursuit-rotor performance of men and women physical education majors. Perceptual and Motor Skills, 38, 785-7.

GANZER, V.J. (1968) Effects of audience presence and test anxiety on learning and retention in a serial situation. Journal of Personality and Social Psychology, 8, 194-9.

GATES, G.S. (1924) The effect of an audience upon performance. Journal of Abnormal and Social Psychology, 18, 334-344.

GEEN, R. (1971) Social facilitation of long-term recall. Psychonomic Science, Vol.24. 89-90.

GEEN, R. (1973) Effects of being observed on short- and long-term recall. Journal of Experimental Psychology, 100, 395-8.

GEEN, R. (1976) Test anxiety, observation and range of cue utilization. British Journal of Social and Clinical Psychology, 15, 253-9.

GELLER, V. & SHAVER, P. (1976) Cognitive consequences of self-awareness. Journal of Experimental Social Psychology, 12, 99.

GELLERMAN, S.W. (1963) Motivation and Productivity, New York: Swets & Zeitlinger.

GERARD, H.B. (1963) Emotional uncertainty and social comparison. Journal of Abnormal and Social Psychology, 66, 568-73.

GIBSON, H.B. & CURRAN, J.D. (1974) The effect of distraction on a psychomotor task studied with reference to personality. The Irish Journal of Psychology, 3, 148-58.

GILL, D.L. & MARTENS, R. (1975) The informational and motivational influence of social reinforcement on motor performance. Journal of Motor Behaviour, 7, 171-83.

GLOVER, C.B. & CRAVENS, R.W. (1974) Trait anxiety, stress and learning. A test of Saltz's hypothesis. Journal of Research in Personality, 8, 243-54.

GOOD, K.J. (1973) Social facilitation: effects of performance anticipation, evaluation and response competition on free association. Journal of Personality and Social Psychology, 28, 270-5.

GRACE, H.A. (1951) Effects of different degrees of knowledge about an audience on the content of communication. Journal of Social Psychology, 34, 111-124.

GRANBERG, D., STEELE, L. & BRENT, E.E. Jr. (1975) Levelling effect of coaction on judgment or retest effect. Journal of Social Psychology, 96, 45-53.

GRAY, J.A. (1971) Strength of the nervous system, introversion-extraversion conditionability and arousal, 49-81. In Eysenck, H.J. (Ed.) Readings in Extraversion-Introversion, Vol.1, New York: Wiley.

GREENBERG, P.J. (1932) Competition in children: an experimental study. American Journal of Psychology, 44, 221-48.

GRUDER, C.L., KORTH, B., DICHTEL, M. & GLOS, B. (1975) Uncertainty and social comparison. Journal of Research in Personality, 9, 85-95.

GUPTA, V.P. (1973) The relationship of age, neuroticism and extraversion to persistence. Psychological Studies, 18, 20-7.

GURIN, G. & GURIN, P. (1972) Expectancy theory in the study of poverty. In Len Sperry (Ed.) Learning, Performance and Individual Differences: Essays and Readings, Glenview, Ill: Scott, Foresman & Co.

HAAS, J. & ROBERTS, G. (1975) Effect of evaluative others upon learning and performance of a complex motor task. Journal of Motor Behavior, 7, 81-91.

HACKMAN, J.R. (1970) Tasks and task performance in research on stress. In McGrath, J.E. (Ed.), Social and Psychological Factors in Stress, New York: Holt, Rinehart and Winston.

HAKE, D.F. & LAWS, D.R. (1967) Social facilitation of responses during a stimulus paired with electric shock. Journal of the Experimental Analysis of Behavior, 10, 387-92.

HAKE, D.F., VAKELICH, R. & KAPLAN, S.J. (1973) Audit responses: responses maintained by access to existing self or coactor scores during non-social, parallel work and cooperation procedures. Journal of the Experimental Analysis of Behavior, 19, 409-23.

HAKMILLER, K.L. (1966) Threat as a determinant of downward comparison. Journal of Experimental Social Psychology, 2, 32-9.

HAMILTON, P., HOCKEY, G.R.J. & REJMAN, M. (in press) The place of the concept of activation in human information processing: an integrative approach. In Dornic, S. & Rabbitt, P.M.A. (Eds.) Attention and Performance VI.

HAMILTON, P., WILKINSON, R.T. & EDWARDS, R.S. (1972) A study of four days partial sleep deprivation, 101-113. In Colquhoun, W.P. (Ed.) Aspects of Human Efficiency, The English Universities Press Ltd.

HANAWALT, N.G. & RUTTIGER, K.F. (1944) The effect of an audience on remembering. Journal of Social Psychology, 19, 259-72.

HARDESTY, D., TRUMBO, D. & BEVEN, W. (1963) Influence of KR on performance in a monitoring task. Perceptual and Motor Skills, 16, 629-34.

HARLOW, H.F. (1932) Social facilitation of feeding in the albino rat. Journal of Genetic Psychology, 43, 211-21.

HARNETT, J.J., GOTTLIEB, J. & HAYES, R.L. (1976) Social facilitation theory and experimenter attractiveness. Journal of Social Psychology, 99, 293-4.

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

HECKHAUSEN, H. (1972) The Anatomy of Achievement Motivation, Academic Press.

HEIDER, F. (1958) The Psychology of Interpersonal Relations, New York: Wiley.

HENCHY, T. & GLASS, D.C. (1968) Evaluation apprehension and the social facilitation of dominant and subordinate responses. Journal of Personality and Social Psychology, 10, 446-454.

HICKS, L.E. (1975) Some effects of anxiety and cognitive style upon pursuit rotor learning. British Journal of Social and Clinical Psychology, 14, 155-69.

HILL, A.B. (1975) Extraversion and variety seeking in a monotonous task. British Journal of Psychology, 66, 9-15.

HILLERY, J.M. & FUGITA, S.S. (1975) Social facilitation in employment testing. Educational and Psychological Measurement, 35, 745-50.

HOCKEY, G.R.J. (1970) Effect of loud noise on attentional selectivity. Quarterly Journal of Experimental Psychology, 22, 28-36.

HOCKEY, G.R.J. (1973) Rate of presentation in running memory and direct manipulation of input-processing strategies. Quarterly Journal of Experimental Psychology, 25, 104-11.

HOLLINGWORTH, H.L. (1935) The Psychology of the Audience, New York: American.

HORNE, W.C. (1971) Comment on Zajonc's social facilitation review including a reanalysis of Travis (1925) data. Perceptual and Motor Skills, 33, 996.

HORNER, M.S. (1974) The measurement and behavioural implications of fear of success in women, 91-117. In Atkinson, J.W. & Raynor (Eds.), Motivation and Achievement, New York: Wiley.

HOUSE, W.C. (1972) Actual and perceived differences in male and female expectancies and minimal goal levels as a function of competition. Journal of Personality, 42, 493-509.

HUNT, P.J. & HILLERY, J.M. (1973) Social facilitation in a coaction setting: an examination of the effects over learning trials. Journal of Experimental Social Psychology, 9, 563-72.

HURLOCK, E.B. (1927) Use of group rivalry as an incentive. Journal of Abnormal and Social Psychology, 22, 278.

HUSBAND, R.W. (1931) Analysis of methods in human maze learning. Journal of Genetic Psychology, 39, 258-77.

INNES, J.M. (1972) The effect of presence of co-workers and evaluative feedback on performance of a simple reaction time task. European Journal of Social Psychology, 2, 466-70.

INNES, J.M. & SAMBROOKS, J.E. (1969) Paired-associate learning as influenced by birth order and the presence of others. Psychonomic Science, 16, 109-10.

INNES, J.M. & YOUNG, R.F. (1975) The effect of presence of an audience, evaluation apprehension and objective self-awareness on learning. Journal of Experimental Social Psychology, 11, 35-43.

JANSSENS, L. & NUTTIN, J.R. (1976) Frequency perception of individual and group successes as a function of competition, coaction and isolation. Journal of Personality and Social Psychology, 34, 830-36.

JOHNSON, D.W. & JOHNSON, R.T. (1975) Learning Together and Alone: Cooperation, Competition and Individualization. Prentice-Hall Inc: Englewood Cliffs, New Jersey.

JONES, E.E. & GERARD, H.B. (1967) Foundations of Social Psychology, John Wiley & Sons Inc., New York, Sidney, London.

KAHN, R.L. (1970) Some propositions toward a researchable conceptualization of stress. In McGrath, J.E. (Ed.), Social and Psychological Factors in Stress, New York: Holt, Rinehart & Winston.

KAHNEMAN, D. (1973) Attention and Effort, Englewood Cliffs, N.J. Prentice-Hall Inc.

KANEKAR, S., NEELAKANTAN, P. & LALKAKA, P.K. (1975) Nominal and real group performance in relation to manifest anxiety and induced stress. Social Behavior and Personality, 3, 197-205.

KARABENICK, S.A. (1972) Effect of sex of competitor on the performance of females following success. Proceedings, 80th Annual Convention, A.P.A.

KARABENICK, S. & MARSHALL, J.M. (1974) Performance of females as a function of fear of success, fear of failure, type of opponent, and performance contingent feedback. Journal of Personality, 42, 220-38.

KARABENICK, S.T., MARSHALL, J.M. & KARABENICK, J.D. (1976) Effects of fear of success, fear of failure, type of opponent and feedback on female achievement performance. Journal of Research in Personality, 10, 369-85.

KARABENICK, S.A. & YOUSSEF, Z.I. (1968) Performance as a function of achievement motive level and perceived difficulty. Journal of Personality and Social Psychology, 10, 414-9.

KATCHMAR, L.T., ROSS, S. & ANDREWS, T.G. (1958) Effects of stress and anxiety on performance of a complex verbal coding task. Journal of Experimental Psychology, 55, 559-64.

KATZ, J., WEINER, H., GALLAGHER, T. & HELLMAN, L. (1970) Stress, distress and ego defences. Archives of General Psychology, 23, 131-42.

KELLEY, H.H. & THIBAUT, J.W. (1969) Group problem solving. In Lindzey,G. & Aronson, E. (Eds.), The Handbook of Social Psychology (2nd Ed.) (Vol.4). Reading, Mass.: Addison-Wesley.

KEPPEL, G. (1973) Design and Analysis: A Researcher's Handbook, Englewood Cliffs, N.J.: Prentice-Hall Inc.

KIESLER, S.B. (1966) Stress, affiliation and performance. Journal of Experimental Research in Personality, 1, 227-35.

KIESLER, C. & KIESLER, S. (1969) Conformity, Addison-Wesley Pub. Co. Inc.

KIMBALL, B. & LEAHY, R.L. (1976) Fear of success in males and females, effects of developmental level and sex-linked course of study. Sex Roles, 2, 273-81.

KIMBLE, G.A. (1950) Evidence for the role of motivation in determining the amount of reminiscence in pursuit rotor learning. Journal of Experimental Psychology, 40, 248-53.

KIRITZ, S. & MOOS, R.H. (1974) Physiological effects of social environments. Psychosomatic Medicine, 36, 96-115.

KLINE, P. & GALE, A. (1971) Extraversion, neuroticism and performance in a psychology examination. British Journal of Educational Psychology, 41, 90-4.

KLINGER, E. (1969) Feedback effects and social facilitation of vigilance performance: mere coaction versus potential evaluation. Psychonomic Science, 14, 161-2.

KOHFIELD, D.L. & WEITZEL, W. (1969) Some relations between personality factors and social facilitation. Journal of Experimental Research in Personality, 3, 287-92.

KOHN, H. (1954) Effects of variations of intensity of experimentally induced stress situations upon certain aspects of perception and performance. Journal of Genetic Psychology, 85, 289-304.

KOZAR, B. (1973) The effects of a supportive and nonsupportive audience upon learning a gross motor skill. International Journal of Sport Psychology, 4, 27-38.

LACEY, J.I. (1967) Somatic response patterning and stress: some revisions of activation theory, 14-37. In Appley, M.H. & Trumbull, R. (Eds.), Psychological Stress, N.Y. Appleton-Century-Crofts.

LACEY, J.I., KAGAN, J., LACEY, B.C. & MOSS, H.A. (1963) The Visceral level: situational determinants and behavioural correlates of autonomic response patterns. In Knapp, P.H. (Ed.) Expression of the Emotions in Man, New York: International Universities Press.

LAMB, D.H. (1976) Usefulness of situation-specific trait and state measures of anxiety. Psychological Reports, 38, 188-91.

LANA, R.E. (1969) Pretest sensitization, 119-51. In Rosenthal, R. & Rosnow, R.L. (Eds.) Artifact in Behavioral Research, New York: Academic.

LATANÉ, B. & DARLEY, J. (1968) Group inhibition of bystander intervention in emergencies. Journal of Personality and Social Psychology, 10, 215-21.

LATHAM, G.P. & BALDES, J.J. (1975) The 'practical significance' of Locke's theory of goal setting. Journal of Applied Psychology, 60, 122-4.

LAUGHLIN, P.R. & JACCARD, J.J. (1975) Social facilitation and observational learning of individual and cooperative pairs. Journal of Personality and Social Psychology, 32, 873-79.

LAUGHLIN, P.R. & WONG-McCARTHY, W.J. (1975) Social inhibition as a function of observation and recording of performance. Journal of Experimental Social Psychology, 11, 560-72.

LAZARUS, R.S. (1966) Psychological Stress and the Coping Process, McGraw Hill, New York.

LAZARUS, R.S. (1967) Cognitive and personality factors underlying threat and coping, 151-180. In Appley, M.H. & Trumbull, R. (Eds.), Psychological Stress, New York: Appleton-Century-Crofts.

LAZARUS, R.S., BAKER, R.W., BROVERMAN, D.M. & MAYER, J. (1957) Personality and psychological stress. Journal of Personality, 25, 559-77.

LAZARUS, R.S., DEESE, J. & OSLER, S. (1952) The effects of psychological stress upon performance. Psychological Bulletin, 49, 293-317.

LEAMON, T.B. (1974) An investigation into the effects of knowledge of results on operator performance. Ergonomics, 17, 639-51.

LEIDERMAN, P.H. & SHAPIRO, D. (1964) Studies on the galvanic skin potential level: some behavioural correlates. Journal of Psychosomatic Research, 7, 277-81.

LEKARCZYK, D.T. & HILL, K.T. (1969) Self-esteem, test anxiety, stress and verbal learning. Developmental Psychology, 1, 147-54.

LEUBA, C.J. (1933) An experimental study of rivalry in young children. Journal of Comparative Psychology, 16, 367-78.

LEVIN, H. & BALDWIN, A.L. (1958) The choice to exhibit. Child Development, 29. 373-380.

LEVIN, H., BALDWIN, A.L. Gallowey, M. & PAIVIO, A. (1960) Audience stress, personality and speech. Journal of Abnormal and Social Psychology, 61, 469-473.

LEWIS, W.D. (1943) Some characteristics of very superior children. Journal of Genetic Psychology, 62, 301-9.

LIEBLING, B.A. & SHAVER, P.(a)(1973) Evaluation, self-awareness and task performance. Journal of Experimental Social Psychology, 9, 297-306.

LIEBLING, B.A. & SHAVER, P.(b)(1973) Social facilitation and social comparison processes. Proceedings of 81st Annual Convention of American Psychol. Association, 8, 327-8.

LOCKE, E.A.(a)(1966) The relationship of intentions to level of performance. Journal of Applied Psychology, 50, 60-6.

LOCKE, E.A.(b)(1966) A closer look at level of aspiration as a training procedure. Journal of Applied Psychology, 50, 417-20.

LOCKE, E.A. & BRYAN, J.F.(a)(1966) Cognitive aspects of psychomotor performance: the effects of performance goals on level of performance. Journal of Applied Psychology, 50, 286-91.

LOCKE, E.A. & BRYAN, J.F.(b)(1966) The effects of goal-setting, rule learning, and knowledge of score on performance. American Journal of Psychology, 79, 451-7.

LOCKE, E.A. & BRYAN, J.F. (1967) Performance goals as determinants of level of performance and boredom. Journal of Applied Psychology, 51, 120-130.

LOCKE, E.A. & BRYAN, J.F. (1968) Grade goals as determinants of academic achievement. Journal of General Psychology, 79, 217-28.

LOCKE, E.A., CARTLEDGE, N. & KOEPPEL, J. (1968) Motivational effects of knowledge of results: a goal setting phenomenon? Psychological Bulletin, 70, 474-85.

LOMBARDO, J.P. & CATALANO, J.F. (1975) The effect of failure and the nature of the audience on performance of a complex motor task. Journal of Motor Behavior, 17, 29-37.

LONG, K.K. & BESSEMER, D.W. (1971) An analytical investigation of instructions designed to elicit test anxiety. Psychological Reports, 29, 283-92.

LORGE, I., FOX, D., DAVITZ, J. & BRENNER, M. (1958) A survey of studies contrasting the quality of group performance and individual performance, 1920-1957. Psychological Bulletin, 55, 337-72.

LURIA, A.R. (1932) The Nature of Human Conflicts, New York: Liveright.

LYNN, R. (1971) Two personality characteristics related to academic achievement, 5-10. In Eysenck, H.J. (Ed.), Readings in Introversion-Extraversion, New York: Wiley.

