SELF-ESTEEM AND REACTIONS TO FEEDBACK:
Theory, Methodology, and Empirical Research

Marilyn Aitkenhead, B.Sc.

Thesis submitted for the degree of
Doctor of Philosophy

University of Leicester,
Department of Psychology

December 1980
To Dad
and
To Raph
with lots of love
(but not forgetting Guizmo Saucepan and his wife!)
ACKNOWLEDGEMENTS

There are many people whose support and advice have been instrumental in the development of this thesis. I should particularly like to thank Professor Geoffrey Stephenson for setting me off along promising lines of research, and for helping me through a difficult period. I am grateful to Dr. Michael Argyle who invited me to seminars at Oxford which were both helpful and stimulating. My supervisor, Dr. Andrew Colman, has helped me in many ways - not least in the final stages of his role as supervisor when, even though extremely busy, he not only read but also made incisive comments on a first draft of the manuscript. Whenever he was needed, he was there: he suggested and advised but did not dominate.

I should also like to thank the many people who served as subjects in my experiments as well as the teachers who gave up their class-time to let me loose on their pupils. My sister, Jean, often served as a guinea-pig in the development of questionnaires, as well as cheerfully helping with data analysis. So thanks, Jean! Anne Robertson did a superb job of typing the thesis, somehow unravelling my illegible scrawl and coping wonderfully with vague instructions.

My greatest debt, however, is to my fiancee, Dr. Raph Gillett. Without him, not only would this thesis never have been completed, but it would also have been twice as long.
ABSTRACT

Despite a great deal of research in the area of self-esteem (SE) and the self-concept, important issues are yet to be resolved. How to define the self-concept and SE is one of these; the personality dynamics underlying reactions to feedback is another. This thesis is divided into two parts. In part I, issues in theory and measurement are discussed; the distinctions between self-concept, SE, and self-assessments are emphasised; and it is suggested that a limitation of much research on SE is its assumption that SE is a relatively unvarying personality trait. This assumption is tested in a number of studies, and found to be unsupported. In addition, an SE questionnaire is developed based on self-ideal discrepancies. Criticisms of this type of SE questionnaire are investigated and found to be empirically unjustified: for example, ideal scores do contribute significantly to the SE score.

Part II tackles the issue of how people with varying levels of SE respond to evaluative feedback. The main theoretical perspectives (consistency, self-enhancement, esteem-protection and the non-motivational hypothesis) are discussed, and predictions derived from each are put to the test. Following Shrauger (1975) a classificatory scheme is proposed whereby affective reactions to feedback are hypothesized to be dominated by self-enhancement needs and non-affective reactions by self-consistency needs. An important modification is introduced which is that consistency needs will be overshadowed by esteem-protection needs when the costs to SE of being consistent are high. Two experiments using repeated trials with one using non-involved high SE and low SE observers, support this modification. A reconceptualization is offered, detailing predictions for three dichotomized independent variables (high vs low SE; positive vs negative feedback; high vs low cost) and a variety of dependent variables (both affective and non-affective). When tested against extant research, the reconceptualization is strongly supported.
## CONTENTS

1. INTRODUCTION TO THE ISSUES AND RESEARCH  
   1.1 Introduction  
   1.2 Towards a Definition of the Self-Concept  
   1.3 Towards a Definition of Self-Esteem, and Self-Assessments  
      1.3(a) The definitions  
      1.3(b) The effects of self-assessments and SE upon behaviour  
   1.4 Predominant Assumptions Made About the Self-Concept, and SE  
   1.5 Are Self-Assessments and SE Personality Traits?  
      1.5(a) The current status of personality traits  
      1.5(b) SE from an 'Interactional' perspective  
   1.6 Methodology in Self-Concept/SE Research  
   1.7 Summary and Proposed Investigations  

2. CONTENT OF SELF-ESTEEM AND COMPARISON-OTHERS AT HOME AND AT SCHOOL  
   Abstract  
   Introduction  
   Method  
   Results  
   Discussion  

3. VARIABILITY OF SELF-ESTEEM ACROSS SITUATIONS  
   Abstract  
   Introduction  
   Method  
   Results  
   Discussion
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. DEVELOPMENT OF A SELF-ESTEEM QUESTIONNAIRE</td>
<td>120</td>
</tr>
<tr>
<td>Abstract</td>
<td>121</td>
</tr>
<tr>
<td>Introduction</td>
<td>122</td>
</tr>
<tr>
<td>1. Development of the SE questionnaire</td>
<td>124</td>
</tr>
<tr>
<td>Introduction</td>
<td>124</td>
</tr>
<tr>
<td>Method</td>
<td>129</td>
</tr>
<tr>
<td>Results</td>
<td>134</td>
</tr>
<tr>
<td>Reliability and validity of the questionnaire</td>
<td>140</td>
</tr>
<tr>
<td>Conclusions</td>
<td>146</td>
</tr>
<tr>
<td>2. An active role-playing study of SE level and content across situations</td>
<td>148</td>
</tr>
<tr>
<td>Introduction</td>
<td>148</td>
</tr>
<tr>
<td>Method</td>
<td>150</td>
</tr>
<tr>
<td>Results</td>
<td>155</td>
</tr>
<tr>
<td>Discussion</td>
<td>161</td>
</tr>
<tr>
<td>SUMMARY AND CONCLUSIONS CHAPTERS 2-4</td>
<td>163</td>
</tr>
<tr>
<td>5. THE SELF-CONSISTENCY/SELF-ENHANCEMENT DEBATE: THEORY AND RESEARCH</td>
<td>167</td>
</tr>
<tr>
<td>Introduction</td>
<td>168</td>
</tr>
<tr>
<td>The theoretical perspectives</td>
<td>169</td>
</tr>
<tr>
<td>(a) The need for consistency: The Theory of Cognitive Dissonance</td>
<td>172</td>
</tr>
<tr>
<td>(b) The need for self-enhancement</td>
<td>192</td>
</tr>
<tr>
<td>(c) The esteem-protection notion</td>
<td>198</td>
</tr>
<tr>
<td>(d) The non-motivational hypothesis</td>
<td>200</td>
</tr>
<tr>
<td>Self-consistency and self-enhancement: Coexistent or mutually exclusive?</td>
<td>203</td>
</tr>
<tr>
<td>Conclusions and proposed investigations</td>
<td>220</td>
</tr>
<tr>
<td>6. THE ART OF COMPROMISE</td>
<td>223</td>
</tr>
<tr>
<td>Abstract</td>
<td>224</td>
</tr>
<tr>
<td>Introduction</td>
<td>225</td>
</tr>
</tbody>
</table>
9. CONCLUSIONS AND DISCUSSION

1. Issues involved in SE measurement
   (a) Implications for investigators
   (b) Implications for measurement
   (c) Implications for theory

2. The relationships between self-assessments; SE; and task-expectancies

3. Reactions to feedback
   (a) Summary of results
   (b) Towards a reconceptualization of the self-enhancement/self-consistency debate
   (c) The reconceptualization and extant research

APPENDICES

List of appendices

Appendices

BIBLIOGRAPHY
VOLUME I

CHAPTERS 1 - 4
CHAPTER 1

Introduction to the Issues and Research
There can be little doubt that the self is one of the most important and well-used concepts within psychology. Even when psychology was just beginning to forge an identity separate from its philosophical roots, the notions of self and self-esteem (SE) played an important role in psychological theorising (eg. James, 1890). In more recent times, the self still occupies the attention of many psychologists (eg. Mischel, 1977; Hamachek, 1978; Kleinke, 1978; Burns, 1979; Harré, 1979), having withstood the onslaught of the behaviourists. During the course of the history of psychology, the notion of self has been made the lynch-pin of several influential theories of personality (eg. Lecky, 1945; Snygg and Combs, 1949; Rogers, 1951; Allport, 1955; Kelly, 1955; Eriksen, 1968). Many of these theories stress the role of the self in personality functioning. For example, Rogers (1951) proposes that unless there is congruence between the organism's experience and its self-structure, psychological tension will occur involving internal strain and anxiety. Another example is Lecky (1945) who suggests that man is constantly striving for unity, or self-consistency. According to Lecky, without a consistent organisation of our experiences we respond to situations as emergencies and are unable to deal effectively with them. This is a similar idea to that of Rogers, implying personality disorganisation unless there is a consistent self-concept.