LYNN, R. & GORDON, I.E. (1961) The relation of neuroticism and extraversion to intellectual and educational attainment. British Journal of Educational Psychology, 31, 194-203.

MACCOBY, E.E. & JACKLIN, C.N. (1975) The Psychology of Sex Differences, Stanford, Calif., Stanford University Press.

MAGNUSSON, D. & EKEHAMMAR, B. (1975) Perceptions of audience reactions to stressful situations. Journal of Personality and Social Psychology, 31, 1147-55.

MAKOSKY, V.P. (1976) Sex-role compatibility of task and of competitor and fear of success as variables affecting women's performance. Sex Roles, 2, 237-238.

MALMO, R.B. (1965) Finger sweat prints in the differentiation of low and high incentive. Psychophysiology, 1, 231-40.

MALMO, R.B., BOAG, T.J. & SMITH, A.A. (1957) Physiological study of personal interaction. Psychosomatic Medicine, 19, 105-119.

MANDLER, G. (1972) Comments on Dr. Sarason's paper, 405-8. In Spielberger, C.D. (Ed.), Anxiety: Current Trends in Theory and Research, Vol.II, Academic Press.

MANDLER, G. & SARASON, S.B. (1952) A study of anxiety and learning. Journal of Abnormal and Social Psychology, 47, 166-73.

MANN, L. (1969) Social Psychology, John Wiley & Sons.

MARTENS, R.(a)(1969) Effects of an audience on learning and performance of a complex motor skill. Journal of Personality and Social Psychology, 12, 252-60.

MARTENS, R.(b)(1969) Palmer sweating and the presence of an audience. Journal of Experimental Social Psychology, 5, 371-74.

MARTENS, R., GILL, D.L., & SCANLAN, T.K. (1976) Competitive trait anxiety, success-failure and sex as determinants of motor performance. Perceptual and Motor Skills, 43, 1199-1208.

MARTENS, R. & LANDERS, D.M. (1969) Coaction effects on a muscular endurance task. Research Quarterly, 40, 733-7.

MARTENS, R. & LANDERS, D.M. (1972) Evaluation potential as a determinant of coaction effects. Journal of Experimental Social Psychology, 8, 347-59.

MASON, J. (1968) A review of psychoendocrine research on the pituitary-adrenal cortical system. Psychosomatic Medicine, 30, 576-607.

MATLIN, M.W. & ZAJONC, R.B. (1968) Social facilitation of word associations. Journal of Personality and Social Psychology, 10, 455-60.

McCLELLAND, D.C., ATKINSON, J.W., CLARK, R.A. & LOWELL, E.L. (1953) The Achievement Motive, New York: Appleton-Century-Crofts.

McCLINTOCK, C.G. & NUTTIN, J.M.Jr. (1969) Development of competitive game behavior in children across two cultures. Journal of Experimental Social Psychology, 5, 203-18.

McCULLAGH, P.D. & LANDERS, D.M. (1976) Size of audience and social facilitation. Perceptual and Motor Skills, 42, 1067-71.

McGRATH, J.E. (1970a) A conceptual formulation for research on stress. In McGrath, J.E. (Ed.) Social and Psychological Factors in Stress, New York: Holt, Rinehart and Winston.

McGRATH, J.E. (1970b) Major substantive issues: time, setting and the coping process. In McGrath, J.E. (Ed.) Social and Psychological Factors in Stress, New York: Holt, Rinehart and Winston.

McGRATH, J.E. (1970c) Major methodological issues. In McGrath, J.E. (Ed.) Social and Psychological Factors in Stress, New York: Holt, Rinehart and Winston.

McGRATH, J.E. (1970d) Settings, measures and themes: an integrative review of some research on social psychological factors in stress. In McGrath, J.E. (Ed.) Social and Psychological Factors in Stress, New York: Holt, Rinehart and Winston.

McGUIRE, W.J. (1969) Suspiciousness of experimenter's intent, 13-57. In Rosenthal, R. & Rosnow, R.L. (Eds.), Artifact in Behavioral Research, New York: Academic Press.

McMANIS, D.L. (1965) Pursuit rotor performance of normal and retarded children in four verbal-incentive conditions. Child Development, 36, 667-83.

McMANIS, D.L. (1966) Marble-sorting rate of elementary school children as a function of verbal-incentive and performance-level pairings. Perceptual and Motor Skills, 23, 499-507.

McNICOL, D. (1972) A Primer of Signal Detection Theory, Australasian Publishing Co., Sydney.

MECHANIC, D. (1970) Some problems in developing a social psychology of adaptation to stress. In McGrath, J.E. (Ed.) Social and Psychological Factors in Stress, New York: Holt, Rinehart and Winston.

MEGLING, B.M. (1976) The effect of evaluation on dominance characteristics: an extension of social facilitation theory. Journal of Psychology, 2, 167-73.

MILGRAM, S. & HOLLANDER, P. (1964) The murder they heard. The Nation, 198, 602-604.

MISCHEL, W. (1968) Personality and Assessment, New York: Wiley.

MISCHEL, W. (1973) Toward a cognitive social learning reconceptualization of personality. Psychological Review, 80, 252-283.

MOLDAWSKY, S. & MOLDAWSKY, P.C. (1952) Digit span as an anxiety indication. Journal of Consulting Psychology, 65, 59-66.

MORGAN, S.W., & MAUSNER, B. Behavioural and fantasied indicators of avoidance of success in men and women. Journal of Personality, 41, 457-470.

MORGENSTERN, F.S., HODGSON, R.J. & LAW, L. (1974) Work efficiency and personality: a comparison of introverted and extraverted subjects exposed to conditions of distraction and distortion of stimulus in learning task. Ergonomics, 17, 211-21.

MORRIS, L.W., SMITH, L.R., ANDREWS, E.S. & MORRIS, N.C. (1975) The relationship of emotionality and worry components of anxiety to motor skills performance. Journal of Motor Behavior, 7, 121-31.

MORRIS, W., WORCHEL, S., BOIS, J., PEARSON, J., ROUNTREE, A., SAMAHA, G., WACHTLER, J. & WRIGHT, S. (1976) Collective coping with stress: group reactions to fear, anxiety and ambiguity. Journal of Personality and Social Psychology, 33, 674-80.

MUKERJI, N.F. (1940) Investigation of ability to work in groups and in isolation. British Journal of Psychology, 30, 352-6.

MUSANTE, G. & ANKER, J.M. (1972) Experimenter's presence: effect on subject's performance. Psychological Reports, 30, 903-4.

NEWCOMB, T.M., TURNER, R.H. & CONVERSE, P.E. (1965) Social Psychology. The study of human interaction. Holt, Rinehart and Winston, Inc., New York, Chicago, San Francisco, Toronto, London.

NICHOLSON, W.M. (1958) The influence of anxiety upon learning: interference or drive increment. Journal of Personality, 26, 303-19.

NOBLE, C.E., FUGHS, J.E. ROBEL, D.P. & CHAMBERS, R.W. (1958) Individual versus social performance on two perceptual-motor tasks. Perceptual and Motor Skills, 8, 131-4.

NORDHOLM, L.A. (1975) Effects of group size and stimulus ambiguity on conformity. Journal of Social Psychology, 97, 123-30.

NOTTELMANN, E.D. & HILL, K.T. (1977) Test anxiety and off-task behaviour in evaluative situations. Child Development, 48, 225-32.

OKUN, M.A. & Di VESTA, F.J. (1975) Cooperation and competition in coacting groups. Journal of Personality and Social Psychology, 3, 615-20.

ORNE, M.T. (1962) On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. American Psychologist, 17, 776-83.

PAGE, M.M. & SCHEIDT, R.J. (1971) The elusive weapons effect: demand awareness, evaluation apprehension and slightly sophisticated subjects. Journal of Personality and Social Psychology, 20, 304-18.

PAIVIO, A. (1963) Audience influence, social isolation and speech. Journal of Abnormal and Social Psychology, 67, 247-53.

PAIVIO, A. (1965) Personality and audience influence, 127-173. In Maher, B. (Ed.), Progress in Experimental Personality Research, Vol.2, New York: Academic Press.

PAIVIO, A. & LAMBERT, W. (1959) Measures and correlates of audience anxiety ('stage fright'). Journal of Personality, 27, 1-17.

PAULUS, P.B. & CORNELIUS, W.L. (1974) An analysis of gymnastic performance under conditions of practice and spectator observation. Research Quarterly, 45, 56-63.

PAULUS, P.B. & MURDOCK, P. (1971) Anticipated evaluation and audience presence in enhancement of dominant responses. Journal of Experimental Social Psychology, 7, 280-91.

PAULUS, P.B., SHANNON, J.C., WILSON, D.L. & BOONE, T.D. (1972) Effect of spectator presence on gymnastic performance in a field situation. Psychonomic Science, 29, 88-90.

PEDERSON, A.M. (1970) Effects of test anxiety and coacting groups on learning and performance. Perceptual and Motor Skills, 30, 55-62.

PEPITONE, A. (1967) Self, social environment and stress, 182-208. In Appley, M. & Trumbull, R. (Eds.), Psychological Stress: Issues and Research, New York: Appleton.

PESSIN, J. (1933) The comparative effects of social and mechanical stimulation on memorizing. American Journal of Psychology, 45, 263-70.

PESSIN, J. & HUSBAND, R. (1933) Effects of social stimulation on human maze learning. Journal of Abnormal and Social Psychology, 28, 148-54.

PETTIGREW, T.F. (1967) Social evaluation theory: convergences and applications. In Levine, D. (Ed), Nebraska Symposium on Motivation, Vol.15, Lincoln University of Nebraska Press.

POULTON, E.C. (1970) Environment and Human Efficiency, Springfield, Ill. Charles C. Thomas.

POULTON, E.C. & FREEMAN, P.R. (1966) Unwanted assymetrical transfer effects with balanced experimental design. Psychological Bulletin, 66, 1-8.

PUTZ, V.R. (1975) The effect of different modes of supervision on vigilance behaviour. British Journal of Psychology, 66, 157-61.

QUARTER, J. & MARCUS, A. (1971) Drive level and the audience effect: a test of Zajonc's theory. Journal of Social Psychology, 83, 99-105.

RAJECKI, D.W., KIDD, R.F. & IVINS, B. (1976) Social facilitation in chickens: a different level of analysis. Journal of Experimental Social Psychology, 12, 233-47.

RAJECKI, D.W., KIDD, R.F., WILDER, D.A. & JAEGER, J. (1975) Social factors in the facilitation of feeding in chickens: effects of imitation, arousal or disinhibition. Journal of Personality and Social Psychology, 32, 510-8.

RAPHELSON, A.C. (1957) The relationships among imaginative, direct verbal, and physiological measures of anxiety in an achievement situation. Journal of Abnormal and Social Psychology, 54, 13-8.

RASMUSSEN, E.W. (1939) Social facilitation. Acta Psychology, 4, 275-91.

REYNOLDS, B. & ADAMS, J.A. (1954) Psychomotor performance as a function of initial level of ability. American Journal of Psychology, 67, 268-77.

RIKI, R. (1974) Effects of experimenter expectancy set and experimenter sex upon grip strength and hand steadiness scores. Research Quarterly, 45, 416-23.

RITTLE, R.H. & BERNARD, N. (1977) Enhancement of response rate by mere physical presence of experimenter. Personality and Social Psychology Bulletin, 3, 127-31.

ROESSLER, R. (1972) Personality, psychophysiology and performance. Psychophysiology, 10, 315-27.

ROETHLISBERGER, F.J. & DICKSON, W.J. (1939) Management and the Worker, Cambridge, Mass.: Harvard University Press.

ROMMETVEIT, R. (1955) Social Norms and Roles: Explorations in the Psychology of Enduring Social Pressures, Minneapolis: University of Minnesota Press.

ROSENDRANTZ, A.L. & VAN DE RIET, V. (1974) The influence of prior contact between child subjects and adult experimenters on subsequent child performance. Journal of Genetic Psychology, 124, 79-91.

ROSENTHAL, R. (1969) Interpersonal expectation: effects of the experimenter's hypothesis. In Rosenthal, R. & Rosnow, R. (Eds.), Artifact in Behavioural Research, New York: Academic Press.

ROSENTHAL, R. & ROSNOW, R.L. (1969) The volunteer subject. In Rosenthal, R. & Rosnow, R. (Eds.), Artifact in Behavioural Research, New York: Academic Press.

ROSNOW, R.L. & ROSENTHAL, R. (1976) The volunteer subject revisited. Australian Journal of Psychology, 28, 97-109.

RUDDOW, E. & HAUTALUOMA, J. (1975) Competition with oneself versus others as a facilitator in the classroom. Journal of Social Psychology, 95, 281-3.

RUSSELL, W.A. (1952) Retention of verbal material as a function of motivating instructions and experimentally induced failure. Journal of Experimental Psychology, 43, 207-216.

RYAN, D.E. & LAKIE, W.L. (1965) Competitive and noncompetitive performance in relation to achievement motive and manifest anxiety. Journal of Personality and Social Psychology, 1, 342-5.

SACKS, E.L. (1952) Intelligence scores as a function of experimentally established social relationships between child and examiner. Journal of Abnormal and Social Psychology, 47, 354-8.

SALTZ, E. (1970) Manifest anxiety: have we misread the data? Psychological Review, 77, 568-73.

SANDERS, G.S. & BARRON, R.S. (1975) The motivating effects of distraction on task performance. Journal of Personality and Social Psychology, 32, 956-63.

SARASON, I.G. (1956) Effect of anxiety, motivational instructions, and failure on serial learning. Journal of Experimental Psychology, 51, 253-60.

SARASON, I.G. (1957) Effect of anxiety and two kinds of motivating instructions on verbal learning. Journal of Abnormal and Social Psychology, 54, 166-171.

SARASON, I.G.(a)(1959) Relationships of measures of anxiety and experimental instructions to word association test performance. Journal of Abnormal and Social Psychology, 59, 37-42.

SARASON, I.G. (1960) Empirical findings and theoretical problems in the use of anxiety scales. Psychological Bulletin, 57, 403-15.

SARASON, I.G. (1961) The effects of anxiety and threat on solution of a difficult task. Journal of Abnormal and Social Psychology, 62, 165-8.

SARASON, I.G. (1972) Experimental approaches to test anxiety: attention and the uses of information, 381-403. In Spielberger, C.D. (Ed.), Anxiety: Current Trends in Theory and Research, Vol.II. Academic Press.

SARASON, I.G. (1975) Test anxiety, attention and the general problem of anxiety, 165-187. In Spielberger, C.D. & Sarason, I.G. (Eds.) Stress and Anxiety, Washington D.C.: Hemisphere Publishing Corp.

SARASON, I.G. & GANZER, V.J. (1970) Effects of delay interval, test anxiety, and task instructions on serial learning. Personality, 1, 265-73.

SARASON, I.G., KESTENBAUM, J.M. & SMITH, D.H. (1972) Test anxiety and the effects of being interviewed. Journal of Personality, 40, 242-9.

SARASON, I.G. & PALOLA, E.G. (1960) The relationship of test and general anxiety, difficulty of task, and experimental instructions to performance. Journal of Experimental Psychology, 59, 185-91.

SARASON, J.G. & SMITH, R.E. (1971) Personality. Annual Review of Psychology, 22, 393-446.

SARASON, S.B. (b)(1959) Intellectual and personality correlates of test anxiety. Journal of Abnormal and Social Psychology, 59, 272-5.

SARASON, S.B., MANDLER, G. & CRAIGHILL, P.G. (1952) The effect of differential instructions on anxiety and learning. Journal of Abnormal and Social Psychology, 47, 561-5.

SCHACHTER, S. (1959) The Psychology of Affiliation, Stanford University Press.

SCHACHTER, S., ELLERTON, N. & GREGORY, D. (1951) An experimental study of cohesiveness and productivity. Human Relations, 4, 229-38.

SCHACHTER, S. & SINGER, J. (1962) Cognitive, social and physiological determinants of emotional state. Psychological Review, 69, 379-99.

SCHMIDT, R.A. (1972) The case against learning and forgetting scores. Journal of Motor Behavior, 4, 79-88.

SCHRAMM, W. & DANIELSON, W. (1958) Anticipated audiences as determinants of recall. Journal of Abnormal and Social Psychology, 56, 282-3.

SCHWARTZ, J.M. & SMITH, W.P. (1976) Social comparison and the inference of ability differences. Journal of Personality and Social Psychology, 34, 1268-76.

SCOTT, J.P. (1968) Social facilitation and allelomimetic behaviour, 55-72. In Simmel, E., Hoppe, R. & Milton, G. (Eds.), Social Facilitation and Imitative Behaviour, Boston: Allyn & Bacon.

SCOTT, W.E.,Jr. & CHERRINGTON, D.J. (1974) Effects of competition, cooperation and individualistic reinforcement contingencies. Journal of Personality and Social Psychology, 3, 748-58.

SECORD, P.F. & BACKMAN, C.W. (1974) Social Psychology, 2nd Ed. McGraw-Hill Inc.