For many of these theorists, and others, the notion of self is important because they argue that individuals usually behave in ways which are consistent with their self-concept. Insofar as successful social behaviour implies the prediction of how others will behave in given circumstances, it can be seen that the notion of self is extremely important. Indeed Lecky (1945) goes so far to say that 'the new goal (within psychology) is the conceptualization of a unit altogether new
to the scientist, in spite of its age in philosophy ... the self' (p.138).

Of course, consideration of the notion of self has not been confined to
psychology. Mead (1934) argues that the self is an agent of social
control, and elaborates upon the symbolic interactionist school begun by
Cooley (1902). Goffman (1957) emphasizes the importance of self in
social interaction. Problems involving the nature of self have occupied
philosophers for many years.

Yet, despite all this attention, the study of the self remains
problematical: so much so that Gergen (1977) writes 'in the case of
self-knowledge there is virtually nothing to know' (p.166) and 'to the
extent that self-psychology is concerned with internal states, that is,
with evaluative feelings, wants, hopes, aversions, tastes, attitudes,
emotions, beliefs, volition and the like, the prospects for a cumulative
science seem dim' (p.168). Gergen's argument is that what we 'know' of
the self is essentially a social construction. Such a social construction
may vary in a myriad of ways from one situation to another and thus we
can make no generalisable statements about the process of self-conceptua-
lization. This pessimistic outlook is hotly contested in this thesis.
But how might it have arisen in the first place?

A major contributary factor is undoubtedly the lack of concern
psychologists have shown towards the conceptualization of social situations.
This omission has had important implications, not just for the study of
the self, but for social psychology in general. Indeed, Forgas (1979)
argues that much of the disenchantment with 'traditional' social psycho-
logy (such disenchantment being epitomized by the work of Harre and
Secord, 1972) is because there has been insufficient attention directed
towards an understanding of naturally-occurring social situations. A
major divide has arisen between theory (which does stress the role of
situations) and research (which, generally, does not). Only recently has
there been much empirical work in this direction (see, for example,
Magnusson 1971; Forgas, 1976; 1978; 1979), and ways of conceptualizing situations are still in their infancy. Insofar as the conceptualization of situations is important to the understanding of the self, it is not surprising that Gergen should think the prospects for a science of the self seem dim.

A second reason why Gergen feels as he does may be due to the poor quality of much of the research in this area. Wylie (1961) writes of the poor methodology of much of this research, and sets out clearly the assumptions upon which such research is based. She points out that such assumptions require to be tested empirically in order for theorising about the self to advance. Yet more than ten years later (Wylie, 1974) she concludes that much the same dismal situation prevails. As well as poor quality measures, poorly conceived research designs, and rampant overgeneralization, there is also an overabundance of terms which appear to have something in common but which are only loosely (if at all) defined. For example, one investigator's term of self-esteem (SE) may have quite different theoretical underpinnings than has the same term for another investigator. Whilst the term is the same, the implications of the results may be quite different, and using identical terminologies will only cloud this. A solution of dubious value but not infrequently adopted by synthesizers of the area is to use as many of the available terms as possible within their writings. Whilst perhaps being a more accurate reflection of the terminologies used by the diverse original sources, such a strategy can only add to the already considerable confusion of those seeking conceptual clarification.

A recent example of such terminological extravagance is provided by Burns (1979). This confuses the reader because the different terms are not all defined as meaning the same thing (the relevant terms are here underlined): 'The rationale of self-consistency theorists is that an individual's actions, attitudes, and his receptivity to information from
other people are strongly affected by a tendency to create and maintain a consistent cognitive state with respect to his evaluations of himself. This argues that individuals with a high self-concept find positive evaluations from others consistent and negative evaluation inconsistent. Individuals with low self-evaluation would, on self-consistency theory, find negative evaluations consistent, but positive evaluations inconsistent'. In contrast, self-esteem theory implies that everyone needs to bear favourable attitudes towards himself, and that the more strongly this need is frustrated the more strongly the individual will wish to have it satisfied. Since low self-esteem individuals would be more hungry for and yet more frustrated in their needs for positive self-evaluations, they should respond more favourably to positive evaluations from others and respond in a more dejected and hostile way to failure than should individuals high in self-esteem. In a review of studies of these two theories Jones (1973) found that self-esteem theory was more appropriate to explain behaviour, with the unhappy self-derogator glowing when praised and unhappy when derided. Self-consistency theory was rarely supported in the numerous studies he reviewed.' (p.242).

In addition to such terminological confusion, there is also a proliferation of measures purporting to assess self-esteem (SE). For many of these, there is no information as to their reliability or validity, and the criteria for item-selection remain the secret of the investigator. Wylie (1974) reviews such measures in some detail. She concludes 'Granting the many practical difficulties, and granting that there are no easy guides to creative thinking about hypotheses and ways to test them, still, many inexcusable and avoidable flaws occur in current research'. (p.330).

Thus, a peculiar situation has arisen in this area. On the one hand, many theorists stress how the self-concept and SE are of central importance to psychology and are crucial to our ability to function adequately
in our society. (Though whether the self is an important construct in other societies remains to be seen). As Gergen (1971) puts it 'An individual who does not possess a sizeable number of labels for himself has a truly difficult time in contemporary society' (p.25). The self is often credited with being the key to the understanding of human behaviour. Nash (1976), for example states 'what a person believes about himself will form part of his motivational structure. That is to say that a person will choose to act or not to act in a given way because he regards himself as the sort of person who does or does not act in that way' (p.8). Such a belief in the power of the self is not uncommon.

Yet, on the other hand, empirical research in this area has only too often been haphazard, devoid of theoretical content, and based upon inadequately tested assumptions. It is one aim of the present thesis to examine empirically three such assumptions. These assumptions will be described later in this chapter, once a definition of the self-concept and SE has been arrived at. As will be seen, these three assumptions are shown to be unjustified. The findings of this thesis have profound implications for both theory and research, and these are discussed.

A second major aim of this thesis is to provide an adequate measure of SE. By 'adequate' is meant that the measure should be reliable enough for the purposes of basic research; should show evidence of validity; and should be meaningful and relevant to the people whose SE it is supposed to measure. It is argued that to measure SE appropriately, the measure must possess certain characteristics. The type of method chosen (a 'discrepancy' measure) has been heavily criticised by a variety of authors (including Wylie, 1974; and Burns, 1979). The criticisms are, however, either open to empirical test or concern aspects of methodology common to all questionnaires (eg, how to select items). Data reported in this thesis suggests that the assumptions specific to discrepancy measures are empirically justified.
The third aim occupies the major part of this thesis. It is to make further progress towards understanding the role of SE in behaviour. There is a major controversy (e.g., Jones, 1973; Shrauger, 1975; Dipboye, 1977) concerning how best to explain people's reactions to evaluative feedback information. Both sides to the controversy assign a major role to SE, but each favours a different explanatory mechanism. Several original experiments are reported in this thesis which attempt to resolve the controversy. It is argued that much of the conflict is for three reasons. First, the theoretical bases to the research have not been carefully scrutinized. Second, inadequate research designs have been used. Third, there has been an appalling lack of attention given to the implications of using different dependent variables.