SEIDMAN, D., BENSEN, S.B., MILLER, I. & MEELAND, T. (1957) Influence of a partner on tolerance for self-administered electric shock. Journal of Abnormal and Social Psychology, 54, 210-2.

SELYE, H. (1950) Stress, Montreal: Acta.

SELYE, H. (1976) The Stress of Life, New York: McGraw-Hill.

SENGUPTA, N.N. & SINKA, C.P.N. (1926) Mental work in isolation and in groups. Indian Journal of Psychology, 1, 106-9.

SETA, J., PAULUS, P.B. & RISNER, H.T. (1977) The effects of group composition and evaluation on task performance. Bulletin of the Psychonomic Society, 9, 115-7.

SETA, J., PAULUS, P.B. & SCHKADE, J.K. (1976) Effects of group size and proximity under cooperative and competitive conditions. Journal of Personality and Social Psychology, 34, 47-53.

SHAVER, P. & LIEBLING, B.A. (1976) Explorations in the drive theory of social facilitation. Journal of Social Psychology, 99, 259-71.

SHERIF, M. (1936) The Psychology of Social Norms, New York: Harper.

SHERIF, M. & SHERIF, C. (1953) Groups in Harmony and Tension, New York: Harper & Row.

SHRAUGER, J.S. (1972) Self-esteem and reactions to being observed by others. Journal of Personality and Social Psychology, 23, 192-200.

SIMONS, L.S. & TURNER, C.W. (1976) Evaluation apprehension: hypothesis awareness and the weapons effect. Aggressive Behaviour, 2, 77-

SIMPSON, H.M. & MOLLOY, F.M. (1971) Effects of audience anxiety on pupil size. Psychophysiology, 8, 491-6.

SIMS, V.M. (1928) The relative influence of two types of motivation on improvement. Journal of Educational Psychology, 19, 480-4.

SINGER, J.E. (1966) Social comparison - progress and issues. Journal of Experimental Social Psychology, 2, 103-10.

SINGER, R.N. (1965) Effect of spectators on athletes and non-athletes performing a gross motor task. Research Quarterly, 36, 473-82.

SINGER, R. (1968) Motor Learning and Human Performance, Mich. St. Un: Collier-Macmillan.

SINGER, R.N. (1970) Effect of an audience on performance of a motor task. Journal of Motor Behaviour, 2, 88-95.

SINGER, R.N. & LLEWELLYN, J.H. (1973) Effects of experimenter's gender on subject performance. Research Quarterly, 44, 185-191.

SMITH, C.P. (1964) Relationships between achievement-related motives and intelligence, performance level and persistence. Journal of Abnormal and Social Psychology, 68, 523-33.

SMITH, L.E. & CRABBE, J. (1976) Experimenter role relative to social facilitation and motor learning. International Journal of Sport Psychology, 7, 158-69.

SMITH, R.E., VANDERBILT, K. & CALLEN, M.B. (1973) Social comparison and bystander intervention in emergencies. Journal of Applied Social Psychology, 3, 186-96.

SMODE, A.F. (1958) Learning and performance in a tracking task under two levels of achievement information feedback. Journal of Experimental Psychology, 56, 297-304.

SOROKIN, P.A., TANQUIST, M., PARTEN, M. & ZIMMERMAN, C.C. (1930) An experimental study of efficiency of work under various specified conditions. American Journal of Sociology, 35, 765-82.

SORRENTINO, R.M. & SHORT, J. (1974) Effects of fear of success on women's performance at masculine versus feminine tasks. Journal of Research in Personality, 8, 277-91.

SPEISMAN, J., LAZARUS, R., MORDKOFF, A. & DAVIDSON, L. (1964) Experimental reduction of stress based on ego-defense theory. Journal of Abnormal and Social Psychology, 68, 367-80.

SPENCE, K.W. (1956) Behavior Theory and Conditioning, New Haven, Conn. Yale University Press.

SPENCE, K.W. (1958) A theory of emotionally based drive (D) and its relation to performance in simple learning situations. American Psychologist, 13, 131-41.

SPENCE, K., FARBER, I. & McFANN, H. (1956) The relation of anxiety (drive) level to performance in competitional and non-competitional paired-associates learning. Journal of Experimental Psychology, 52.296-306.

SPENCE, J.T. & SPENCE, K.W. (1966) The motivational components of manifest anxiety: drive and drive stimuli, 291-326. In Spielberger, C.D. (Ed.), Anxiety and Behaviour, New York: Academic Press.

SPIELBERGER, C.D. (1966) Theory and research on anxiety, 3-19. In Spielberger, C.D. (Ed.), Anxiety and Behaviour, New York: Academic Press.

SPIELBERGER, C.D. (1975) State - trait - process, 115-143. In Spielberger, C.D. & Sarason, I.G. (Eds.), Stress and Anxiety, Vol.I, Washington D.C. Hemisphere Publishing Corporation.

STAKE, J.E. (1976) Effect of probability of forthcoming success on sex differences in goal setting: a test of the fear of success hypothesis. Journal of Consulting and Clinical Psychology, 44, 444-9.

STAMM, J.S. (1961) Social facilitation in monkeys. Psychological Reports, 8, 479-84.

TAYLOR, J. & SPENCE, K. (1952) The relationship of anxiety to level of performance in serial learning. Journal of Experimental Psychology, 44, 61-4.

TEEVAN, R.C. & SMITH, B.D. (1975) Relationships of fear of failure and need achievement motivation to a confirming-interval measure of aspiration levels. Psychological Reports, 36, 967-77.

THACKRAY, R.I., JONES, K.N. & TOUCHSTONE, R.M. (1974) Personality and physiological correlates of performance decrement on a monotonous task requiring sustained attention. British Journal of Psychology, 65, 351-9.

THACKRAY, R.I. & PEARSON, D.W. (1968) Effects of cognitive appraisal of stress on heart rate and task performance. Perceptual and Motor Skills, 27, 651-8.

THAYER, R.E. & MOORE, L.E. (1972) Reported activation and verbal learning as a function of group size (social facilitation) and anxiety-inducing instructions. Journal of Social Psychology, 88, 277-87.

THOMAS, J.R. & FORTSON, M.P. (1975) Experimenter effects on children's motor performance. Journal of Motor Behavior, 7, 65-73.

THOMAS, J.R. & HALLIWELL, W. (1976) Individual differences in motor skill acquisition. Journal of Motor Behavior, 8, 89-99.

THOMPSON, G. (1962) Child Psychology, Boston: Houghton Mifflin Co.

THORNTON, D. & ARROWOOD, A.J. (1966) Self-evaluation, self-enhancement and the locus of social comparison. Journal of Experimental Social Psychology, 2, 40-8.

TOLMAN, C.W. (1968) The role of the companion in the social facilitation of animal behaviour, 33-54. In Simmel, E.C., Hoppe, R.A. & Milton, G.A. (Eds.), Social Facilitation and Imitative Behaviour, Boston: Allyn & Bacon.

TORRANCE, E.P. (1962) Guiding Creative Talent, Englewood Cliffs, N.J. Prentice-Hall.

TRAVIS, L.E. (1925) The effect of a small audience upon eye-hand coordination. Journal of Abnormal and Social Psychology, 47, 222-229.

TRESEMER, D. (1976) The cumulative record of research on "fear of success". Sex Roles, 2, 217-36.

TRIPLETT, N. (1897) The dynamogenic factors in pacemaking and competition. American Journal of Psychology, IX, 507-32.

TYLER, L.E. (1956) The Psychology of Human Differences, New York: Appleton-Century-Crofts, 2nd Ed.

ULLRICH, C. (1957) Measurement of stress evidenced by college women in situations involving competition. Research Quarterly, 28, 160-72.

ULLRICH, C. & BURKE, R. (1957) Effect of motivational stress upon physical performance. Research Quarterly, 28, 403-12.

UNDERWOOD, B.J. (1975) Individual differences as a crucible in theory construction. American Psychologist, 30, 128-35.

UNDERWOOD, G.L. (1976) Social facilitation of individual learning and performance. International Journal of Sport Psychology, 7, 143-58.

VAUGHT, G.M. & NEWMAN, S.E. (1966) The effects of anxiety on motor-steadiness in competitive and non-competitive conditions. Psychonomic Science, 6, 519-20.

VUKELICH, R. & HAKE, D.F. (1974) Effects of the difference between self and coactor scores upon the audit responses that allow access to these scores. Journal of the Experimental Analysis of Behavior, 22, 61-73.

WACHTEL, P.L. (1968) Anxiety, attention and coping with threat. Journal of Abnormal Psychology, 73, 137-43.

WALKER, E.L. & TARTE, R.D. (1963) Memory storage as a function of arousal and time with homogenous and heterogenous lists. Journal of Verbal Learning and Verbal Behaviour, 2, 113-20.

WALTERS, R.H. (1968) Some conditions facilitating the occurrence of imitative behavior, 7-30. In Simmel, E., Hoppe, R. & Milton, G. (Eds.), Social Facilitation and Imitative Behaviour, Boston: Allyn & Bacon.

WANKEL, L.M. (1972) Competition in motor performance: an experimental analysis of motivational components. Journal of Experimental Social Psychology, 8, 427-37.

WANKEL, L.M. (1975) The effects of social reinforcement and audience presence upon the motor performance of boys with different levels of initial ability. Journal of Motor Behavior, 7, 207-17.

WAPNER, S. & ALPER, T. (1952) The effect of an audience on behavior in a choice situation. Journal of Abnormal and Social Psychology, 47, 222-9.

WAUGH, N. (1960) Serial position and the memory-span. American Journal of Psychology, 73, 68-79.

WEBER, S.J. & COOK, T.D. (1972) Subject effects in laboratory research: an examination of subject roles, demand characteristics, and valid inference. Psychological Bulletin, 77, 273-95.

WEBSTER, M. & SOBIESZEK, B. (1974) Sources of Self-Evaluation: A Formal Theory of Significant Others and Social Influence, New York: Wiley.

WEICK, K.E. (1970) The 'ess' in stress: some conceptual and methodological problems. In McGrath, J.E., Social and Psychological Factors in Stress, New York: Holt, Rinehart & Winston.

WEINER, B. (1970) New conceptions in the study of achievement motivation, 67-109. In Maher, B. (Ed.), Progress in Experimental Personality Research, Vol.5. New York: Academic Press.

WEISS, R. & MILLER, R. (1971) The drive theory of social facilitation. Psychological Review, 78, 44-57.

WEITZENKORN, S.D. (1974) An adjusted measure of achievement motivation for males and females and effects of future orientation on level of performance. Journal of Research in Personality, 8, 361-78.

WELFORD, A.T. (1976) Skilled Performance: Perceptual and Motor Skills, Scott, Foresman & Co.

WESTON, S.B. & ENGLISH, H.B. (1926) Influence of the group on psychological test scores. American Journal of Psychology, 37, 600-1.

WHEELER, L. (1966) Motivation as a determinant of upward comparison. Journal of Experimental Social Psychology, 2, 27-31.

WHEELER, L. (1968) Behavioural contagion: theory and research. In Simmel, E., Hoppe, R. & Milton, G. (Eds.), Social Facilitation and Imitative Behaviour, Boston: Allyn & Bacon.

WHITING, H.F. & ENGLISH, H.B. (1925) Fatigue - tests and incentives. Journal of Experimental Psychology, 8, 33-49.

WHITTEMORE, I.C. (1924) The influence of competition on performance: an experimental study. Journal of Abnormal and Social Psychology, 19, 236-53.

WHITTEMORE, I.C. (1925) The competitive consciousness. Journal of Abnormal and Social Psychology, 20, 17-23.

WHYTE, W.F. (1943) Street Corner Society, Chicago University Press.

WICKLUND, F.A. & DUVAL, S. (1971) Opinion change and performance facilitation as a result of objective self awareness. Journal of Experimental Social Psychology, 7, 319-42.

WILKINSON, R. (1969) Some factors influencing the effect of environmental stressors upon performance. Psychological Bulletin, 72, 260-72.

WILLIAMS, D.C.S. (1956) Effects of competition between groups in a training situation. Occupational Psychology, 30, 85-93.

WILMORE, H.H. (1968) Influence of motivation on physical work capacity and performance. Journal of Applied Psychology, 24, 459-63.

WINE, J. (1971) Test anxiety and direction of attention. Psychological Bulletin, 76, 92-104.

WINER, B.J. (1962) Statistical Principles in Experimental Design, McGraw-Hill Book Co. Inc.

WINKEL, G.H. & SARASON, I.G. (1964) Subject, experimenter and situational variables in research on anxiety. Journal of Abnormal and Social Psychology, 68, 601-8.

WOLFSON, S. (1977) Machiavellianism and helping behaviour. Paper presented at the British Psychological Society Annual Conference, Exeter.

WOLPE, J. (1958) Psychotherapy by Reciprocal Inhibition, Stanford University Press, Stanford, Calif.

WRIGHTSMAN, L.S. (1960) Effects of waiting with others on changes in felt level of anxiety. Journal of Abnormal and Social Psychology, 61, 216-22.

YERKES, R.M. & DODSON, J.D. (1908) The relation of strength of stimulus to rapidity of habit formation. Journal of Comparative Neurological Psychology, 18, 459-82.

ZAFFY, D.J. & BRUNING, J.E. (1966) Drive and the range of cue utilization. Journal of Experimental Psychology, 71, 382-4.

ZAJONC, R.B. (1965) Social facilitation, Science, 149, 269-74.

ZAJONC, R.B., HEINGARTNER, A. & HERMAN, E.M. (1969) Social enhancement and impairment of performance in the cockroach. Journal of Personality and Social Psychology, 13, 83-92.

ZAJONC, R.B. & SALES, S.M. (1966) Social facilitation of dominant and subordinate responses. Journal of Experimental Social Psychology, 2, 160-8.

ZAJONC, R.B., WOLOSIN, R.J., WOLOSIN, M.A. & LOL, W.D. (1970) Social facilitation and imitation in group risk-taking. Journal of Experimental Social Psychology, 6, 26-46.

ZENTALL, T.R. & HOGAN, D.E. (1976) Imitation and social facilitation in the pigion. Animal Learning and Behaviour, 4, 427-30.

ZIMMERMAN, C. & BAUER, R.A. (1956) The influence of an audience on what is remembered. Public Opinion Quarterly, 20, 238-48.

ZORBAUGH, H., BOARDMAN, R.K. & SHELDON, P. (1951) Some observations of highly gifted children, 86-105. In Witty, P. (Ed.), The Gifted Child, (The American Association for Gifted Children) Boston: Heath.

ZWEIDELSON, I. (1956) Test anxiety and intelligent test performance. Journal of Consulting Psychology, 20, 479-81.

APPENDICES

APPENDIX I

PEARSON CORRELATIONS FOR PERSONALITY TEST SCORES AND PERFORMANCE MEASURES

The following tables (A.1 to A.4) present all correlation coefficients calculated between the separate measures of Extraversion, Neuroticism, Fear of Failure and Audience Sensitivity and all performance measures in each experiment. Each table contains values obtained for one of these four personality measures. In experiments in which both Hits and Output were analysed, only the values for Output are presented because of the high positive relationship between these two performance measures.

The data within experiments are organized roughly along a continuum ranging from A and E conditions to Ex-Comp(F) as in Table 7.2, the rationale being the same as that discussed in Section 7.1.2. An attempt has been made to order similar conditions from each experiment in the columns of the tables to facilitate comparison between experiments and enable the identification of any trends under analogous conditions of testing. The analogies are not perfect, however, although the conditions are believed to be roughly equivalent in their stressfulness.

The tables are sub-divided according to the various performance measures used and organized in terms of task components. All tasks which provided measures of Output and Error are grouped under the same general sub-heading. However, data for RDS, Tracking and Aiming (darts) are given discrete sub-headings since these measures are not directly comparable to any other performance component. In some cases the correlations are based on personality test scores and raw score performance data (r.s.), while in other experiments the values are based on comparisons between the personality measures and difference scores (d.s.).