In the following sections of this chapter, definitions of the self-concept and SE are elaborated, and some of the most relevant issues and research are discussed. It must be pointed out here that it is not the intention in this chapter to provide a comprehensive review of the literature on the self-concept or self-esteem (such a review would require several volumes). Instead, important issues in this literature are highlighted and discussed.
1.2: Towards a Definition of the Self-Concept

It would seem fair to say that defining the self has not been an easy task for either psychologists or philosophers. There has been much concern over the issue of man as both the knower and the known, i.e., man does the knowing and yet is also the object of his knowing. At first glance this appears to produce a paradox: how am I ever to know myself when part of myself is the 'I'? It appears to lead to an infinite regress of self reflecting upon self - each one a more 'inner' self examining and evaluating the one before it; or an infinite number of homunculi situated somewhere in the head of a person and scrutinising each other. To define the self along the lines of 'all that is known about the self' is to fall prey to what seems to be an inescapable circularity - for in order to define the self we need to know what the self is. Such definitions include the ingenious solution offered by Allport (1955) who coined a new word, an effective way of removing any surplus meaning which might be attached to a word due to its use in everyday speech. He defined the proprium as 'all the regions of our life that we regard as particularly ours' (p.40). It is evident, however, that this solution does not work - because in order to identify the proprium we need to know what people regard as essentially theirs which in turn requires identification of the proprium. Similarly, the definition of James (1890) that the self is whatever the individual views as belonging to himself (including the material, social and spiritual selves); the definition by Rogers; that of Snygg and Combs (1949), who defined the self-concept as 'those parts of the phenomenal field which the individual has differentiated as definite and fairly stable characteristics of himself'; and others, do not avoid the circularity issue.

Indeed, many writers who are aware of the problem of 'self as knower and known' place it firmly out of bounds and leave it for the philosopher to deal with. It may be conjectured that this omission is more
for the reason that so far psychologists have been unable to conceptualize 'self as knower' rather than because they thought that this aspect of human functioning would not be a useful one for the understanding of human behaviour. Mead, for example, says 'I do not mean to raise the metaphysical question of how a person can be both "I" and "me"' (1974 eq. p.173) and Harré (1977) regards the 'I'-'me' debate as 'arbitrary and intractible' within the realm of metaphysics, and opts for the solution of accounting for the self-concept in terms of the ways in which we learn to distribute personal pronouns in our speech.

However, perhaps this problem of the self as both the 'knower' and the 'known' is a pseudo-issue. Alston (1977) is rather puzzled as to why psychologists have been so concerned over the 'homunculus' issue, and feels that their concern has been somewhat misguided, and based upon a misconception of the self-concept. He, too, draws attention to the inherent circularity of various attempts by psychologists to define the self. He says that it is misleading to suggest that the 'self-as knower' and 'self-as-known' are two different entities, and that 'in over-reaction to the inner agent, one has shied away from any agent and left the thoughts without a thinker' (1977, p.67). He points out that in everyday life there are no problems associated with the concept of self: when a person praises 'himself' the object of his praise is the same as it would be if I were praising this person. This seems sensible, even though it must be recognised that the 'same' object can appear very different to different people. Alston goes on to say: 'The 'self' of whom we speak in everyday discourse is in no way to be distinguished from the individual man and is no more problematical. If we know who (what) we are talking about when we use a proper name or other designation of a human being, we are in possession of all we need to talk about a self ... There is no need to postulate any mysterious inner agent in order to attribute the various inner and outer doings, functions and characteristics of a person to
something. We simply attribute them to the same familiar agent we are accustomed to attribute them to in daily life ... By maintaining our hold on the everyday conceptual scheme, we will avoid conceptual quagmires and get on to the real problems (eg) what are the antecedents and consequences of positive and negative self-esteem' (p.68). Alston's solution at first appears a simple one - we merely ignore the issue. It causes no undue discomfort in our everyday lives to speak of the self, therefore why should social scientists bother themselves further with pseudo-philosophical wrangles? However, Alston is saying more than this - he suggests we do not regard the 'knower' and the 'known' as two distinct entities. This implies that the problem is a semantic one, imposed by the way in which the problem has been formulated. But if this solution is to be more than a convenient semantic twist, what is needed is a model of man within which the problem of man as both the 'knower' and the 'known' does not arise.

In many respects, the issue of 'homunculi' within the study of the self is similar to the issue of homunculi of the 'New Look' theorists in the area of perception. These theorists were concerned with the influences of psychodynamic events (such as expectancies, needs, and defence-systems) upon the perception of external stimuli. Here the apparent logical contradiction concerned both perceptual defence (ie. the idea that the perception of stimuli may be inhibited as a function of the input's emotionality) and perceptual vigilance (ie. the enhancing of a stimulus due to its emotionality). Given that these phenomena do occur (Erdelyi, 1974, presents evidence that they do) they appear to give rise to the following paradox: 'if perceptual defence is really perceptual, how can the perceiver selectively defend himself against a particular stimulus unless he first perceives the stimulus against which he should defend himself?' (Erdelyi, 1974, p.3). As Erdelyi points out, the issue was created by the way in which the question was posed: he goes on to
say that the posing of the question in the above form makes several assumptions which may well turn out to be false. Firstly, that perception is a single process. Secondly, that the flow of information is unidirectional from the perceptual system to the response system and, thirdly, that there are only two possible loci in which the phenomena of perceptual defence and perceptual vigilance may occur: the perceptual system and the response system. As soon as one recognises a multi-stage concept of information-processing, the paradox falls away and it becomes meaningful to ask whether only certain inputs are selected for further processing, ultimately to appear in consciousness. A system with control processes for internal regulation, including regulation of input and the responses made to it, 'violates no sacrosanct edict of science, nor does it imply the literal existence of little men and demons in the head' (Erdelyi, 1974, p.4).

Erdelyi's model will be briefly elaborated upon here in order to provide an example of the type of information-processing model into which the concept of self may be fitted without too much difficulty, and without the generation of metaphysical problems of definition. First of all, Erdelyi makes it clear that he regards cognitive processing as consisting of a great many systems through which information flows in a multi-dimensional manner, and is processed in a variety of ways.

Incoming information is first of all processed by the peripheral adjuncts and receptor systems (eg. eyes, ears) which in themselves are significant loci of selectivity (eg. we can close the eyes in order to shut out certain stimuli, and we can control the direction of looking). Then various storage systems are proposed, such as the 'afferent sensory storage system'; 'iconic storage'; an 'encoding system'; and a 'short-term-storage system' which is in close contact with a 'rehearsal and consolidation system'. This latter, according to the model, is closely linked to the 'long-term-memory' system. The long-term-memory system
has the capacity to issue control commands over all the other systems and so to determine the sequence of storage operations imposed upon the incoming information (eg. whether the input will result in conscious awareness). Finally, there is an 'output generator' which, in response to commands from long-term-memory, produces the required outputs. According to Erdelyi (1974), this model is only one of many possibilities available for explaining the required cognitive phenomena.

Two important points made by Erdelyi are: (a) that all the proposed systems can examine the contents of the other systems. This examination does not mean that the contents of the examined system are transferred to the examining system, or that a direct result of the examination is conscious awareness of the materials examined; and (b) that 'selectivity ... starts at the beginning and ends only at the very end of information-processing' (p.21). In other words, only some of the incoming information will be selected for further processing, and the selection can occur at any of the stages. The examination of one system by another may lead to disruption of ongoing storage operations. For example, the long-term-memory system may examine the afferent sensory storage system and prevent further processing of certain information held there. Erdelyi proposes that conscious awareness occurs in a processing region somewhere around the short-term-storage system - before long-term-memory but after the various encoding systems. He does not expand fully on the 'tricky problem of awareness', but makes the important point that a tremendous degree of in-depth analysis of inputs can occur without necessarily resulting in conscious awareness.

One may wonder why, if the systems can do all this analysis in order to select which information to process further, there should be any selection at all? Erdelyi proposes that we select because there is only a limited amount of storage space available. Therefore, the more carefully we select what we wish to store, the better. Indeed, the material
which is finally allowed to reach conscious awareness is a much reduced version of the information originally encoded. Hence 'consequentially, while the span of consciousness or conscious perception is small, it does not follow that the span of perceptual analysis is not vast.' (p.20).

Finally, Erdelyi equates selectivity with positive or negative biases in the way in which information is processed. These biases 'alter the probability of the material's being transferred to short-term storage, the region of consciousness'. It may be surmised that one set of biasing factors arises from the opinions a person holds about himself. This notion of SE as a biasing factor in the processing of information will be elaborated upon further in a later section.

How can the notion of self be fitted into this model? Erdelyi makes no reference to the concept, and it is not usually discussed in such a context. For the purposes of defining the self, the important feature in Erdelyi's model is that one system can examine the contents of another. Let us make the reasonable assumption that in the long-term-memory of an individual there is a vast store of information, laid down over the years, concerning the nature of that individual. The conscious part of the system can call up (ie. examine) bits of this reservoir at a time and make it accessible to awareness. If the labels 'I' or 'me' are attached to this information, then it can be defined as part of the self-concept. For example, the selection mechanisms in the processing system might bring into conscious awareness the information 'I am a woman'. It is suggested here that this piece of information is exactly equivalent in status to the piece of information 'You are a woman' or 'X is a woman'. This is the argument of Alston (1977) previously mentioned. That is, information with the pronouns 'I' or 'me' attached to it is, in terms of the processing mechanisms involved in its selection from long-term-storage, qualitatively identical to information with any other object-as-subject attached to it.
The self-concept can therefore be defined as: all the information held in the cognitive processing system which has the labels 'I' or 'me' attached to it, and which is potentially available to conscious awareness.

After this consideration of Erdelyi's model and its implications for the self-concept, it can be seen that the problems of circularity in the definition of the self-concept, the infinite regress of self scrutinising self, and the homunculi issue, all disappear because an appropriately sophisticated view of the cognitive processing capacities of human-beings is used. There is no reason whatsoever why one part of the system should not reflect upon another part, and this in no way implies a 'fleeing self'. Within this context the interesting question of how criteria for selectivity are developed becomes meaningful.

The real problem is not how to define the self, but to discover answers to questions like:

(a) How is the information with the labels 'I' or 'me' attached to it selected for storage? Or, in older-style parlance, how does the self-concept develop?

(b) What internal or external stimuli bring a certain piece of information about the individual into conscious awareness (given that the individual cannot hold in his limited span of consciousness all the information about himself at once)?

(c) What implications does this information have for behaviour?

It should be pointed out here that the information a person has about himself is processed in exactly the same multi-process system as any other information, but whether it is affected by the same biasing factors (i.e. criteria for selectivity) to the same extent is a matter for empirical research.

The definition proposed above of the self-concept is all very well, but can it be effectively operationalized so that the self-concept can
become a useful construct within the science of human behaviour?

According to Gergen (1977) the self does not possess sufficient 'factuality' for a science of the self to be possible - he suggests that a person can, through biased scanning, find confirming instances of any self-description he wishes. If this is so, and there are no limits on the self-characterizations a person will use to describe himself, then Gergen is probably correct and the 'possibilities for an exact science of mental life in general, and a science of self-experience in particular, seem limited' (1977, p.168). This statement, however, begs the question. And even if it is true that self-descriptions vary from one moment to another, it does not follow from this that there are no rules governing how they vary (see question 'b' above).

Many operational definitions of the self-concept spring to mind, though it is unlikely that any of them measure the self without error. For example, how can we bring to light all of the information a person possesses about himself, when he can only hold a small portion of it in awareness at any one point in time, and when he himself might not be clear about the criteria he uses for selection? It is even probable that the information itself is influenced by the questions asked about it. Suppose, for example, that the person is asked to write down as much information about himself as he can. This 'free-description' approach appears very direct, but as investigators we can never be sure that the individual has told us everything, or even that what he has told us is the truth. The individual may omit certain information because it would be embarrassing for him to reveal it, or he may just forget to mention something of vital importance. Jourard (1971) and his coworkers show clearly that how much, and what, a person is prepared to reveal about himself depends upon who is seeking the information. Wylie (1957) found that subjects who wrote open-ended essays missed out aspects of themselves which other investigators had found to be important. Whatever the operational definition
used, therefore, we shall have to accept its limitations. Further dis-
cussion of methodology and accompanying problems will be postponed for
a later section.

One final point about the definition of the self-concept used
here is that it makes no reference to self knowledge, only to information
the individual possesses. In accordance with the arguments of Hamlyn
(1977) this information is assigned the status of beliefs rather than
knowledge.
1.3: Towards a Definition of Self-Esteem, and Self-Assessments

(a) The definitions

There is a multitude of terms which appear to have some meaning in common in referring to a construct which seems to describe the extent to which the individual values (or otherwise) the information which he possesses about himself. These include terms such as self-esteem, positive and negative self-regard, high and low self-concept, self-evaluation and self-worth. It is not exactly clear whether these terms are synonymous, and their diversity only serves to create confusion and to make integration of the results of different studies difficult. Different definitions have produced a vast array of different measures some of which measure self-esteem as defined by one investigator but do not measure self-esteem as defined by a different investigator.

In the previous section, the self-concept was defined as the information held in the processing systems of the individual which has the labels 'I' or 'me' attached to it. Suppose a person describes himself as 'quite honest'. How does he arrive at this self-characterization? He must have some notion of what it is to be honest, as well as of what it is to be dishonest. That is, he must be aware that there is a scale of honesty. Placement of himself on this scale implies acceptance of various criteria (which could be implicit, or explicit; unique to that individual or common to all mankind), as well as the ability to apply these criteria to himself. When an individual tells a friend that he is quite honest, he is describing himself and providing his friend with an aspect of his self-concept. In a similar way, he is providing an investigator with an aspect of his self-concept when he endorses an item which reads 'I am not very confident'. The individual will be said to be making a self-assessment when he places himself at a particular point on a scale with
respect to a particular attribute.

However, it is possible to make another type of judgement apart from the simple one of where we would place ourselves upon a scale of 'intensity' with respect to an attribute. This second type of evaluation concerns the value we place upon our self-assessments. In the above example of an individual who regards himself as being 'quite honest', it is possible for him to say how satisfied he is with being 'quite honest'. Is it a good, or a bad, thing in the eyes of the individual to possess this degree of honesty? Obviously, the person may evaluate many of his self-assessments in this way, and the set of these evaluations is what is called self-esteem (SE). In a similar way to the self-concept, the SE of an individual as experienced at a particular moment in time will depend upon the values he places on the self-assessments brought into awareness.