TABLE A.1

Pearson Correlations between Performance Measures and Extraversion

MEASURES OF SPEED (OUTPUT)

	CONDITIONS					EXPERIMENT NO.
	A				Ex-Comp	
Cancellation (r.s.)	-.36				-.21	I (N = 8)
	A	A-T	AUD	Im-Comp	Ex-Comp	
Cancellation (r.s.)	-.07	.14	-.34	-.01	.17	II (N = 20)
	E			Im-Comp	Ex-Comp	
Cancellation (d.s.)	.30			.17	-.01	IV (N = 30)
	E*			Im-Comp	Im-Comp (F)	
Transformation (d.s.)	.20			.11	.15	V (N = 20)
Logical Syllogism(d.s.)	.19			-.31	.17	*N = 19

MEASURES OF ERROR

	CONDITIONS					EXPERIMENT NO.
	A				Ex-Comp	
Cancellation (r.s.)	.10				.10	I (N = 8)
	A	A-T	AUD	Im-Comp	Ex-Comp	
Cancellation (r.s.)	-.09	-.21	.18	-.01	.03	II (N = 20)
	E			Im-Comp	Ex-Comp	
Cancellation (d.s.)	-.09			.00	-.11	IV (N = 30)
	E*			Im-Comp	Im-Comp (F)	
Transformation (d.s.)	-.05			-.05	-.27	V (N = 20)
Logical Syllogism (d.s.)	-.12			.05	-.25	*N = 19

MEMORY TESTS

	CONDITIONS					EXPERIMENT NO.
	A	A-T	AUD	Im-Comp	Ex-Comp	
RDS (r.s.)	.16	.06	.28	.06	.01	II (N = 20)
	E			Im-Comp	Ex-Comp	
Correct Digits (d.s.)	.13			-.09	-.31	IV (N = 30)

TRACKING (TOT)

	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Comp (F)	EXPERIMENT NO.
Phase I	.47	-.18	-.04	.56	.02	.09	.38	.28	VI (N = 12)
Phase II	-.25	-.36	.31	-.32	.00	-.31	.11	-.02	
			N=10	N=11	N=10				

TABLE A.2

Pearson Correlations between Performance Measures and Neuroticism

	CONDITIONS					EXPERIMENT NO.
MEASURES OF SPEED (OUTPUT)						
	A				Ex-Comp	
Cancellation (r.s.)	.33				-.15	I (N = 8)
	A	A-T	AUD	Im-Comp	Ex-Comp	
Cancellation (r.s.)	.39	.04	.11	.20	-.05	II (N = 20)
	E			Im-Comp	Ex-Comp	
Cancellation (d.s.)	$.58^{++}$			-.35	-.15	IV (N = 30)
	E^{*}			Im-Comp	Im-Comp(F)	
Transformation (d.s.)	-.09			-.28	-.12	V (N = 20)
Logical Syllogism(d.s.)	.18			-.03	.00	*N = 19
MEASURES OF ERROR						
	A				Ex-Comp	
Cancellation (r.s.)	.52				.77	I (N = 8)
	A	A-T	AUD	Im-Comp	Ex-Comp	
Cancellation (r.s.)	.32	-.21	-.34	-.13	.37	II (N = 20)
	E			Im-Comp	Ex-Comp	
Cancellation (d.s.)	$-.46^{+}$			.14	.21	IV (N = 30)
	E^{*}			Im-Comp	Im-Comp(F)	
Transformation (d.s.)	-.17			.10	$-.45^{+}$	V (N = 20)
Logical Syllogism(d.s.)	.15			$-.44^{+}$	-.25	*N = 19
MEMORY TESTS						
	A	A-T	AUD	Im-Comp	Ex-Comp	
RDS (r.s.)	.03	.25	.03	-.27	.21	II (N = 20)
	E			Im-Comp	Ex-Comp	
Correct Digits (d.s.)	-.13			.04	-.31	IV (N = 30)

TRACKING (TOT)

	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Comp (F)	
Phase I	-.46	.11	.15	-.05	-.19	-.02	.16	.27	VI (N = 12)
Phase II	.07	-.22	.17	-.13	.35	.58	-.02	- .09	
			N=10	N=11	N=10				

$+$ $p < .05$

$++$ $p < .01$

TABLE A.3

Pearson Correlations between Performance Measures and Fear of Failure

			CONDITIONS			EXPERIMENT NO.
MEASURES OF SPEED (OUTPUT)						
	A				Ex-Comp	
Cancellation (r.s.)	.42				-.20	I (N = 8)
	A	A-T	AUD	Im-Comp	Ex-Comp	
Cancellation (r.s.)	.24	-.09	.42	-.08	.12	II (N = 20)
	E-T		E-T(F)			
Cancellation (r.s.)	-.31		.01			III (N = 12)
				Ex-Comp	Ex-Comp(F)	
Cancellation (d.s.)				.10	-.06	III (N = 12)
	E			Im-Comp	Ex-Comp	
Cancellation (d.s.)	.35			$-.40^{+}$	.02	IV (N = 30)
	E^{*}			Im-Comp	Im-Comp(F)	
Transformation (d.s.)	.10			.10	.15	V (N = 20)
Logical Syllogisms (d.s.)	-.09			-.22	-.18	*N = 19
MEASURES OF ERROR						
	A				Ex-Comp	
Cancellation (r.s.)	.60				-.16	I (N = 8)
	A	A-T	AUD	Im-Comp	Ex-Comp	
Cancellation (r.s.)	.07	-.14	-.07	-.37	.37	II (N = 20)
	E-T		E-T(F)			
Cancellation (r.s.)	-.10		.41			III (N = 12)
				Ex-Comp	Ex-Comp(F)	
Cancellation (d.s.)				-.25	.14	III (N = 12)
	E			Im-Comp	Ex-Comp	
Cancellation (d.s.)	$-.40^{+}$			.16	.11	IV (N = 30)
	E^{*}			Im-Comp	Im-Comp(F)	
Transformation (d.s.)	-.02			.30	$-.71^{++}$	V (N = 20)
Logical Syllogism (d.s.)	.36			-.15	-.29	*N = 19

Continued/

TABLE A.3 (CONTINUED)

	CONDITIONS					EXPERIMENT NO.
MEMORY TASKS						
	A	A-T	AUD	Im-Comp	Ex-Comp	
RDS (r.s.)	-.26	.11	-.23	.00	.27	II (N = 20)
	E-T		E-T(F)			
RDS (r.s.)	.45		-.14			III (N = 12)
				Ex-Comp	Ex-Comp(F)	
RDS (r.s.)				-.49	.16	III (N = 12)
	E			Im-Comp	Ex-Comp	
Correct Digits (d.s.)	-.31			.00	-.17	IV (N = 30)

TRACKING (TOT)

	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Comp (F)	
Phase I	-.15	.38	$.67^{+}$	-.48	-.37	-.12	-.32	-.26	VI (N = 12)
Phase II	.10	.13	-.24 N=11	-.15	.05	.44	-.39	-.30	

AIMING (DARTS)

	Video 1	Video 2	Post-Vid 1	Post-Vid 2	Vid/Comp	
M (ratio)	$-.76^{++}$	.42	$-.56^{+}$	.32	.10	
Var(y) (ratio)	$-.54^{+}$	.05	$-.73^{++}$	.22	.40	VII (N = 20)
Var(x) (ratio)	-.43	.40	-.02	.25	.21	

$^{+}$ $p < .05$

$^{++}$ $p < .01$

TABLE A.4

Pearson Correlations between Performance Measures and Audience Sensitivity

			CONDITIONS		EXPERIMENT NO.
MEASURES OF SPEED					
	E-T	E-T(F)			
Cancellation (r.s.)	.15	-.27			III (N = 12)
			Ex-Comp	Ex-Comp (F)	
Cancellation (d.s.)			.08	.04	III (N = 12)
	E		Im-Comp	Ex-Comp	
Cancellation (d.s.)	-.23		-.13	.24	IV (N = 30)
	E*		Im-Comp	Im-Comp (F)	
Transformation (d.s.)	-.11		-.03	-.49+	V (N = 20)
Logical Syllogism (d.s.)	.00		.10	-.44+	*N = 19
MEASURES OF ERROR					
	E-T	E-T(F)			
Cancellation (r.s.)	-.57+	.13			III (N = 12)
			Ex-Comp	Ex-Comp (F)	
Cancellation (d.s.)			-.40	-.30	III (N = 12)
	E		Im-Comp	Ex-Comp	
Cancellation (d.s.)	-.03		-.05	.04	IV (N = 30)
	E*		Im-Comp	Im-Comp (F)	
Transformation (d.s.)	-.01		.03	-.01	V (N = 20)
Logical Syllogism (d.s.)	.11		-.11	.46+	*N = 19
MEMORY TASKS					
	E-T	E-T(F)			
RDS (r.s.)	.33	.15			III (N = 12)
			Ex-Comp	Ex-Comp (F)	
RDS (d.s.)			.42	.13	III (N = 12)
	E		Im-Comp	Ex-Comp	
Correct Digits (d.s.)	.01		.13	-.18	IV (N = 30)

Continued/

TABLE A.4 (CONTINUED)

	CONDITIONS								EXPERIMENT NO.
TRACKING (TOT)									
	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Comp (F)	
Phase I	-.29	.46	.37	-.51	.15	-.05	-.40	-.01	VI (N = 12)
Phase II	.26	.06	-.05 N=11	.35	-.05	.26	-.11	.25	

AIMING (DARTS)						
	Video 1	Video 2	Post Vid 1	Post Vid 2	Vid/Comp	
M	-.06	-.16	-.19	.09	.26	
Var(y)	.07	-.27	-.14	.02	.32	VII (N = 20)
Var(x)	.23	.11	.04	.24	.51++	

+ $p < .05$

++ $p < .02$

SUMMARY OF CORRELATIONAL DATA

Extraversion:

Generally the values obtained between Extraversion and performance measures were low, the highest single r value being -.36 (for experiments based on 20 or more subjects per cell). Somewhat higher values appear from the pursuit rotor data, but these are based on only 12 subjects per cell and are not significant.

Measures of Output yielded little in the way of single significant correlations or consistent patterns of correlations between experiments. The values for r on measures of Error are near zero in most cases. However, in Ex-Comp conditions there is some tendency for E to be positively related to the Errors measures when raw scores are used and negatively related to improvement in accuracy (r = -.11 to -.27). There is no consistent pattern between RDS or Correct Digits and Extraversion in any condition apart from Tracking. In this latter instance, some individual values are of high magnitude (.56), but no pattern is apparent between the various conditions within the experiment.

Neuroticism:

Values tend to be rather higher for this measure than for Extraversion. On Output measures there is a positive relationship between N and speed of performance in alone and experimenter only conditions (r = -.09 to .58), while the relationship between these variables in group situations (both Im-Comp and Ex-Comp) tends to be negative (r = .20 to -.35) with only one of the eight values being positive. A similar pattern is apparent for Errors measures. There is a tendency for N to be positively related to error scores and to be negatively related to improvement in accuracy in alone and experimenter only conditions (4 out of 5 cases), and this same pattern is suggested in Ex-Comp conditions (again 4 out of 5 cases), but the correlations are somewhat higher in magnitude. In general then, there is a tendency for N to be related to greater speed and lower accuracy in non-social conditions, while in groups this same factor is related to less speed and even greater inaccuracy.

There is no particular pattern between N and measures of RDS, and, as was true of Extraversion, although there are some individual r values which are somewhat high for Tracking, these follow no logical progression over the experimental conditions.

Fear of Failure:

Fear of Failure showed some tendency to be positively related to Output in alone and experimenter only conditions (4 out of 6 cases), values generally being higher for positive r's. However, there is no discernible pattern in any of the social conditions for this factor. For measures of Error there are some individual r's of significant magnitude, but there is no convincing pattern to the correlations between experiments. Similarly, measures of RDS yield no interesting patterns. There is some tendency for FF to be negatively related to Tracking performance (both Phases I and II) in group conditions (6 out of 8 comparisons), although the pattern of correlations in alone conditions appears random.

Interestingly, high negative values were obtained between FF and improvement in Darts performance under the Video condition on Day 1 ($r = -.43$ to $-.76$), and this pattern carried over to the Post-Video 1 condition ($r = -.02$ to $-.73$). Under these same testing conditions on Day 2, however, FF is generally related to performance decrement for both Video ($r = .05$ to $.42$) and Post Video 2 ($r = .22$ to $.32$). Similarly, under competition (Day 3) FF is associated with performance decrement, most notably for Var(y), $r = .40$. The reversal of the pattern between

Days 1 and 2 may be due to subjects anticipating the stress on subsequent testing days; those who were initially favourably influenced by the stress, when suddenly exposed to it, gave a negative response when given advance warning.

Audience Sensitivity:

For measures of Output r values are of insufficient magnitude to allow any speculations about the effects of this personality variable. Only in the Im-Comp(F) conditions of Experiment V are correlations of a substantial magnitude, negative in this case. No particular pattern emerged between AS and Errors measures. However, for RDS there is a positive association between AS and RDS on Day 1 and improvement in RDS on Day 2 in both the oral and written conditions of Experiment III, magnitudes being higher in the written conditions. Tracking performance does not appear to be related to AS in any consistent way. However, there are sizeable correlations between AS and darts performance under competition. In this case AS, like FF, is associated with decrement in performance, particularly for Var(x).

* * *

The above discussion of results only points out some possible trends in the correlational data. Neuroticism (and presumably Fear of Failure because of its high positive relationship to Neuroticism) would seem to merit further research with tasks for which speed and accuracy can be measured. N gives the strongest indication of being sensitive to the experimental situation, being associated with increased Output and inaccuracy in low stress situations and more general impairment in high stress conditions. Similarly, Fear of Failure yielded resonably high relationships with darts performance, suggesting a facilitative effect when stress is suddenly introduced but impairment when forewarning is available.

The values for Extraversion and Audience Sensitivity (negatively related to each other) were low in magnitude and sporadic both between and within experiments. There is little information available from this correlational analysis indicating that these two personality variables are of much importance in the social situations employed in this research.

PERSONALITY DATA: MEANS AND STANDARD DEVIATIONS

Experiment No.	Personality Measure	Conditions				
		A	Ex-Comp			
I (N = 8 per cell)	E	13.4	11.1			
	(S.D.)	(2.8)	(3.9)			
	N	12.1	11.0			
	(S.D.)	(4.0)	(4.1)			
	Argyle-Robinson FF	14.3	13.4			
	(S.D.)	(2.2)	(4.6)			
		A	A-T	Aud	Im-Comp	Ex-Comp
II (N = 20 per cell)	E	13.6	11.8	12.1	11.9	12.7
	(S.D.)	(4.4)	(2.7)	(4.3)	(3.9)	(4.6)
	N	9.1	11.1	11.8	9.5	11.5
	(S.D.)	(4.4)	(3.8)	(4.4)	(6.1)	(4.9)
	Argyle-Robinson FF	11.5	12.8	12.5	12.5	13.1
	(S.D.)	(3.3)	(3.7)	(2.0)	(3.3)	(3.5)
		E-T	E-T(F)			
III (N = 12 per cell)	TAS	16.5	15.1			
	(S.D.)	(5.5)	(4.8)			
	ASI	14.5	10.9			
	(S.D.)	(4.1)	(4.8)			
		E	Im-Comp	Ex-Comp		
IV (N = 30 per cell)	E	15.5	15.4	15.0		
	(S.D.)	(3.3)	(3.3)	(3.6)		
	N	13.2	13.7	12.9		
	(S.D.)	(3.3)	(4.9)	(4.1)		
	TAS	16.5	19.2	18.3		
	(S.D.)	(5.9)	(7.9)	(6.9)		
	ASI	13.6	15.2	12.5		
	(S.D.)	(6.7)	(5.5)	(5.5)		

Continued/

Experiment No.	Personality Measure	Conditions		
		E	Im-Comp	Im-Comp(F)
V (N = 20 per cell)* *For condition E, N = 19	E (S.D.)	13.1 (4.7)	15.5 (3.6)	14.6 (2.7)
	N (S.D.)	13.2 (4.1)	14.3 (4.2)	12.8 (4.6)
	TAS (S.D.)	18.2 (7.4)	18.8 (6.7)	14.8 (7.7)
	ASI (S.D.)	17.6 (6.0)	15.1 (5.5)	14.3 (4.7)

Experiment No.	Personality Measure	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Com (F)
VI (N = 12 per cell) *N = 10 **N = 11	E (S.D.)	11.2 (4.3)	12.3 (6.4)	11.4* (5.5)	12.2** (6.3)	12.1* (3.3)	12.3 (4.8)	12.0 (3.2)	10.5 (4.9)
	N (S.D.)	10.4 (5.9)	11.0 (5.3)	10.7* (5.4)	10.6** (4.7)	8.6* (4.9)	9.1 (2.8)	9.5 (5.8)	9.7 (4.1)
	TAS (S.D.)	13.4 (7.5)	13.2 (5.9)	13.6** (5.4)	13.4 (5.7)	16.9* (6.6)	14.3 (4.2)	15.9 (4.9)	16.4 (7.7)
	ASI (S.D.)	17.7 (6.4)	11.7 (6.4)	12.0** (6.6)	13.2 (6.2)	12.8* (6.9)	12.5 (6.3)	14.0 (6.2)	13.7 (5.7)

Experiment No.	Personality Measure	Same subjects in all conditions
VII (N = 20)	TAS (S.D.)	15.9 (5.2)
	ASI (S.D.)	11.7 (5.7)

APPENDIX II

TASK AND EXPERIMENTAL INSTRUCTIONS

EXPERIMENT I

Task Instructions (for alone and group conditions):

You (each of you) will be given several pages of printed material. You are to go through this material and mark out with a single diagonal line all of the 'e's in each line of printing. You should try to complete as many lines as possible while being careful not to miss any of the 'e's. Before the actual experiment begins I would like you to practice for a few minutes so that I will be sure you understand what you are to do. You should begin on the first page of printed material before you for the practice trial. You may begin when I say begin and continue working until I say stop.

Experimental Instructions:

Alone

I would now like you to begin the actual experiment. During the experiment you will be given 20 minutes in which to work. You should continue working as you did during the practice period. When you have finished with a page, put it underneath the other pages in your stack. After the experiment begins I will leave the room. A buzzer will sound several times during the experiment from outside the door. Whenever you hear the buzzer place an X in the right hand margin of the page to indicate the last line you have completed at that point. Your scores on this part of the experiment will be compared with the scores on two personality tests which you will be given at a later date. As you were told earlier, you should try to work as quickly and as carefully as possible. You may begin the experiment when I say begin and continue working until I say stop.