In order to evaluate anything along the dimension of worth, we must (a) be aware of alternatives (if I am not aware that there is any state other than the state I am in, how can I be satisfied or dissatisfied with that state?), and (b) be able to compare a current state with a desired state (unless I can decide whether I want to be, for example, extremely honest, then how can I decide whether I am satisfied or dissatisfied with being only quite honest?). SE, therefore, involves an assessment of what we are like relative to what we are not like along the dimension of worth. Because worth is a term hardly used by psychologists, it may be better to talk instead of values (worth also implies moral excellence, according to the O.E.D.). Of course, what one person values may not be what another person values, and therefore when empirically assessing SE it is essential to allow the individual to express his own, idiosyncratic notions of what he regards as valuable (Rosenberg, 1968, argues along very similar lines).

It should be clear from the discussion so far that under the present definitions, simple self-descriptions are not equivalent to SE. Taking
the argument to its logical conclusion, an individual is making a self-assessment when answering the question 'How good am I?', but is engaging his SE when answering the question 'How satisfied am I that I am this good?' In other words, an adequate measuring device for SE should encompass the individual's satisfaction or dissatisfaction with his self-descriptions - without measuring the values the individual places on these we are measuring no more than the self-descriptions themselves. As Hamachek (1978) puts it: 'it is not simply how good we think we are with regard to some quality, but how good we want to be that counts' (p.263).

If an individual feels that he is quite honest, and he feels that this is how he wants to be, his SE for honesty will be high. On the other hand, if he thinks he is quite honest, but that it is desirable to be extremely honest, his SE for honesty will be low.

SE can therefore be defined as: the feeling(s) of personal satisfaction (or worth) which arise(s) when an individual compares what he believes himself to be like with what he feels he would like to be like, in relation to his self-assessments.

There are two consequences of this definition which should be briefly mentioned here. Firstly, measures which do not ask an individual to evaluate his self-assessments are not measuring SE and, secondly, that SE is theoretically unrelated to performance, or to any other outside criterial assessment of the individual. It is not how well an individual performs which is an indicant of his SE, but how well he performs in relation to how well he would like to perform. It is perfectly possible, for example, for someone to achieve an extremely high mark in an examination and for the person to be very dissatisfied with this mark. It is only insofar as the 'objective' assessments of the person are positively related to his own value system that these objective assessments can, in any way, be considered an indicant of SE. Whether such a relationship exists must be discovered through empirical research.
(b) The effects of self-assessments and SE upon behaviour

There is not much point in making a theoretical distinction between concepts unless they are hypothesized (a) to exert different effects upon behaviour or (b) to interact in determining behaviour. The important distinguishing feature between SE and self-assessments, as conceptualized here, is that the former takes into account the values of the individual whereas the latter does not. Obviously, insofar as people in general (or within an experimental sample) hold the same values, SE and self-assessments will be highly correlated. Therefore, classification on the basis of self-assessments would be highly similar to classification on the basis of SE. The assumption that values are identical within an experimental sample is a popular, though implicit, assumption in much of the literature. For example, Simpson and Boyle (1975) say: 'situational self-esteem refers to evaluations of more restricted sets of behaviours in specific situations. This constraint can be conceptualized in one way as the expectations by the individual of his performance in a task-specific situation.' (p.898). Thus, these authors are equating 'SE' with 'expectations'. Values are not mentioned. Similarly, Diggory (1966) equates SE with the subjective probability of success on a given task. This also assumes that SE can be measured without taking values into account — in Diggory's case the assumption is made that everyone wishes to succeed to the same degree.

Nevertheless, there are probably some circumstances in which people do have highly similar values, and in these self-assessments will be a good indicant of SE. For example, it may be reasonable to assume that most college students wish to do well in their chosen courses. Presumably people who voluntarily enter competitions desire to win, and people who join a particular political party share many similar ideals. Whether or not self-assessments are valid indicants of SE, therefore, will depend
upon the sample used. However, there is certainly no lack of evidence to suggest that people's values do differ (e.g. Rosenberg, 1965; Rokeach, 1968; 1973). It seems to the present author intuitively unsatisfactory to assume that values are identical when measuring SE - particularly when there is no need to make such an assumption.

In what ways, therefore, are self-assessments and SE hypothesized to exert their own, independent, effects upon behaviour? Crucial to the following exposition is the assumption that to have LSE is distressing, because it is unpleasant to regard oneself as different from what one would like to be like. In other words, to have LSE is to possess a 'chronic sense of failure', because one is just not achieving one's goals. On the other hand, to have HSE is to have a 'chronic sense of success', because an HSE person does achieve his goals. It is being proposed here that this chronic sense of failure of LSE's, accompanied by the associated self-dissatisfaction, will produce effects upon behaviour different from the effects of their self-assessments alone. Similarly, for HSE's, their chronic sense of success will produce effects upon behaviour different from their self-assessments alone. It has already been pointed out that a person who has 'low' self-assessments may have HSE because he does not wish to be any different from what he is already. Similarly, a person with 'high' self-assessments may have HSE because he likes how he is. Though the self-assessments of these two people are very different, their SE level is highly similar. Thus in some ways they should behave similarly. A similar case may be made for people who have different self-assessments but who have LSE. That is, people with similar levels of SE should in some ways behave similarly - no matter what the actual self-assessments are. Some of these hypothesized independent effects of SE and self-assessments are documented below.
1. Influence of self-assessments upon goal-setting

Self-assessments may be considered as expectancies a person has about his behaviour, that is, as the subjective probability he attaches to his reaching a particular goal. For example, a person who evaluates himself as 'quite intelligent' would regard it as less probable that he would succeed in a moderately difficult intellectual task than would a person who evaluates himself as 'very intelligent'. There has been a great deal of research into 'level of aspiration' (i.e. goal-setting), which clearly demonstrates that the subjective probability of success will influence the goals chosen (see, for example, Escalona, 1940; Festinger, 1942; Lewin, Dembo, Festinger and Sears, 1944). Studies in this area, and in the strongly related area of achievement motivation, show that typical individuals will not choose goals where their subjective probability of success is either too low (because then they would never experience the delights of success), or too high (because then the task is too easy and success loses its challenge).

Unfortunately, in later work (e.g. Atkinson, 1964; Atkinson and Raynor, 1974) level of aspiration has become synonymous with probability of success, at least in the area of achievement motivation (see, for example, Weiner, 1980, p.198). This confuses the issue as far as the present author is concerned. Nevertheless, sufficient early work has been done to show that goal-setting does seem to be influenced by self-assessments.

2. Influence of SE upon selection of situations

One way in which SE rather than self-assessments is hypothesized to affect behaviour is its role in the choosing of situations by the individual. If an individual possesses LSE on a particular set of attributes engaged (i.e. brought into awareness) by a particular type of situation, then he will try to avoid such situations as far as possible. In addition, he will actively seek out those situations where he has HSE.
Theoretically, this choice of situations should be independent of self-assessments. Consider the following example: suppose Person A considers himself to be highly socially skilled in relation to his peers, but wishes strongly to be even more socially skilled. Person B, on the other hand, thinks she is about average in social skills compared to her peers and is perfectly satisfied with this level of skill. Given a choice, Person A should avoid social situations whereas Person B should not, because the self-dissatisfaction aroused in Person A by participating in such social situations is unpleasant whereas for Person B no such unpleasant feelings are aroused.

Obviously, however, people do not have a totally free choice over the situations in which they participate - people do not have much choice over their family or their workmates, for example. Social situations are difficult to avoid altogether, and most people have to go to work (or to school) where, often, there is little choice over the situations in which one engages. In addition, the definition of situations often depends upon how the other participants view the interaction. A person who is engaged in serious (but enjoyable) intellectual discussion with an opposite-sex companion may be extremely discomfitted when she learns that he views it as a flirtation. If her level of SE is high in intellectual situations, but low in heterosexual flirtations, she may well choose to end the encounter, or make strong efforts to change the definition of the situation.

The influence of SE level upon choice of situations has, sadly, not been the subject of much research. This may be because of the previously-mentioned lack of attention to situations by social psychologists. One area in which there is indirect evidence of such an effect comes from American research on school dropouts. In a four year study, Bachman, Green and Wirtanen (1972) found that the SE of dropouts rose once they had left school. Their SE was not significantly lower than the SE of those
who stayed on to graduate. Rosenberg (1957) examined occupational choice, and found that people tended to choose occupations where they thought they had a reasonable chance of doing well. In other words, they chose careers where their SE would be maintained or enhanced. In later work, Rosenberg (1965) demonstrated that there is a tendency for people to value what they are good at - again suggesting that, where there is a choice, people will choose so as to have HSE.

It must be stressed that the desire to maintain or enhance SE is probably not the only factor governing choice of situations. There are all sorts of pressures upon people to enter particular types of situations, no matter what their SE level is. For example, ambitious parents may push their offspring into going to university when they would otherwise not choose to go; one person might persuade her friend to go to a party, even though her friend has LSE in such situations. In such cases, competing desires win over the desire to maintain or enhance SE.

The influence of SE in the selection of situations will only be seen, therefore, when either: (a) the person has a completely free choice. That is: there are no other competing motives; or (b) when the competing motives are equally strong for all the situations from which the individual has to choose.

3. Influence of SE upon selection of comparison-others

To reiterate, to arrive at a self-assessment an individual needs to place himself on a scale with respect to a particular attribute. To arrive at this self-assessment, the individual may compare his behaviour with his evaluations of other people's behaviour. Therefore, such self-assessments will be largely determined by whom he chooses for comparison (the comparison-others). Who will such comparison-others be? It is hypothesized here that the individual will, given a free choice, choose comparison-others who allow him to have HSE. To illustrate this using
an extreme example, if an individual wishes to be very honest, he may choose as comparison-others those who are less honest than he is. By comparison with these people, therefore, he will appear very honest and thus have HSE for honesty. In other words, the person's placement of himself on a scale with reference to a particular attribute may depend upon the comparison-others he chooses.

Such an hypothesis is not an original one, of course. It is suggested by Festinger (1954) in his theory of social comparison processes. He postulates that people have a drive to evaluate themselves by comparing themselves with others, so that a valid self-assessment is derived. But he postulates further that the comparison-others will be chosen so that the individual is able to maintain or enhance his SE. As Billig (1976) says, when describing Festinger's (1954) theory, 'There will be a tendency for individuals to compare themselves to others of equal ability or of slightly lower ability. They will not compare themselves to those of infinitely greater ability than themselves as this would damage their own self-esteem.' (p.349). Unfortunately, however, people are not always totally free to choose their comparison-others. Morse and Gergen (1970) demonstrate the effects on self-assessments of particular constraints upon the availability of others for comparison.

Tajfel (1972) and Turner (1975) extend Festinger's social comparison theory to intergroup processes. Basically, they postulate that groups attempt to create and maintain a positive social identity. In order to have such a positive social identity, groups need to compare themselves with other groups and, further more, each group somehow has to come to the conclusion that it is 'better' than the others. Each does this by assigning values to the dimensions of comparison. If, by such means, positive group identity does not occur, then value-change will take place.

Thus another effect of SE upon behaviour concerns the selection of comparison-others. In this way, self-assessments may be arrived at in
order to allow the individual to maintain or enhance his SE. Taken in conjunction with the 'level of aspiration' literature (see page 22), it can be seen that SE and self-assessments influence each other: self-assessments play a major role in the selection of values; and values play a major role in the selection of situations and comparison-others so as to have self-assessments which are not too far removed from the goals of the individual.

It should be pointed out here, however, that individuals are not completely free to select their own self-assessments by a judicious choice of comparison-others. According to the Symbolic Interaction tradition in sociology (eg. Cooley, 1902; Mead, 1934; Shrauger and Schoeneman, 1979), self-assessments are determined by the evaluations of 'significant-others'. Highly critical 'significant-others' will produce self-assessments which are far removed from 'ideals', and thus will generate LSE (see, for example, Coopersmith, 1967, for an account of childrearing practices and SE among adolescent boys). Insofar as an individual internalises the assessments made of him by significant-others, and insofar as he internalises their values, his SE level will be externally determined.

4. The roles of SE and self-assessments in determining reactions to feedback

The second half of this thesis is concerned with an examination of how SE affects reactions to feedback, and therefore a full chapter (Chapter 5) is devoted to a detailed discussion of this issue. The main points are summarized here, however, in order to consider also the effects of self-assessments.

There are several theoretical positions from which predictions may be derived regarding the role of SE and self-assessments in determining reactions to feedback. One is called self-enhancement theory (eg. Dittes,
The basic postulate of this theory is that everyone has a need to think well of himself, and that if this need is not satisfied it becomes stronger. This is basically a drive-reduction model, and attempts to satisfy the need for self-enhancement are hypothesized to operate in ways conceptually similar to attempts to reduce, say, a hunger or thirst drive. LSE's (by definition) are in a chronic state of need deprivation, and therefore their attempts to be self-enhancing should be even stronger than are the attempts made by HSE's. This may have important effects on the processing of feedback information. For example, suppose LSE's and HSE's are both given identical, but ambiguous, feedback information concerning their performance on a task. According to self-enhancement theory, LSE's should think they have done better than HSE's on the basis of the ambiguous information, because LSE's should leap at any opportunity to think well of themselves.

A second theoretical notion of relevance here is what I have termed (in Chapter 5) 'esteem-protection' theory. This rests on two assumptions. Firstly, that people have a need to think well of themselves and that this need increases when it is not satisfied (this is exactly the same assumption as that made by self-enhancement theory) Secondly, that satisfaction of this need is tempered by a consideration of the costs in doing so. Such costs include, for example, the embarrassment or loss in esteem of not living up to one's publicly declared expectations. LSE's would, by definition, be extremely dissatisfied if they lost any more SE. This means that the costs of losing SE would be greater for LSE's than for HSE's, and therefore that (under certain circumstances) LSE's should be more willing to 'accept' criticism than should HSE's.

The third relevant theoretical position is that of the consistency theorists (eg. Festinger, 1957; Heider, 1958). Their basic assumption is that people have a need for consistency in their cognitive elements (eg. their beliefs and attitudes), and that inconsistency produces an
unpleasant state of tension which people will try to reduce. If we assume that the chronic sense of failure possessed by LSE's, and the chronic sense of success possessed by HSE's are cognitive elements, then information discrepant with these elements should be unpleasant. There are a variety of reactions to such discrepant information which could take place in order to reduce the inconsistency. For example, the information source could be discredited, or one of the inconsistent elements could be changed.