Explicit Competition

These instructions were identical with those above apart from the underlined section which was substituted with:

I will be scoring the work you complete to see which one of you does best, so it is important that you work as quickly and as carefully as possible.

EXPERIMENT II

Task Instructions (all conditions):

You (each of you) will be given several pages of printed material. You are to go through this material and cross out with a single diagonal line all of the 'e's in each line of printing. You should try to work as quickly and carefully as you can. While you are working a tape recorder will be playing lists of digits. After the completion of each list you are to write the last eight digits you have heard in the right hand margin of the page. Write the digits at the end of the line on which you are working. The numbers will be presented at a rate slightly less than 1 per second, and the lists are of different lengths so you will not be able to anticipate the end of the list. All of the last eight digits are different and you may guess if you like when recalling them. Write the digits you can remember as quickly as possible and return immediately to your 'e' crossing. I will be analysing both your 'e' crossing and your digit recall scores. However, you should bear in mind that 'e' crossing is the primary task and will be weighted much more heavily in your total score. At this time I would like you to practice for a few minutes so I will be sure you understand the task. You may begin when I say begin and continue working until I say stop.

Experimental Instructions:

Conditions A, AUD and Im-Comp

I would now like you to begin the actual experiment. During the experiment you will be given approximately 20 minutes in which to work. As stated in the previous instructions, you should work as quickly and carefully as possible on your 'e' crossing while attending to and recalling the digits. There is evidence indicating a correlation between personality traits and response patterns on this task. After the completion of the task you will be given two personality tests whose scores I will be comparing with your task scores. I will be leaving the room after you begin the experiment. You may continue working until you hear the phrase "end of test" on the tape.

Conditions A-T and Ex-Comp

These instructions were identical with those above apart from the underlined section which was substituted with:

> After the completion of the task you will be given two personality tests. There is evidence indicating a correlation between personality traits and success on this task. I will be scoring your work in conjunction with your personality tests to see how well you compare with the other students participating in the experiment.

EXPERIMENT III

Preliminary Explanation (all subjects):

You have been asked to come here today to work on a rather unusual task. This task requires you to do two things at once. One part of the task is a digit span exercise which is a modification of the Digit Span subtest of the Wechsler-Bellevue Intelligence Scale. The other part is a cancellation task, which essentially involves proofreading and is correlated with reading and vocabulary skills. I am interested in combining an auditory and a visual task in order to determine how performance on one task interacts with and influences performance on the other.

Task Instructions for Pretest (E-T and E-T(F)):

You will be given several sheets of printed material. You are to go through this material and cross out all the 'e's with a single diagonal line. At the same time a tape with lists of digits will be playing. Whenever the list stops you are to stop 'e' crossing and write (say out) the last eight digits you have heard. The lists are of different lengths so you will not be able to anticipate the end of the list. You will have 15 seconds in which to recall your answers. The numbers are presented at a rate of slightly less than one per second and all of the last eight are different. You will probably not always be able to remember all eight digits; in these cases write down (report) as many as you do remember - do not random guess, put down (report) only those digits you are reasonably sure about. As soon as you have recorded the digits you can return to your 'e' crossing; you do not have to wait until the numbers begin again. Do not mull over numbers you cannot remember, as this type of memory fades very quickly and what you don't remember almost immediately you probably won't be able to recall and you will sacrifice time from your 'e' crossing by trying to do so. Neither part of the test is more important than the other; they both count equally, so divide your efforts between the two as best you can.

Day II (Ex-Comp and Ex-Comp(F)):

As I mentioned before, I am interested in variables affecting performance on complex mental tasks. You may remember that the two tasks you have done previously were a digit span test, which is a sub-test of the Wechsler-Bellevue Intelligence Scale, and a cancellation task, which is related to reading and vocabulary skills. There is evidence that performance tasks of this nature are sometimes affected by subtle features of the environment, such as the number of persons present during the testing session. This is a fairly important variable to investigate, since performance rarely occurs in absolute isolation. For this reason you have been asked to come today in a group. Your task will be the same as in the previous session. Today, however, I will ask you to work at the task for 10 minutes each in turns. When you are not actually engaged in doing the task, you are asked to sit quietly and just observe.

EXPERIMENT IV

Task Instructions for Pretest (all subjects):

This is a task which requires you to pay attention to two things at once. The first and most important thing you must do is to cross out with a single line all of the 'e's you see as you scan the pages in front of you. You should do this as quickly as you can but try not to miss any 'e's. While you are 'e' crossing you will also be hearing lists of numbers on the tape recorder. At the end of each list there will be a brief pause of about 5 seconds. During this pause you will be asked to recall the third number back from the end of the list. This number you are to write down on the page next to the line where you stopped 'e' crossing. A completed page would look like this (show sample page). You will have to write the number down quickly, as the next list will begin after only a very short interval. If you cannot remember the number just guess, but always write down something. Lists are of different lengths, so you will not be able to know for sure when the end will occur. None of the numbers are zero. Both the 'e' crossing and the number task are important, but the 'e' crossing is about twice as important as the number task, so you should pay more attention to 'e' crossing. It should be similar to reading a book with the radio on. Your actual performance on the task today is not going to be evaluated in any way. I am only interested in seeing how people do the task. Therefore do not concern yourself with whether you are doing well or not, just do your best, working quickly and carefully at all times.

To get you started you will have a brief practice which will include 10 number lists. In all this test will take about 15 minutes. You may begin when I say begin and continue working until I say stop.

Experimental Instructions:

Conditions E and Im-Comp

During the first testing session you were probably not aware that your performance was changing as you became more experienced with the task. In a task of this sort a lot of learning takes place over the first hour or so of practice, but these changes are difficult to analyse because people become tired and fatigue also affects performance. Therefore, you are being tested in two relatively short sessions so that I can see more clearly the different patterns of performance that emerge as people become more skilled. Once again, as in the first session, your performance is not being evaluated. I am merely interested in seeing how people do the task. Just try to do your best, working quickly and carefully at all times. Not all of your friends are being tested in exactly this same way. This is because I am looking at several different aspects of performance and your test is only part of the total experiment.

Condition Ex-Comp

This test, which you all have done before, is an important one because it is related to several different skills and abilities. 'E' crossing is highly dependent on reading and vocabulary skills, and the number task is a modified version of a widely used intelligence test. After you have finished the test you will be given several questionnaires which ask you questions about yourself, as I wish to see if certain types of people are more successful on the test than others. The first test you did was for practice, as some individuals do not learn the task as easily as others and are therefore disadvantaged. But, today's session is a test, and I will be looking at your papers to see which students are best. Therefore, it is important that you do your best, working quickly and carefully at all times. Not all of your friends are being tested in this same way. This is because I am looking at several different aspects of performance, and your test is only part of the total experiment.

Subjects in this condition were asked not to reveal to yet untested subjects that the test was related to intelligence.

EXPERIMENT V

All subjects were informally told during the recruitment meeting that the experiment was concerned with personality and performance.

Task Instructions for Pretest (all subjects):

In a few minutes you will be asked to work briefly on two different tasks. I have studied the way in which university students do these tasks and I am now interested in seeing whether there are any differences in the ways in which somewhat younger people do them. The first task essentially involves counting. You will be given a letter of the alphabet plus a number, and your job will be to count forward in the alphabet as the number specifies and write down the letter answer. For example (write on blackboard):

A + 3 = D

Y + 2 = A

The numbers will always be between 2 and 5, and if you count beyond the end of the alphabet, simply start again at the beginning as in example two. You should try to work as quickly as possible without making any mistakes. Do not guess at answers - you should be reasonably sure you have the right answer before you write it down.

The second task involves reading a series of statements about a combination of letters, either AB or BA and deciding if the statement is true. For example (write on blackboard):

A follows B - AB, F

B is not preceded by A - BA, T

Once again accuracy is important, and you should not record an answer unless you are sure it is the right one.

You will be given 2 minutes to work on each task. Printed sheets with the questions will be provided for you and you are to write your answers in the spaces next to each question.

At this point test sheets were passed out and pupils were given an opportunity to ask questions.

You may begin when I say go and continue working until I say stop.

Experimental Instructions (all conditions, E, Im-Comp and Im-Comp(F)):

As you were told before, this is an experiment concerning the effects of personality differences on the performance of routine mental tasks. There is evidence that different types of people approach tasks differently. I am also interested in seeing whether these type tasks are best done in isolation or in groups, and therefore, some of the pupils will be tested alone and some in small groups. Today you will be asked to do a longer version of the two tests you did previously, but there will be some procedural differences. All the questions today will appear on cards with 4 questions to a card. You will do the tests by turning up one card at a time and writing your answers on a separate answer sheet. You are not to make any marks on the cards. When you have finished a card put it in a separate pile and go on to the next card. Your answer sheets are marked out with numbers that correspond to the numbers on each card. Next to each number are 4 lines, one for each problem on the card. (Show sample card and answer sheet.) Today you will spend 6 minutes working on each task, the alphabetic task first and the true-false statements second. Each task will have a minute rest period half-way through.

At this point the task instructions were reviewed.

You may go when I say go and continue working until I say stop. As mentioned, midway through each test you will have a minute rest period, and I will tell you when it occurs. You should sit quietly and not talk during the rest period.

EXPERIMENT VI

Task Instructions for Pretest (all subjects):

You have been asked here today to participate in an experiment involving the development of a motor skill. You will be working on an apparatus called a pursuit rotor. Your task will be to trace a moving spot of light with this stylus like so (demonstration). The light follows a specific course which is completed every six seconds. The pattern will remain the same throughout the experiment. You should only use your preferred hand to hold the stylus and you may not change hands during the course of the experiment. Every 20 seconds a number will appear in this window. This number will be the amount of time you were on target in the previous 20 seconds. Your goal is to remain on target as much of the time as possible. You should not try to watch the score indicator, however, as this will distract you and make you do worse.

Before we begin the experiment I will give you (each of you) one minute of practice with the task. This is so I will be sure you understand the task and also to give you some degree of familiarity with it before the experiment begins. It is necessary that I remain in the room during the experiment in order to record your score. Therefore, I will stay here while you are practicing to accustom you to my presence and activity nearby. This type of task requires a great deal of attention, so try not to be distracted by what I am doing. If you have questions concerning the scoring procedure or the apparatus I will answer them when the experiment is over. (Groups only - I will ask each of you to wait outside until it is your turn to practice. This is so no one will approach the task with the advantage of having watched someone else and therefore have more information available.) Remember, your goal is to keep the stylus on the light spot as much of the time as possible.

Experimental Instructions:

Conditions E and Im-Comp

I am interested in studying the early stages in the learning and development of motor skill and the conditions under which people learn such skills best. Subjects are being tested in several different conditions varying the speed of target rotation, pattern of target, and type of practice. How well you do individually is of no special importance since your score will be averaged in with the scores of the other subjects tested in your condition. Also, individuals have markedly different patterns and rates of learning this task, so if you think you are not performing very well, this does not reflect on your ultimate ability to master this or any other perceptual-motor task. It is important however, that you try to do your very best so that I will obtain an accurate estimate of your group's score. * *
You will be given a total of 3 minutes to work on the task. At the end of each minute you will be given a rest period of 20 seconds. * *
In all you will have 3 performance trials and 2 interpolated rests. 5 seconds before it is time for you to begin again after the rest, I will say "ready". Do not begin, however, until I say "go".

(Group only) - As only one of you will be able to do the task at a time I would like the rest of you to sit quietly and wait your turn. When you have finished will you please return to your seat and wait until the others have finished as there is a short questionnaire for you to complete before you go.

Conditions E(F) and Im-Comp(F)

Instructions were the same as those above apart from the two insertions indicated by the asterisks.

* During the experiment you will continue to work on the task as you did in the practice. The only difference will be that this time whenever you are off target, you will hear this noise (demonstrate). This will alert you when you are making an error.*

* During this rest period you may check and see what your score was for the previous minute as subjects usually like to know how they are doing.*

Conditions E-T and Ex-Comp

I am interested in studying the early stages in the learning and development of motor skill and how personality and individual differences contribute to successful and unsuccessful performance. Good performance on the pursuit rotor is dependent upon the basic ability to translate visual information and coordinate appropriate muscular responses to this information. The same basic ability is required for successful performance in such games as darts and especially games involving movement, such as tennis. It is also related to more practical skills such as driving a car. There is great individual variation in this ability, and today I will be interested in seeing how well you (each of you) as an individual are able to do the task. Later you will be given some questionnaires from which I will try to find relationships between what type of person you are and how successfully you learn and perform this task. Remember, I will not be averaging your score in with the scores of the other subjects but will be looking at your individual score to see how favourably it compares with the scores of the other subjects. Therefore, it is important that you try to do your very best. * * You will be given a total of 3 minutes to work on the task. At the end of each minute you will be given a rest period of 20 seconds. * *

The remainder of the instructions are the same as those given for conditions E and Im-Comp.

Conditions E-T(F) and Ex-Comp(F)

These instructions were identical to those given above for E-T and Ex-Comp apart from the two insertions regarding auditory feedback and knowledge of results noted by the asterisks and listed under conditions E(F) and Im-Comp(F).

EXPERIMENT VII

Preliminary Explanation:

You have been asked here to participate in an experiment investigating the role of personality factors and individual movement patterns in game playing. The experiment will take place in three sessions. Your task will be to throw darts at a dartboard and record your score on a scoring sheet. You will also be given several personality tests later in the experiment.
(Subjects choose darts and warm up.)

Instructions for Task (all subjects):

You will use the darts you have chosen today in the experiment and in the remainder of the experiment. You are to throw your darts from behind the farthest (nearest) line marked with tape on the floor. You may play in any manner you like and vary your style if you wish. Each time you have completed throwing the three darts you will be asked to record the position of the darts on a scoring sheet which consists of several small replicas of dart boards. In order to insure that scoring be done as accurately as possible you will need to take the score sheet up to the board and record the score before you remove any darts. This should be done in the following manner - (demonstration of scoring method).

The 'bull' will be considered the target in this experiment and the closer your darts are to the bull the higher your score will be. Your score will be computed as the average distance your darts land from the bull. Therefore, there will be no advantage in landing in segments associated with high numbers or in the doubles or trebles rings. If you throw a dart and it bounces off the board you may pick it up and throw it again. However, if your dart misses the board and lands in or bounces off the gray backboard you should not throw it again and record only the two in the board.

Instructions for Baseline 1

I would now like you to prepare yourself for the experiment by practising darts throwing and scoring. It is important that you try to throw and score accurately during the practice because your practice performance sometimes sets a trend for later performance. You are to throw 24 darts during the practice and record your score on a separate dart board replica after each 3 darts you throw. Remember to take your score sheet up to the board and record your score before pulling out any darts. You may take as long as you like in throwing, but you will be signalled when to begin by two flashes of this light. I will be leaving the room before you begin throwing your darts. When you have finished please sit down and wait for my return. Do not throw any additional darts.

Instructions for Video 1

As I mentioned before, part of the purpose of this experiment is to investigate the role of stylistic movement patterns in games playing. There is evidence indicating that individuals have different styles and methods of playing games. These styles are relatively stable characteristics which individuals use in any game or task with similar requirements, such as those involving eye-hand coordination. I am interested in studying the relationship between these playing styles and personality and in determining which individuals and which styles are the most successful. In a few minutes I will ask you to throw darts again and this time I will make a video recording of you. I have made arrangements with a group of students to analyse these recordings and classify the different styles and movement patterns that you and other subjects are using in playing darts. You should try to throw your darts and record your score as accurately as possible because your score will be compared and coordinated with the results of the tape analysis and the scores of the other participants in the experiment. You should continue throwing as you did during the practice session - you will again throw 24 darts. I will be staying in the room this time in order to operate the equipment and I will be taking certain recordings myself regarding your performance. You may take as long as you like in throwing, but you should not begin until I tell you to. Do not try to watch the television monitor.

Instructions for Post-Video 1

The main part of today's session is now over. However, I would like you to throw 24 more darts and record your scores. This is to see whether there has been any change in your performance as a result of practice today.

Instructions for Baseline 2, Video 2 and Post-Video 2

For this session subjects were told informally that they would be doing the same thing as in the previous session, in order to insure that sufficient data had been collected and to combat possible effects due to fatigue. The instructions for the task were briefly reviewed.

Instructions for Baseline 3

Subjects were informally told that they were to 'practice' again in preparation for the experiment.

Instructions for Video-Competition

As you know I am interested in studying individual movement patterns in games playing. I now have an adequate sample of your style of darts playing in an isolated situation. However, there is evidence that these styles vary somewhat in actual game conditions. Therefore, today I would like the two of you to play a rather simple dart game. The task will be the same as it has been when you were tested before. The only difference is that you will be keeping score for your partner rather than for yourself. Each time one of you throws 3 darts your partner will record your score. You will then become scorekeeper while your partner throws 3 darts. You are to alternate in this fashion until you have each thrown 24 darts. In addition to plotting the position of the darts, I would also like you to keep a numerical score. Numerical values for each concentric ring are as indicated on this sample. (Show sample.) When you have recorded the position of the darts you are also to calculate the numerical score and write it beside the respective dart board miniature. Once again I will be making a video recording of you, and therefore please call out the numerical score for each trial so it will be recorded on the tape. You may talk to each other as you like and play in any manner that you wish. However, it is important that you try to do your best as this test is the most similar to an actual performance situation and therefore will be counted more heavily.

APPENDIX III

MEANS, STANDARD DEVIATIONS AND ANALYSIS OF VARIANCE TABLES

This appendix contains the summarized data for all measures considered in the previous experiments. In some cases data which appear in the text are reproduced so that all the experiments can be surveyed without referring to previous sections. A standard format has been employed to facilitate reading. Each performance measure is presented separately (for each experiment) with the corresponding means, standard deviations and analysis of variance summary tables. Standard deviations refer only to the overall means for the various social conditions. In the analysis of variance summary tables, values have been rounded off to two decimal places. However, the actual analyses were conducted on data carried out to four significant figures.

EXPERIMENT I

Task : Cancellation; N = 16

Variables: Between - Sex(2) and Order(2), Within - Social Conditions(2)
and Periods(4)

PERFORMANCE MEASURE - OUTPUT

Means and Standard Deviations:

Order 1 (Alone → Group)

	Alone				Group			
Periods	1	2	3	4	1	2	3	4
Males	258.3	264.8	264.5	258.8	325.0	292.3	304.8	298.0
Females	243.0	232.0	238.5	252.3	311.3	292.5	280.3	287.8
All over mean	250.6	248.4	251.5	255.5	318.1	292.4	292.5	292.9
(S.D.)	(47.04)	(45.43)	(37.14)	(35.21)	(29.29)	(24.11)	(28.42)	(22.77)

Order 2 (Group → Alone)

	Alone				Group			
Periods	1	2	3	4	1	2	3	4
Males	275.5	281.8	297.0	303.5	294.0	269.5	271.0	280.3
Females	264.5	244.5	249.3	246.3	258.5	229.0	222.8	216.5
All over mean	270.0	263.1	273.1	274.9	276.3	249.3	246.9	248.4
(S.D.)	(50.28)	(49.50)	(51.09)	(55.50)	(49.57)	(42.04)	(53.12)	(52.79)

OUTPUT (CONTINUED)

Analysis of Variance Summary Table:

Order = 0, Sex = S, Social Conditions = SC, Periods = P

Source	DF	SS	MS	F	Probability
Subj.	15	186647.22			
0	1	5000.00	5000.00	<1	n.s.
S	1	27612.50	27612.50	2.23	n.s.
0 x S	1	5644.53	5644.53	<1	n.s.
Error	12	148390.19	12365.85		
SC	1	8385.13	8385.13	10.98	<.01
SC x 0	1	31312.53	31312.53	41.00	<.001
SC x S	1	0.78	0.78	<1	n.s.
SC x 0 x S	1	561.13	561.13	<1	n.s.
Error	12	9165.69	763.81		
P	3	4414.78	1471.60	5.15	.005
P x 0	3	37.31	12.44	<1	n.s.
P x S	3	1533.81	511.27	1.79	n.s.
P x 0 x S	3	2001.53	667.18	2.34	<.09
Error	36	10279.31	285.54		
SC x P	3	4979.69	1659.90	9.52	n.s.
SC x P x 0	3	94.78	31.59	<1	n.s.
SC x P x S	3	759.28	253.09	1.45	n.s.
SC x P x 0 x S	3	439.69	146.56	<1	n.s.
Error	36	6275.31	174.31		

PERFORMANCE MEASURE - OMISSIONS (PER LINE)

Means and Standard Deviations:

Order 1 (Alone → Group)

	Alone				Group			
Periods	1	2	3	4	1	2	3	4
Males	.35	.43	.32	.37	.22	.22	.22	.23
Females	.52	.63	.56	.57	.68	.59	.42	.43
All over mean	.43	.53	.44	.47	.45	.40	.32	.33
(S.D.)	(.30)	(.37)	(.24)	(.32)	(.33)	(.24)	(.22)	(.20)

Order 2 (Group → Alone)

	Alone				Group			
Periods	1	2	3	4	1	2	3	4
Males	.15	.19	.23	.12	.36	.31	.27	.29
Females	.14	.16	.17	.11	.34	.36	.23	.21
All over mean	.14	.17	.20	.12	.35	.34	.25	.25
(S.D.)	(.10)	(.14)	(.13)	(.11)	(.16)	(.16)	(.19)	(.24)

OMISSIONS (CONTINUED)

Analysis of Variance Summary Table:

Order = 0, Sex = S, Social Conditions = SC, Periods = P

Source	DF	SS	MS	F	Probability
Subj	15	5.47			
0	1	1.23	1.23	4.65	.05
S	1	0.43	0.43	1.61	n.s.
0 x S	1	0.63	0.63	2.39	n.s.
Error	12	3.18	0.27		
SC	1	0.02	0.02	<1	n.s.
SC x 0	1	0.42	0.42	10.07	<.01
SC x S	1	0.02	0.02	<1	n.s.
SC x 0 x S	1	0.21	0.21	<1	n.s.
Error	12	0.50	0.04		
P	3	0.11	0.04	2.75	<.06
P x 0	3	0.02	0.01	<1	n.s.
P x S	3	0.04	0.01	1.01	n.s.
P x 0 x S	3	0.01	0.00	<1	n.s.
Error	36	0.46	0.01		
SC x P	3	0.10	0.03	2.27	<.1
SC x P x 0	3	0.02	0.01	<1	n.s.
SC x P x S	3	0.05	0.02	1.14	n.s.
SC x P x 0 x S	3	0.03	0.01	<1	n.s.
Error	36	0.51	0.01		

PERFORMANCE MEASURE - HITS

Means and Standard Deviations:

Order 1 (Alone - Group)

	Alone				Group			
Periods	1	2	3	4	1	2	3	4
Males	243.5	245.3	250.0	242.8	313.3	282.3	293.8	291.5
Females	221.0	207.3	217.0	228.3	276.5	265.0	261.3	268.3
Overall mean	232.3	226.3	233.5	235.5	294.9	273.6	277.5	279.9
(S.D.)	(41.15)	(38.38)	(34.79)	(30.33)	(30.57)	(27.72)	(30.53)	(19.60)

Order 2 (Group - Alone)

	Alone				Group			
Periods	1	2	3	4	1	2	3	4
Males	269.5	274.0	286.5	297.5	277.0	256.5	258.5	267.5
Females	258.0	237.5	244.8	266.0	243.8	215.5	214.5	233.8
Overall mean	263.8	255.8	265.6	281.8	260.4	236.0	236.5	250.6
(S.D.)	(48.20)	(48.45)	(50.21)	(69.23)	(46.35)	(42.41)	(49.38)	(50.62)

HITS (CONTINUED)

Analysis of Variance Summary Table:

Order = O, Sex = S, Social Conditions = SC, Periods = P

Source	DF	SS	MS	F	Probability
Subj	15	170277.47			
O	1	4.50	4.50	<1	n.s.
S	1	30135.13	30135.13	2.59	n.s.
O x S	1	385.03	385.03	<1	n.s.
Error	12	139752.81	11646.07		
SC	1	6612.50	6612.50	7.12	<.02
SC x O	1	39691.53	39691.53	42.74	<.001
SC x S	1	132.03	132.03	<1	n.s.
SC x O x S	1	105.13	105.13	<1	n.s.
Error	12	11143.31	928.61		
P	3	4916.03	1638.68	5.71	<.01
P x O	3	881.06	293.69	1.02	n.s.
P x S	3	826.94	275.65	<1	n.s.
P x O x S	3	550.78	183.59	<1	n.s.
Error	36	10329.69	286.94		
SC x P	3	2728.56	909.52	3.15	<.05
SC x P x O	3	127.03	42.34	<1	n.s.
SC x P x S	3	710.03	236.68	<1	n.s.
SC x P x O x S	3	266.69	88.90	<1	n.s.
Error	36	10391.19	288.64		

EXPERIMENT II

Task : Cancellation/RDS, N = 100

Variables: Between - Social Conditions(5), Instructions(2), Sex(2)

PERFORMANCE MEASURE - OUTPUT

Means and Standard Deviations:

	Conditions				
	A	A-T	Aud	Im-Comp	Ex-Comp
Males	887.6	911.9	946.4	1100.5	985.1
Females	869.1	1026.9	1035.0	1104.3	1015.7
Overall Mean	878.4	969.4	990.7	1102.4	1000.4
(S.D.)	(195.9)	(193.8)	(155.0)	(183.0)	(122.7)

OUTPUT (CONTINUED)

Planned Comparisons Summary Table:

Social Conditions = SC, Sex = S

Comparisons:

SC_1 - Alone vs Group

SC_2 - Ego-Threat vs Neutral Instructions

SC_3 - $SC_1 \times SC_2$

SC_4 - Audience vs Alone

SC_5 - Audience vs Im-Comp

Source	DF	SS	MS	F	Probability
SC Comparisons					
SC_1	1	325252.51	325252.51	10.24	<.001
SC_2	1	599.51	599.51	<1	n.s.
SC_3	1	186341.51	186341.51	5.87	<.025
SC_4	1	126225.20	126225.20	3.98	.05
SC_5	1	124768.90	124768.90	3.93	.05
S	1	48180.20	48180.20	1.52	n.s.
$S \times SC_1$	1	4820.51	4820.51	<1	n.s.
$S \times SC_2$	1	8176.91	8176.91	<1	n.s.
$S \times SC_3$	1	14231.11	14231.11	<1	n.s.
$S \times SC_4$	1	28676.05	28676.05	<1	n.s.
$S \times SC_5$	1	17977.60	17977.60	<1	n.s.
Error	90	2857799.10	31753.32		

PERFORMANCE MEASURE - OMISSIONS (PER LINE)

Means and Standard Deviations:

	Conditions				
	A	A-T	Aud	Im-Comp	Ex-Comp
Males	.47	.90	.68	.60	.53
Females	.58	.69	.44	.68	.45
Overall Mean	.52	.79	.56	.64	.49
(S.D.)	(.43)	(.34)	(.33)	(.35)	(.29)

OMISSIONS (CONTINUED)

Planned Comparisons Summary Table:

Social Conditions = SC, Sex = S

Comparisons:

SC_1 - Alone vs Group

SC_2 - Ego-Threat vs Neutral Instructions

SC_3 - $SC_1 \times SC_2$

SC_4 - Audience vs Alone

SC_5 - Audience vs Im-Comp

Source	DF	SS	MS	F	Probability
SC Comparisons					
SC_1	1	0.17	0.17	1.41	n.s.
SC_2	1	0.07	0.07	<1	n.s.
SC_3	1	0.89	0.89	7.44	<.005
SC_4	1	0.01	0.01	<1	n.s.
SC_5	1	0.07	0.07	<1	n.s.
S	1	0.12	0.12	<1	n.s.
$S \times SC_1$	1	0.01	0.01	<1	n.s.
$S \times SC_2$	1	0.28	0.28	2.36	n.s.
$S \times SC_3$	1	0.92	0.92	<1	n.s.
$S \times SC_4$	1	0.29	0.29	2.44	n.s.
$S \times SC_5$	1	0.04	0.04	<1	n.s.
Error	90	10.73	0.12		

PERFORMANCE MEASURE - HITS

Means and Standard Deviations:

	Conditions				
	A	A-T	Aud	Im-Comp	Ex-Comp
Males	815.1	779.8	839.9	990.9	899.9
Females	794.1	911.8	960.1	979.2	942.7
Overall Mean	804.6	845.8	900.0	985.1	921.3
(S.D.)	(175.4)	(187.0)	(150.5)	(165.5)	(128.51)

HITS (CONTINUED)

Planned Comparisons Summary Table:

Social Conditions = SC, Sex = S

Comparisons:

- SC_1 - Alone vs Group
- SC_2 - Ego-Threat vs Neutral Instructions
- SC_3 - $SC_1 \times SC_2$
- SC_4 - Audience vs Alone
- SC_5 - Audience vs Im-Comp

Source	DF	SS	MS	F	Probability
SC Comparisons					
SC_1	1	327552.01	327552.01	12.59	<.001
SC_2	1	2542.51	2542.51	<1	n.s.
SC_3	1	55072.51	55072.51	2.12	n.s.
SC_4	1	91011.60	91011.60	3.50	n.s.
SC_5	1	72335.03	72335.03	2.78	n.s.
S	1	68801.29	68801.29	2.64	n.s.
$S \times SC_1$	1	7980.01	7980.01	<1	n.s.
$S \times SC_2$	1	20853.11	20853.11	<1	n.s.
$S \times SC_3$	1	12127.81	12127.81	<1	n.s.
$S \times SC_4$	1	49843.60	49843.60	1.92	n.s.
$S \times SC_5$	1	43494.03	43494.03	1.67	n.s.
Error	90	2342010.30	26022.34		

MEASURE - RDS

Means and Standard Deviations:

	Conditions				
	A	A-T	Aud	Im-Comp	Ex-Comp
Males	2.80	3.30	2.80	2.80	3.34
Females	2.95	3.88	3.20	2.96	2.80
Overall Mean	2.88	3.59	3.00	2.88	3.07
(S.D.)	(.82)	(.84)	(.88)	(.88)	(.69)

MEASURE RDS (CONTINUED)

Planned Comparisons Summary Table:

Social Conditions = SC, Sex = S

Comparisons:

SC_1 - Alone vs Group

SC_2 - Ego-Threat vs Neutral Instructions

SC_3 = $SC_1 \times SC_2$

SC_4 - Audience vs Alone

SC_5 - Audience vs Im-Comp

Source	DF	SS	MS	F	Probability
SC Comparisons					
SC_1	1	2576.45	2576.45	1.98	n.s.
SC_2	1	7960.05	7960.05	6.10	<.025
SC_3	1	2622.05	2622.05	2.01	n.s.
SC_4	1	297.03	297.03	<1	n.s.
SC_5	1	286.23	286.23	<1	n.s.
S	1	1108.89	1108.89	<1	n.s.
$S \times SC_1$	1	2894.25	2894.25	2.22	n.s.
$S \times SC_2$	1	174.05	174.05	<1	n.s.
$S \times SC_3$	1	3034.25	3034.25	2.33	n.s.
$S \times SC_4$	1	297.03	297.03	<1	n.s.
$S \times SC_5$	1	265.23	265.23	<1	n.s.
Error	90	117396.50	1304.41		

EXPERIMENT III

Task: Cancellation/RDS, N = 24

Variables: Between - Mode of Reporting RDS(2) and Sex(2)

DAY 1 - EXPERIMENTER ONLY (E)

PERFORMANCE MEASURE - OUTPUT

Means and Standard Deviations:

	Written (E-T)	Oral (E-T(F))
Males	374.3	352.7
Females	414.0	405.2
Overall Mean	394.2	378.9
(S.D.)	(60.9)	(80.9)

Analysis of Variance Summary Table:

Mode of Reporting RDS = MR, Sex = S

Source	df	SS	MS	F	Probability
MR	1	1395.00	1395.00	<1	n.s.
S	1	12742.00	12742.00	2.32	n.s.
MR x S	1	247.00	247.00	<1	n.s.
Error	20	110073.00	5503.65		
Total	23	124457.00			

PERFORMANCE MEASURE - OMISSIONS (PER LINE)

Means and Standard Deviations:

	Written (E-T)	Oral (E-T(F))
Males	.57	.71
Females	.81	.96
Overall Mean (S.D.)	.69 (.37)	.84 (.51)

Analysis of Variance Summary Table:

Mode of Reporting RDS = MR, Sex = S

Source	df	SS	MS	F	Probability
MR	1	0.14	0.14	<1	n.s.
S	1	0.35	0.35	1.59	n.s.
MR x S	1	0.00	0.00	<1	n.s.
Error	20	4.41	0.22		
Total	23	4.89			

PERFORMANCE MEASURE - HITS

Means and Standard Deviations:

	Written (E-T)	Oral (E-T(F))
Males	338.0	316.7
Females	361.3	339.0
Overall Mean (S.D.)	349.7 (56.1)	327.8 (75.1)

Analysis of Variance Summary Table:

Mode of Reporting RDS = MR, Sex = S

Source	df	SS	MS	F	Probability
MR	1	2860.00	2860.00	<1	n.s.
S	1	3128.00	3128.00	<1	n.s.
MR x S	1	1.00	1.00	<1	n.s.
Error	20	102298.00	5114.90		
Total	23	108287.00			

PERFORMANCE MEASURE - RDS

Means and Standard Deviations:

	Written (E-T)	Oral (E-T(F))
Males	3.83	3.59
Females	3.33	3.96
Overall Mean (S.D.)	3.58 (1.01)	3.78 (0.89)

Analysis of Variance Summary Table:

Mode of Reporting RDS = MR, Sex = S

Source	df	SS	MS	F	Probability
MR	1	96.00	96.00	<1	n.s.
S	1	10.63	10.63	<1	n.s.
MR x S	1	450.69	450.69	1.19	n.s.
Error	20	7586.69	379.33		
Total	23	8144.00			