But, of course, self-assessments themselves have already been described as expectancies. Therefore these, too, may be regarded as cognitive elements. Now, insofar as self-assessments and SE are uncorrelated, a person who expects to do badly on a task will not necessarily have LSE. Similarly, a person who expects to do well on a task will not necessarily have HSE. This raises the interesting question of whether self-assessments or SE will exert the stronger influence in determining reactions to feedback. Insofar as self-assessments and SE are correlated, of course, their determining influences on reactions to feedback will be similar. In order to elucidate the differential effects of self-assessments and SE, however, let us assume for the moment that they are uncorrelated.

According to the self-enhancement notion, people whose need to think well of themselves is frustrated will have this need increased. If self-assessments are the important factor in determining reactions to feedback then people with a low self-assessment on the particular (task-relevant) attributes, should respond more 'favourably' to (eg. should like it more) positive feedback than should people with high self-assessments on these attributes, irrespective of their level of SE. If, on the other hand, SE is the important factor determining reactions to feedback, then LSE's should respond more 'favourably' to positive feedback than should HSE's.
A similar argument may be made for predictions derived from the esteem-protection and cognitive consistency theoretical positions. If self-assessments are stronger than SE level in determining reactions to feedback then, irrespective of level of SE, people with low self-assessments should (under certain circumstances) be more prepared to 'accept' criticism (e.g. agree with it, or regard it as valid) than should people with high self-assessments. The same prediction may be derived from the cognitive consistency theories.

One intention of this thesis is to examine the effects of SE upon reactions to feedback: in particular, to discover the limiting conditions of the theoretical positions outlined above. The separate effects of self-assessments are not examined: these will be investigated in later research. However, correlations between specific task-expectancies and SE are presented. As will be seen, such correlations are generally low, though positive. Therefore, insofar as results support the predictions, the effects of SE will have been demonstrated.

A useful research strategy for an investigator wishing to examine the separate effects of SE and self-assessments would be to use a 2 x 2 x 2 ANOVA design. The first factor would be SE level (low vs high), the second factor would be self-assessments on task-relevant attributes (low vs high); and the third factor would be feedback (positive vs negative). Eight groups of subjects would be needed. The ANOVA framework would allow any main effects (of SE; self-assessments; or feedback) to be revealed, as well as any interactions between them.

In sum, SE is proposed to act as a 'locus of selectivity' (after Erdelyi, 1974) in the processing of feedback information because people with LSE possess a chronic sense of failure, and people with HSE possess a chronic sense of success. On the other hand, self-assessments are also expected to influence reactions to feedback. The present thesis focusses on SE, though self-assessments are not totally ignored. Two important
issues are not examined in this thesis. These are: (i) the relative strengths of the effects of SE as opposed to self-assessments in determining reactions to feedback; and (ii) the circumstances under which the effects of SE or self-assessments would be expected to be strong or weak. The following research on SE and reactions to feedback is based upon the assumption that a person's chronic sense of success (if he is an HSE) or of failure (if he is an LSE) will be strongly engaged in a self-relevant feedback situation.
1.4: Predominant assumptions made about the self-concept, and SE

Several assumptions are nearly always made in connection with self-concept and self-esteem research which are (a) not necessarily theoretically derived and (b) inadequately (if at all) tested empirically. The present thesis aims to examine some of these assumptions, to test their justification. Obviously, it is difficult to know quite what assumptions are being made if definitions are ambiguous and theoretical statements regarding, for example, the variables which are supposed to interact with SE (or the self-concept) are lacking. However, a close examination of current research reveals, with a few notable exceptions (eg. Gergen, 1971), that the following assumptions are made by empirical researchers in this area:

i) SE and the self-concept are assumed to be personality traits which remain stable from one situation to another. For example, a person has a particular level of SE which characterises him in all of his everyday situations. Within the theoretical notions outlined earlier (and, indeed, within many theories of the self, eg. James, 1890) such a view of SE is untenable. The self-concept, it is here proposed, varies with the addition of new information into the processing systems of the individual; and those aspects of the self-concept available to awareness vary across situations. Because different personal attributes are relevant in different situations, and because comparison-others (ie. the people chosen for comparison) may vary from one situation to another, it is perfectly possible for the self-assessments and SE to vary both in content (ie. the attributes focussed upon) and in level across situations. The assumption of lack of variability in the self-concept and SE is such an important one that it will be discussed in its own section below.

ii) It is not necessary to specify a reference group with whom subjects should compare themselves when making their self-assessments.
In other words, comparison-others do not vary across situations or, if they do, this has no consequences for the resulting self-assessments or SE. The importance of comparison-others has long been recognised (eg. Festinger, 1954), but there appears to be little work examining the hypothesis that comparison-others vary from one situation to another. This hypothesis is examined in the first study to be reported here.

iii) Every item in a questionnaire (or all items endorsed by subjects in those questionnaires allowing them choice such as the adjective check list: Gough, 1960) is equally important, relevant and salient to all subjects, and means the same to them all. This, however, may not be so. For example, Oswalt (1974) found a striking lack of agreement between investor-determined rating scales and the actual words used by subjects to describe other people. The two lists only overlapped to the degree of 12%. Loehlin (1961, 1967) has found that self-descriptive words do not mean (in terms of Osgood, Suci, and Tannenbaum's 1957 semantic-differential dimensions) the same to all subjects, and that this may affect whether or not they use a particular word to describe themselves. If words do not mean, at least approximately, the same to different people, and if the items are not equally relevant, then significant sources of error may be introduced into the data collected. The two problems of meaning and of relevance are tackled in the following ways in this thesis.

a) It is assumed that if there is no consensus as to the meaning of words, human verbal communication would be impossible. It is recognised that different populations of individuals may use different words, or attach different meanings to the same words. Hence, steps were taken to use as items in a questionnaire frequent words and phrases elicited from a sample drawn from the same populations as the subjects. After taking this precaution, a certain degree of consensus in meaning was assumed.
b) The relevance issue was tackled by allowing subjects to specify how relevant they thought a particular item was for a particular situation. Therefore, different aspects of the self-concept could be examined for their relevance to different situations and the degree of subject variability in the relevance of items could be measured.

iv) In discrepancy measures (such as the semantic differential), the assumption is made that the greater the discrepancy, the greater the dissatisfaction associated with it - no matter what the item. But it may be the case that a discrepancy of a particular size is associated with more or less dissatisfaction depending upon how important the particular attribute is to the individual. For example, a discrepancy of six points between self and ideal on the attribute 'I am friendly' may be associated with a great deal of dissatisfaction for a particular individual, but a discrepancy of the same size on the attribute 'I am careful' may be associated with hardly any dissatisfaction at all.

Thus four variables are involved in the measurement of SE using discrepancy measures. These are: (a) SE (b) dissatisfaction (c) the size of the discrepancy and (d) the importance of the item. To assume that the SE score is a good indicator of SE involves the additional assumption that dissatisfaction increases as the size of the discrepancy increases, irrespective of the importance of the item. This additional assumption is tested in Chapter 4 of this thesis.
1.5: Are self-assessments and SE personality traits?