DAY II - GROUP

PERFORMANCE MEASURE - OUTPUT (DAY II - DAY I)

Means and Standard Deviations:

	Written (Ex-Comp)	Oral (Ex-Comp(F))
Males	38.2	55.5
Females	39.7	48.7
Overall Mean	38.9	52.1
(S.D.)	(49.1)	(64.5)

Analysis of Variance Summary Table:

Mode of Reporting RDS = MR, Sex = S

Source	df	SS	MS	F	Probability
MR	1	1040.16	1040.16	<1	n.s.
S	1	42.66	42.66	<1	n.s.
MR x S	1	104.17	104.17	<1	n.s.
Error	20	78780.94	3939.05		
Total	23	79968.00			

PERFORMANCE MEASURE - OMISSIONS (PER LINE) (DAY II - DAY I)

Means and Standard Deviations:

	Written (Ex-Comp)	Oral (Ex-Comp(F))
Males	-.18	-.16
Females	-.44	-.53
Overall Mean (S.D.)	-.31 (.29)	-.35 (.32)

Analysis of Variance Summary Table:

Mode of Reporting RDS = MR, Sex = S

Source	df	SS	MS	F	Probability
MR	1	0.08	0.08	1.29	n.s.
S	1	0.33	0.33	5.00	<.05
MR x S	1	0.11	0.11	1.76	n.s.
Error	20	1.31	0.07		
Total	23	1.83			

PERFORMANCE MEASURE - HITS (DAY II - DAY I)

Means and Standard Deviations:

	Written (Ex-Comp)	Oral (Ex-Comp(F))
Males	46.7	56.5
Females	65.3	83.8
Overall Mean (S.D.)	56.0 (44.0)	70.2 (50.4)

Analysis of Variance Summary Table:

Mode of Reporting RDS = MR, Sex = S

Source	df	SS	MS	F	Probability
MR	1	1204.19	1204.19	<1	n.s.
S	1	3174.00	3174.00	1.26	n.s.
MR x S	1	112.63	112.63	<1	n.s.
Error	20	50431.06	2521.55		
Total	23	54921.88			

PERFORMANCE MEASURE - RDS (DAY II - DAY I)

Means and Standard Deviations:

	Written (Ex-Comp)	Oral (Ex-Comp(F))
Males	.15	.42
Females	-.17	.10
Overall Mean	-.01	.26
(S.D.)	(.43)	(.71)

Analysis of Variance Summary Table:

Mode of Reporting RDS = MR, Sex = S

Source	df	SS	MS	F	Probability
MR	1	181.50	181.50	1.37	n.s.
S	1	253.50	253.50	1.89	n.s.
MR x S	1	4.17	4.17	<1	n.s.
Error	20	2676.67	2676.67		
Total	23	3115.83			

EXPERIMENT IV

Task: Cancellation/RDS, N = 90

Variables: Between - Social Conditions(3) and Sex(2)

MEASURE - OUTPUT (DAY II - PRETEST)

Means and Standard Deviations:

	Conditions E	Im-Comp	Ex-Comp
Males	-1.20	39.67	22.33
Females	50.07	15.67	11.67
Overall Mean (S.D.)	24.43 (43.1)	27.67 (46.7)	17.00 (44.1)

Analysis of Variance Summary Table:

Social Conditions = SC, Sex = S

Source	df	SS	MS	F	Probability
SC	2	1794.87	897.43	<1	n.s.
S	1	688.90	688.90	<1	n.s.
SC x S	2	24196.47	12098.23	6.58	<.005
Error	84	154492.67	1839.20		
Total	89	181172.90			

MEASURE - OMISSIONS (PER LINE) (DAY II - PRETEST)

Means and Standard Deviations:

	Conditions E	Im-Comp	Ex-Comp
Males	0.42	0.34	0.46
Females	0.25	0.38	0.46
Overall Mean (S.D.)	0.33 (0.19)	0.36 (0.31)	0.46 (0.36)

Analysis of Variance Summary Table:

Social Conditions = SC, Sex = S

Source	df	SS	MS	F	Probability
SC	2	0.26	0.13	1.38	n.s.
S	1	0.03	0.03	<1	n.s.
SC x S	2	0.18	0.09	<1	n.s.
Error	84	7.77	0.09		
Total	89	8.24			

MEASURE - HITS (DAY II - PRETEST)

Means and Standard Deviations:

	Conditions E	Im-Comp	Ex-Comp
Males	28.07	60.13	50.47
Females	64.93	48.67	49.47
Overall Mean (S.D.)	46.50 (35.17)	54.40 (38.79)	49.97 (34.24)

Analysis of Variance Summary Table:

Social Conditions = SC, Sex = S

Source	df	SS	MS	F	Probability
SC	2	940.82	470.41	<1	n.s.
S	1	1488.40	1488.40	1.18	n.s.
SC x S	2	9698.87	4849.43	3.83	<.025
Error	84	106238.40	1264.74		
Total	89	118366.49			

MEASURE - NUMBER OF CORRECT DIGITS (DAY II - PRETEST)

Means and Standard Deviations:

	E	Conditions Im-Comp	Ex-Comp
Males	.60	1.20	2.60
Females	-.13	-0.20	.93
Overall Mean	.23	.50	1.80
(S.D.)	(3.7)	(4.3)	(3.5)

Analysis of Variance Summary Table:

Social Conditions = SC, Sex = S

Source	df	SS	MS	F	Probability
SC	2	40.27	20.13	1.41	n.s.
S	1	36.10	36.10	2.52	n.s.
SC x S	2	79.83	39.92	2.79	<.10
Error	84	1202.30	14.31		
Total	89	1358.50			

EXPERIMENT V

Task: Alphabetic Transformation, Syllogistic Reasoning, N = 60

Variables: Between - Social Conditions(3), Sex(2), Periods(2)

PERFORMANCE MEASURE - TRANSFORMATION OUTPUT (PER MINUTE) (DAY II - PRETEST)

Means and Standard Deviations:

Conditions

	E			Im-Comp			Im-Comp (F)		
Periods	1	2	Mean	1	2	Mean	1	2	Mean
Males	-2.02	-0.35	-1.18	-2.32	-0.45	-1.38	-1.65	-0.02	-0.83
Females	-1.40	0.07	-0.67	-4.00	-1.57	-2.78	-1.77	-0.03	-0.90
Overall Mean	-1.71	-0.14	-0.93	-3.16	-1.01	-2.08	-1.71	-0.03	-0.87
(S.D.)	(1.6)	(1.4)	(1.38)	(2.7)	(2.5)	(2.51)	(2.2)	(2.4)	(2.22)

Analysis of Variance Summary Table:

Social Conditions = SC, Sex = S, Periods = P

Source	df	SS	MS	F	Probability
Subj	59	558.70			
SC	2	37.68	18.84	2.04	n.s.
S	1	3.01	3.01	<1	n.s.
SC x S	2	19.32	9.66	1.05	n.s.
Error	54	498.69	9.24		
P	1	97.25	97.25	98.74	<.001
P x SC	2	1.91	0.96	<1	n.s.
P x S	1	0.18	0.18	<1	n.s.
P x SC x S	2	0.75	0.37	<1	n.s.
Error	54	53.18	0.99		

PERFORMANCE MEASURE - TRANSFORMATION ERRORS (PER MINUTE) (DAY II - PRETEST)

Means and Standard Deviations:

Conditions

	E			Im-Comp			Im-Comp (F)		
Periods	1	2	Mean	1	2	Mean	1	2	Mean
Males Females	0.25 0.15	0.22 -0.05	0.23 0.05	0.07 -0.05	0.20 -0.02	0.13 - .03	0.13 0.00	0.47 0.23	0.30 0.12
Overall Mean (S.D.)	0.20 (.60)	0.08 (.50)	0.14 (0.51)	0.01 (.57)	0.09 (.41)	0.05 (0.46)	0.07 (.70)	0.35 (.60)	0.21 (0.63)

Analysis of Variance Summary Table:

Social Conditions = SC, Sex = S, Periods = P

Source	df	SS	MS	F	Probability
Subj	59	34.62			
SC	2	0.50	0.25	<1	n.s.
S	1	0.95	0.95	1.54	n.s.
SC x S	2	0.00	0.00	<1	n.s.
Error	54	33.17	0.61		
P	1	0.21	0.21	2.63	n.s.
P x SC	2	0.81	0.40	5.10	<.01
P x S	1	0.11	0.11	1.39	n.s.
P x SC x S	2	0.01	0.00	<1	n.s.
Error	54	4.26	0.08		

PERFORMANCE MEASURE - TRANSFORMATION HITS (PER MINUTE) (DAY II - PRETEST)

Means and Standard Deviations:

Conditions

	E			Im-Comp			Im-Comp (F)		
Periods	1	2	Mean	1	2	Mean	1	2	Mean
Males	-1.77	-0.13	-0.95	-2.28	-0.25	-1.27	-1.08	0.45	-0.32
Females	-1.25	0.02	-0.62	-4.05	-1.58	-2.82	-1.77	0.20	-0.78
Overall Mean	-1.51	-0.06	-0.78	-3.17	-0.92	-2.04	-1.43	0.33	-0.55
(S.D.)	(1.65)	(1.34)	(1.37)	(2.79)	(2.55)	(2.59)	(2.35)	(2.29)	(2.16

Analysis of Variance Summary Table:

Social Conditions = SC, Sex = S, Periods = P

Source	df	SS	MS	F	Probability
Subj	59	581.25			
SC	2	51.47	25.74	2.77	<.07
S	1	9.43	9.43	1.01	n.s.
SC x S	2	17.89	8.94	<1	n.s.
Error	54	502.46	9.30		
P	1	99.04	99.04	93.01	<.001
P x SC	2	3.37	1.63	1.53	n.s.
P x S	1	0.21	0.21	<1	n.s.
P x SC x S	2	1.07	0.54	<1	n.s.
Error	54	57.50	1.06		

PERFORMANCE MEASURE - LOGICAL SYLLOGISM OUTPUT (PER MINUTE) (DAY II -PRETEST)

Means and Standard Deviations:

Conditions

	E			Im-Comp			Im-Comp (F)		
Periods	1	2	Mean	1	2	Mean	1	2	Mean
Males Females	-0.90 -0.82	-0.93 0.05	-0.92 -0.38	-0.40 -1.50	0.43 0.30	0.02 -0.60	-0.47 0.55	-0.07 1.35	-0.27 0.95
Overall Mean (S.D.)	-0.86 (2.44)	-0.44 (1.88)	-0.65 (2.07)	-0.95 (1.62)	0.37 (2.27)	-0.29 (1.74)	0.04 (1.95)	0.64 (2.42)	0.34 (2.02)

Analysis of Variance Summary Table:

Social Conditions = SC, Sex = S, Periods = P

Source	df	SS	MS	F	Probability
Subj	59	468.41			
SC	2	20.18	10.09	1.28	n.s.
S	1	4.28	4.28	<1	n.s.
SC x S	2	17.22	8.61	1.09	n.s.
Error	54	426.74	7.90		
P	1	18.12	18.12	12.75	.001
P x SC	2	4.51	2.25	1.59	n.s.
P x S	1	4.29	4.29	3.02	n.s.
P x SC x S	2	0.48	0.24	<1	n.s.
Error	54	76.74	1.42		

PERFORMANCE MEASURE - LOGICAL SYLLOGISMS ERRORS (PER MINUTE) (DAY II - PRETEST)

Means and Standard Deviations:

Conditions

	E			Im-Comp			Im-Comp (F)		
Periods	1	2	Mean	1	2	Mean	1	2	Mean
Males	1.03	0.97	1.00	0.43	0.67	0.30	0.53	0.83	0.68
Females	0.68	0.65	0.67	0.50	0.33	0.42	-0.08	-0.35	-0.22
Overall Mean	0.86	0.81	0.83	0.47	0.25	0.36	0.22	0.24	0.23
(S.D.)	(1.12)	(1.13)	(1.04)	(0.85)	(1.43)	(1.03)	(1.08)	(1.28)	(1.10)

Analysis of Variance Summary Table:

Social Conditions = SC, Sex = S, Periods = P

Source	df	SS	MS	F	Probability
Subj	59	140.78			
SC	2	8.05	4.02	1.76	n.s.
S	1	4.16	4.16	1.82	n.s.
SC x S	2	5.20	2.60	1.14	n.s.
Error	54	123.38	2.28		
P	1	0.20	0.20	<1	n.s.
P x SC	2	0.29	0.15	<1	n.s.
P x S	1	0.16	0.16	<1	n.s.
P x SC x S	2	0.67	0.34	<1	n.s.
Error	54	26.71	0.49		

PERFORMANCE MEASURE - LOGICAL SYLLOGISMS HITS (PER MINUTE) (DAY II - PRETEST)

Means and Standard Deviations:

Conditions

	E			Im-Comp			Im-Comp.(F)		
Periods	1	2	Mean	1	2	Mean	1	2	Mean
Males	0.13	0.04	0.08	0.04	0.60	0.32	0.07	0.77	0.42
Females	-0.13	0.70	0.28	-1.00	0.63	-0.18	0.47	1.00	0.73
Overall Mean (S.D.)	-0.00 (2.10)	0.37 (1.95)	0.18 (1.88)	-0.48 (1.31)	0.62 (2.22)	0.07 (1.52)	0.27 (2.07)	0.88 (1.93)	0.58 (1.80)

Analysis of Variance Summary Table:

Social Conditions = SC, Sex = S, Periods = P

Source	df	SS	MS	F	Probability
Subj	59	366.67			
SC	2	5.66	2.83	<1	n.s.
S	1	0.00	0.00	<1	n.s.
SC x S	2	3.91	1.96	<1	n.s.
Error	54	357.09	6.61		
P	1	14.48	14.48	9.08	<.005
P x SC	2	2.75	1.37	<1	n.s.
P x S	1	2.80	2.80	1.76	n.s.
P x SC x S	2	2.28	1.14	<1	n.s.
Error	54	86.08	1.59		

EXPERIMENT VI

Task: Pursuit Rotor, N = 96

Variables: Between - Social Conditions(2), Instructions(2), Feedback(2), Sex(2)

PERFORMANCE MEASURE - T.O.T. PHASE I (PERIOD 1 - PRACTICE)

Means and Standard Deviations:

	Conditions							
	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Comp (F)
Males	10.53	9.93	8.08	11.30	4.82	9.55	9.70	7.95
Females	6.27	11.75	9.62	9.95	4.83	9.57	8.43	6.23
Overall Mean (S.D.)	8.40 (5.5)	10.84 (2.4)	8.85 (2.0)	10.63 (4.0)	4.83 (3.6)	9.56 (3.6)	9.07 (4.0)	7.09 (4.1)

Intrasubject Variance Means:

	Conditions							
	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Comp (F)
	2.56	1.51	1.34	1.42	2.42	1.83	2.30	2.04

Analysis of Variance Summary Table:

Social Conditions = SC, Instructions = I, Feedback = F, Sex = S

Source	df	SS	MS	F	Probability
SC	1	100.25	100.25	6.24	<.025
I	1	6.05	6.05	<1	n.s.
SC x I	1	3.57	3.57	<1	n.s.
F	1	72.98	72.98	4.54	<.05
SC x F	1	3.19	3.19	<1	n.s.
S	1	10.21	10.21	<1	n.s.
SC x S	1	0.18	0.18	<1	n.s.
I x F	1	81.59	81.59	5.08	<.025
SC x I x F	1	54.75	54.75	3.41	<.065
I x S	1	0.06	0.06	<1	n.s.
SC x I x S	1	11.97	11.97	<1	n.s.
F x S	1	2.84	2.84	<1	n.s.
SC x F x S	1	5.00	5.00	<1	n.s.
I x F x S	1	33.25	33.25	2.07	n.s.
SC x I x F x S	1	27.20	27.20	1.69	n.s.
Error	80	1284.66	1284.66		
Total	95	1697.71			

PERFORMANCE MEASURE - T.O.T. PHASE II (PERIOD 3 - PERIOD 1)

Means and Standard Deviations:

	Conditions							
	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Comp (F)
Males	-0.33	1.33	4.42	3.48	4.38	4.08	6.47	4.47
Females	6.05	2.63	1.52	3.20	6.67	2.82	4.37	4.42
Overall Mean	2.86	1.98	2.97	3.34	5.53	3.45	5.42	4.44
(S.D.)	(6.9)	(5.1)	(3.6)	(3.9)	(3.1)	(5.3)	(3.5)	(4.3)

Intrasubject Variance Means:

	Conditions							
	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp (F)	Ex-Comp	Ex-Comp (F)
	2.56	1.51	1.34	1.42	2.42	1.83	2.30	2.04

Analysis of Variance Summary Table:

Social Conditions = SC, Instructions = I, Feedback = F, Sex = S

Source	df	SS	MS	F	Probability
SC	1	88.55	88.55	3.83	<.05
I	1	8.28	8.28	<1	n.s.
SC x I	1	0.51	0.51	<1	n.s.
F	1	18.90	18.90	<1	n.s.
SC x F	1	9.75	9.75	<1	n.s.
S	1	4.25	4.25	<1	n.s.
SC x S	1	11.90	11.90	<1	n.s.
I x F	1	8.28	8.28	<1	n.s.
SC x I x F	1	0.03	0.03	<1	n.s.
I x S	1	73.85	73.85	3.19	<.08
SC x I x S	1	22.23	22.23	<1	n.s.
F x S	1	5.90	5.90	<1	n.s.
SC x F x S	1	0.35	0.35	<1	n.s.
I x F x S	1	66.33	66.33	2.87	<.1
SC x I x F x S	1	1.65	1.65	<1	n.s.
Error	80	1850.79	23.13		
Total	95	2171.58			

EXPERIMENT VII

Task: Darts, N = 20

Variables: Between - Sex(2), Within - Social Conditions(3), Days(3),
Axis of Dispersal(2)

Means and Standard Deviations*

Raw Score Means for all Conditions (millimetres)

	Baseline 1			Video 1			Post-Video 1		
	M	Var(x)	Var(y)	M	Var(x)	Var(y)	M	Var(x)	Var(y)
Males	14.0	71.1	140.0	14.8	85.9	133.8	13.9	67.7	108.6
Females	17.1	139.9	194.3	15.4	122.6	125.2	17.3	153.0	159.7
Overall Mean	15.6	105.5	167.2	15.1	104.3	129.5	15.6	110.4	134.2

	Baseline 2			Video 2			Post-Video 2		
	M	Var(x)	Var(y)	M	Var(x)	Var(y)	M	Var(x)	Var(y)
Males	15.1	79.7	96.5	11.7	70.4	97.7	13.7	66.6	111.1
Females	16.5	128.2	145.5	16.1	122.4	110.2	16.5	121.3	159.7
Overall Mean	15.8	104.0	121.0	13.9	96.4	104.0	15.1	94.0	135.4

	Baseline 3			Video/Competition		
	M	Var(x)	Var(y)	M	Var(x)	Var(y)
Males	14.0	71.7	98.8	14.4	65.7	110.7
Females	16.9	126.2	145.7	17.6	124.1	159.9
Overall Mean	15.5	99.0	122.3	16.0	94.9	135.3

Continued/

Continued

Mean Ratios for all Conditions
(With Log_{10} Transformations in Brackets)

	Video 1/Baseline 1			Post-Video 1/Baseline 1		
	M	Var(x)	Var(y)	M	Var(x)	Var(y)
Males (log)	1.07 (0.026)	1.40 (0.055)	1.02 (-0.038)	1.04 (0.005)	1.10 (-0.001)	1.06 (-0.081)
Females (log)	0.90 (-0.055)	1.16 (0.023)	0.85 (-0.151)	0.98 (-0.024)	1.40 (0.069)	0.79 (-0.115)
Overall Mean (log) (S.D.)*	0.99 (-.015) (.09)	1.28 (.039) (.26)	.94 (-.095) (.24)	1.01 (-.009) (.11)	1.25 (.034) (.23)	.98 (-.098) (.28)

	Video 2/Baseline 2			Post-Video 2/Baseline 2		
	M	Var(x)	Var(y)	M	Var(x)	Var(y)
Males (log)	0.96 (-0.026)	1.13 (-0.025)	1.03 (-0.026)	1.01 (-0.002)	1.18 (0.017)	1.19 (0.043)
Females (log)	1.01 (-0.004)	1.16 (0.005)	0.92 (-0.124)	1.03 (0.008)	1.24 (0.008)	1.21 (0.020)
Overall Mean (log) S.D.*	.97 (-.015) (.08)	1.15 (-.010) (.24)	.98 (-.075) (.26)	1.02 (.003) (.06)	1.21 (.013) (.26)	1.20 (.032) (.21)

Continued/

Continued

	Video-Competition/Baseline 3		
	M	Var(x)	Var(y)
Males (log)	1.07 (0.010)	1.44 (0.064)	1.40 (0.077)
Females (log)	1.05 (0.018)	1.08 (-0.003)	1.18 (0.039)
Overall Mean (log) S.D.*	1.06 (.014) (.10)	1.26 (.031) (.23)	1.29 (.058) (.22)

*S.D.'s are for the distribution of logged scores

Analysis of Variance Summary Table:

Between: Sex = S
Within: Social Conditions (Video, Post-Video) = SC, Days (1 and 2) = D, Axis of Dispersal (Horizontal, Vertical) = AD

Source	df	SS	MS	F	Probability
Subj	19	1.839			
S	1	0.027	0.027	<1	n.s.
Error	18	1.813	0.101		
SC	1	0.037	0.037	1.36	n.s.
SC x S	1	0.029	0.029	1.09	n.s.
Error	18	0.484	0.027		
D	1	0.016	0.016	<1	n.s.
D x S	1	0.000	0.000	<1	n.s.
Error	18	3.067	0.170		
AD	1	0.242	0.242	2.62	n.s.
AD x S	1	0.067	0.067	<1	n.s.
Error	18	1.657	0.092		
SC x D	1	0.048	0.048	1.98	n.s.
SC x D x S	1	0.013	0.013	<1	n.s.
Error	18	0.434	0.024		
SC x AD	1	0.018	0.018	<1	n.s.
SC x AD x S	1	0.005	0.005	<1	n.s.
Error	18	0.628	0.035		
D x AD	1	0.120	0.120	1.99	n.s.
D x AD x S	1	0.001	0.001	<1	n.s.
Error	18	1.089	0.061		
SC x D x AD	1	0.017	0.017	<1	n.s.
SC x D x AD x S	1	0.012	0.012	<1	n.s.
Error	18	0.473	0.026		

APPENDIX IV

BADDELEY'S (1968) 3-MINUTE REASONING TEST (ADAPTED FOR USE IN EXPERIMENT V)

SAMPLE TEST

Sentence Checking Name ____________________

In the following test there are a number of short sentences each followed by a pair of letters (AB or BA). The sentences claim to describe the order of the two letters i.e. to say which comes first. They can do this in several different ways. Thus the order AB can be correctly described by saying either (1) A precedes B, or (2) B follows A, or (3) B does not precede A, or (4) A does not follow B. All these are correct descriptions of the pair AB but are incorrect when applied to the other pair, BA.

Your job is to read each sentence and to decide whether it is a true or false description of the letter pair which follows it. If you think that the sentence describes the letter pair correctly put a tick in the first column (labelled True). If you think the sentence does not give a true description of the letter order, put a tick in the second ("False") column.

This is illustrated in examples 1 and 2 below. When you have read 1 and 2, try examples 3, 4, 5, and 6.

Examples	True	False
1. A follows B - BA	✓	
2. B precedes A - AB		✓
3. A is followed by B - AB		
4. B is not followed by A - BA		
5. B is preceded by A - BA		
6. A does not precede B - BA		

When you start the main test, work as quickly as you can without making mistakes. Start with sentence 1 and work systematically through the test leaving no blank spaces.

Do NOT begin the main test until you are given the start signal.

Name

	True	False		True	False		True	False
1. A does not precede B - AB			22. A does not precede B - BA			43. B is preceded by A - AB		
2. A is not preceded by B - BA			23. B precedes A - BA			44. A is not preceded by B - AB		
3. B is followed by A - BA			24. A precedes B - AB			45. B is followed by A - AB		
4. B does not follow A - BA			25. A is not followed by B - BA			46. B precedes A - BA		
5. A is preceded by B - AB			26. B follows A - AB			47. B does not precede A - BA		
6. A does not follow B - BA			27. B is preceded by A - AB			48. B follows A - AB		
7. B does not precede A - BA			28. B follows A - BA			49. A does not follow B - BA		
8. A is preceded by B - BA			29. A is followed by B - BA			50. B precedes A - AB		
9. A does not follow B - AB			30. B does not follow A - AB			51. B is not preceded by A - AB		
10. A follows B - BA			31. B is not followed by A - BA			52. B does not follow A - BA		
11. B precedes A - AB			32. A is not followed by B - AB			53. A does not follow B - AB		
12. B is not followed by A - AB			33. A follows B - AB			54. A is preceded by B - BA		
13. B is not preceded by A - BA			34. A is followed by B - BA			55. A is followed by B - AB		
14. A follows B - AB			35. B follows A - BA			56. B is not preceded by A - BA		
15. B is not preceded by A - AB			36. A is not preceded by B - BA			57. B does not precede A - AB		
16. B is followed by A - AB			37. A follows B - BA			58. A is not followed by B - BA		
17. B is preceded by A - BA			38. A is preceded by B - AB			59. A does not precede B - AB		
18. A is not preceded by B - AB			39. B is followed by A - BA			60. A precedes B - BA		
19. A precedes B - BA			40. B is not preceded by A - AB			61. B is not followed by A - AB		
20. B does not precede A - AB			41. A does not precede B - BA			62. A precedes B - AB		
21. A is followed by B - AB			42. B does not follow A - AB			63. B is preceded by A - BA		
						64. B is not followed by A - BA		

APPENDIX V

SELF-REPORT DATA FROM EXPERIMENT VI: QUESTIONNAIRE AND SUMMARY OF RESULTS

SAMPLE QUESTIONNAIRE

Name:

Sex:

Age:

	Not at all				Very much so
1. Did you feel uneasy about being a subject in a psychology experiment before arriving for the actual experiment?					
2. During the experiment, were you worried that your performance might not be as good as the performance of the other subjects?					
3. Did the presence of others in the room while you were performing make you uncomfortable?					
4. Were you aware of trying to beat your own previous performance standard?					
5. Do you think you could do better on the task in the privacy of your own room?					
6. Were you worried that your performance was being evaluated by the experimenter?					
7. Did you feel distracted by the noise and activity around you?					
8. Were you aware of trying to do better on the task than the other subjects?					
9. Did you feel you were being deceived as to the real purpose of the experiment?					
10. Were you worried that your performance was being evaluated by the other subjects?					

SUMMARY OF MEANS ON SELF-REPORT QUESTIONNAIRE FOR SOCIAL CONDITIONS

	E	E(F)	E-T	E-T(F)	Im-Comp	Im-Comp(F)	Ex-Comp	Ex-Comp(F)
1. Did you feel uneasy about being a subject in a psychology experiment before arriving for the actual experiment?	.5	.5	.17	.33	.58	.5	.33	.42
2. During the experiment were you worried that your performance might not be as good as the performance of the other subjects?	.58	1.5	1.42	1.00	.67	1.42	1.17	1.42
3. Did the presence of others in the room while you were performing make you uncomfortable?	.08	.17	.17	.08	.25	.33	.75	.83
4. Were you aware of trying to beat your own previous performance standard?	3.2	2.7	2.0	2.92	2.58	2.67	2.83	1.58
5. Do you think you could do better on the task in the privacy of your own room?	1.1	.92	.33	.42	.50	.92	1.00	1.50
6. Were you worried that your performance was being evaluated by the experimenter?	.42	.67	.42	.25	.25	.42	.83	.42
7. Did you feel distracted by the noise and activity around you?	.33	.58	.50	.08	.58	.25	.58	1.00
8. Were you aware of trying to do better on the task than the other subjects?	.67	1.00	1.67	.92	.58	1.67	1.25	1.75
9. Did you feel you were being deceived as to the real purpose of the experiment?	.50	.50	.25	.08	.75	.58	1.00	.75
10. (Group only) Were you worried that your performance was being evaluated by the other subjects?	-	-	-	-	.25	.58	.50	.33
Overall Means	.82 N=12	.95 N=12	.77 N=12	.68 N=12	.75 N=12	.97 N=12	1.08 N=12	1.07 N=12

DISCUSSION OF QUESTIONNAIRE RESULTS

Question 1 was designed to assess subjects' subjective arousal level before the administration of any of the testing procedures. As can be seen the ratings are uniformly low. In fact, the subjects were involved with quite a lot of experimental participation as part of their course work, and therefore probably did not regard the experiment with any prior suspicion or anxiety.

Question 2 was intended to tap feelings of evaluation stress, and ratings were expected to be higher in the conditions involving ego-threatening instructions and/or feedback and to be generally higher in the group conditions. The table reveals little difference between group and alone conditions, although ratings for E and Im-Comp are lower than for conditions involving feedback and/or ego-threat. It would seem that these treatments did have some effect in stimulating evaluation fear, although the low ratings indicate that this was not experienced as 'stressful'.

Question 3 was intended to tap any general feelings of discomfort associated with others presence but not perhaps identified with evaluation. As expected the ratings for subjects in group conditions, although low, are uniformly higher than subjects tested with just the experimenter.

Question 4 was designed to assess both the impact of the ego-threatening instructions and feedback in motivating subjects. Unexpectedly, subjects in all conditions gave this question relatively high ratings. These results indicate that, as expected, the task was intrinsically motivating.

Question 5 was intended to indirectly assess subjects' discomfort in the experimental situation by assessing their preference to perform in more private surroundings. The answers are less uniform, and not particularly related to the stress manipulations.

Question 6 attempted to assess any evaluation fear associated with the experimenter's presence. Ratings are uniformly low and do not vary as a function of the ego-threatening instructions.

Question 7 was an indirect measure of subjects' discomfort in the situations. Ratings were expected to increase with the number of people present and sources of potential evaluation. However, only subjects in Ex-Comp(F) show a rating which is noticeably higher than other conditions.

Question 8 was a direct measure of competitiveness. Both E and Im-Comp give ratings which are lower than other conditions. There does not seem to be a progressive effect with the addition of other subjects or evaluative stimuli. However, Ex-Comp(F) subjects yield the highest mean rating on the question.

Question 9 was a measure of subjects' suspiciousness regarding the experimental manipulations, particularly in regard to disbelief in the explanations of the task or the competitive instructions. Subjects tested alone show uniform acceptance of the experimental manipulations. However, subjects' ratings when tested in groups are somewhat higher. This would seem to indicate that the reason for the others' presence was regarded with suspicion, although the ratings may indicate uncertainty rather than disbelief.

Question 10 applied only to subjects tested in groups and was designed to assess discomfort at being observed and perhaps evaluated by the other subjects. The ratings are uniformly low, suggesting that even though subjects were motivated by the task the prospect of giving a poor performance before the other subjects was not alarming.

An overview and further discussion of these results can be found in Section 6.2.4. of the text.

APPENDIX VI

EXPERIMENT VII (DARTS) DATA
ACCORDING TO FREQUENCY OF PLAY: SUMMARY TABLE AND DISCUSSION

Number of individuals showing facilitation(+), decrement(-), and no change(0) in Video and Competition Conditions according to frequency of play.

Group 1 - plays about every day (N = 4)
Group 2 - plays several times a month (N = 8)
Group 3 - plays several times a year (N = 8)

Change in performance from Baseline to Video (Days 1 + 2)

	Group 1 (N = 8)			Group 2 (N = 16)			Group 3 (N = 16)		
	+	-	0	+	-	0	+	-	0
M	5	1	2	7	7	2	6	6	4
Var(x)	2	6	0	6	9	1	10	4	2
Var(y)	6	1	1	11	5	0	7	8	1

Change in performance from Baseline to Post Video (Days 1 + 2)

	Group 1 (N = 8)			Group 2 (N = 16)			Group 3 (N = 16)		
	+	-	0	+	-	0	+	-	0
M	2	2	4	7	7	2	8	8	0
Var(x)	2	6	0	4	11	1	12	4	0
Var(y)	4	3	1	8	6	2	7	9	0

Change in performance from Baseline to Video-Competition

	Group 1 (N = 4)			Group 2 (N = 8)			Group 3 (N = 8)		
	+	-	0	+	-	0	+	-	0
M	1	3	0	5	1	2	1	6	1
Var (x)	1	2	1	5	2	1	2	5	1
Var(y)	0	4	0	4	4	0	4	4	0

DISCUSSION OF FREQUENCY OF PLAY DATA

Looking firstly at the Video conditions, in terms of the stated goal of performance (nearness to the bull), only Group 1 members show any particular advantage: in only one instance (out of 8) was decrement noted for M. Both Groups 1 and 2 have some debilitation in Var(x) and improvement in Var(y), while Group 3 members show improvement in Var(x).

The Post-Video data indicate little difference in the general pattern of performance changes from that of Video, at least for Groups 2 and 3. The facilitation shown by Group 2 for Var(y) is lost in this condition, as the frequencies are roughly equal in the '+' and '-' categories. However, the results for Group 1 are rather different in Post-Video compared to Video. Firstly, the number of cases showing facilitation in M is reduced from 5 to 2. Secondly, only half the cases show facilitation in Var(y). In summary, it would appear that only very frequent players showed any general tendency for better performance under stress and that this advantage was lost (in fact some relative decrement indicated) following the removal of the stress.

The patterns under the Video-Competition category are different yet. In this case the majority of subjects in Groups 1 and 3 demonstate decrement for M. The pattern is most striking in the case of Group 1, where there are only two cases of facilitation for all three measures. Group 2 on the other hand demonstrates some degree of facilitation, at least for M and Var(x). Only one subject shows less overall accuracy (M), neither is there any tendency for the majority of subjects in this group to show decrement on any of the dispersal measures.

No conclusions are drawn from this analysis due to the small number of subjects. An overview of the results and their possible implications can be found in Section 6.3.4 of the text.