(a) The current status of personality traits

It can readily be ascertained, from even a cursory examination of the literature in the area of SE, that this construct is usually regarded as a personality trait. Traits have been variously defined (eg. Cattell, 1965; Allport, 1966) and, as Argyle and Little (1972) point out, how personality traits are conceived will have a major bearing upon both the research hypotheses investigated and the ways in which they are studied. Few theorists would quarrel with Mischel's (1973) definition, however, that 'traits are relatively stable, highly consistent attributes that exert widely generalised causal effects on behaviour' (p.253). In other words, they are attributes ascribed to persons rather than to samples of his behaviour (Argyle and Little, 1972). That SE is a highly consistent attribute of a person has hardly ever been questioned (though Gergen, 1971, 1977; appears to be a notable exception). Consequently, in the instructions given to subjects the situations to which their self-ratings are supposed to apply are never specified. It is therefore possible that subjects are referring to different situations when making their self-ratings. This could have undesirable consequences if the assumption that SE does not vary across situations is false. For example, it could result in the lowering of the relationship between SE and some other variable being studied. The trait assumption also implies that people can be labelled as having LSE or HSE without the need for qualifying statements.

The lack of attention paid to situational variables in this area characterises the type of study which looks for relationships between SE and some other variable such as ratings of peers (eg. Miyamoto and Dornbusch, 1956), romantic love (Dion and Dion, 1975), parents' self-concepts (eg. Gecas, Calonico, and Thomas, 1974), susceptibility to

Despite the fact that the bulk of the literature in the area accepts uncritically the assumption that SE is an unvarying attribute of persons, there is also the widespread belief that a person's SE may change over time - for example, as he gets older (eg. Katz and Zigler, 1967; Hess and Bradshaw, 1970; Kaplan and Pokorny, 1970), or as he receives treatment for psychiatric problems (eg. Rogers and Dymond, 1954; Laxer, 1964; Hartlage and Hale, 1968; Harrow, Fox, Markhus, Stillman and Hallowell, 1968), or is given blind mobility training (Delafield, 1975). However, the systematic study of how the self-concept and/or SE may vary according to the situation as well as to the person has been lacking.

As Mischel (1968) points out, after an extensive review of the research on personality traits, the utility of a trait conceptualization of personality is open to question - if psychologists wish to predict behaviour accurately, validity coefficients in the region of 0.25 to 0.40 will not be much help. Yet these exemplify the magnitude of correlations obtained when personality measures are related to other variables. He showed that only in the sphere of cognitive variables (such as intelligence and problem-solving strategies) is there any adequate evidence of a reasonable degree of cross-situational consistency. Traits such as
conformity, dependency, aggression, and attitudes to authority showed substantial situational specificity. 

His critique of trait theory, and his emphasis upon the importance of situations in determining behaviour, led some theorists to believe that he was saying that there were no consistencies in behaviour, that individual differences are unimportant, and that situations are the main determinants of behaviour (eg. Alker, 1972; Bowers, 1973), but, as Mischel himself (1973), and Bem (1972) point out, this represents a gross misunderstanding of his position. The misunderstanding arises from the misleadingly posed question 'is behaviour due to situations or to persons?' As Endler (1973) puts it 'asking whether behavioural variance is due to either situations or to persons ... is analogous to asking whether air or blood is more essential to life, or asking one to define the area of a rectangle in terms of length or width. The more sensible question is "How do individual differences and situations interact in evoking behaviour?" (p.289). Alker (1972) accuses Mischel, among other things, of ignoring the moderator variable approach, and the interactionist position. Yet, as Bem (1972) points out, Mischel ignored neither. A moderator variable is 'any variable which affects the relationship between two other variables.' (Bem, 1972, p.20). For example, SE affects the relationship between authoritarianism and influencability, such that it is positive for those with HSE and negative for those with LSE (Berkowitz and Lundy, 1957). The moderator variable approach consists of dividing a population into sub-groups on the basis of hypothesized moderating variables, and then examining the correlations between the other variables within each subgroup. If the population were not subdivided the relationships discovered by taking moderator variables into account would be masked.

Mischel (1968) does mention the above strategy, but is rather unhappy with it. In a later paper (Mischel, 1973), he points out that the number of moderator variables required to predict behaviour and the
complexity of their inter-relationships would be enormous, and that 'the
more moderators required to qualify a trait, the more the "trait" becomes
a relatively specific description of a behaviour-situation unit.' (p.257).
In addition, to have scientific predictive utility, moderators need to
be specified in advance, rather than applied post-hoc, as is so often
the case. Even proponents of the moderator-variable strategy recognise
its severe limitations. Wallach and Leggett (1972), for example, say
'The empirical basis for recommending moderators as the answer to the
search for consistency thus seems more apparent than real.' (p.313).

If the moderator variable approach is not the answer to examining
the trans situational consistency of traits, what is?

Argyle and Little (1972) point out that there are at least four
types of variability in behaviour. One is that behaviour remains constant
for each individual in every situation, regardless of its nature. This
is an extreme view, and assumes that all variability in behaviour is due
to differences between people. No-one who is a human being could ratio-
nally endorse this view, for it is an undisputed fact of life that beha-
viour does vary from situation to situation. A second type of variability
is that which arises when all behaviour is totally determined by the
situation. This implies that there are no individual differences in
behaviour: ie. within a situation, everyone behaves in exactly the same
way. This view is called situationism, of which Mischel (1968) was
accused, but, again, is one which few would endorse.

A third type is called 'dispositional' by Argyle and Little (1972),
and assumes that the rank order of individuals stays the same across situa-
tions, but that the absolute 'amount' of a given behaviour may vary, as
may the extent to which individuals change their behaviour across situa-
tions (as long as the rank order stays the same). This is the view of
behaviour said to characterise the trait position (eg. Cattell, 1965).
Within this model, it is legitimate to examine situations for measurable
dimensions (eg. how stressful they are) which will account for the variations in behaviour shown. However, there should be consistency in the sense that the correlations across situations should be high. Mischel's (1968) review shows that these correlations are generally small, and on the basis of this evidence, he suggests we should reconceptualise personality (1973, 1977a, 1977b).

The fourth kind of variability mentioned by Argyle and Little (1972) is the so-called interaction type. The model is that there are individual differences in the ways people vary across situations. For example, one person might not change his aggressive behaviour at all across a sample of situations; another might show increasing aggression across them whilst a third might show decreasing aggression across them. In this model, correlations across situations would therefore be low when many people are measured, but may well be quite high for a particular individual.

A third approach (apart from the moderator variable and correlational strategies) has been used to examine the utility of traits, that of analysis of variance (eg. Rausch, Dittman and Taylor, 1959(b); Moos, 1968; Endler and Hunt, 1966, 1968, 1969; Argyle and Little, 1972; Bishop and Witt, 1970; Trickett and Moos, 1970). This strategy partitions the total variance into the percentage attributable to persons, situations, and their interactions. It was pioneered by Endler and Hunt (1966) in their study of self-reported symptoms of anxiety experienced in different situations. Trait theory would suggest that the overwhelming proportion of variance would be due to persons, situationism that most of the variance would be due to the situations, and the interactionist view that most of the variance would be attributable to the interaction terms in the analysis of variance equations.

Bowers (1973) in a review of 11 such studies found that the mean proportion of variance attributable to persons was 12.71%, to situations 10.17%, and to the interactions 20.77%. In 14 out of 18 comparisons
reported by Bowers, the variance attributable to the person x situation (P x S) interaction terms is greater than for the person and situation terms. Bowers concludes 'to some considerable extent the person and the situation are codeterminers of behaviour, and they need to be specified simultaneously if predictive accuracy is required.' (p.322). This conclusion is convincing, because the studies mentioned use three different types of data namely: self-reports based upon past experience in certain situations (eg. those by Endler and Hunt on anxiety, 1966, 1968, 1969); self-reports after participation in certain situations (eg. Moos, 1968), and observations of actual behaviour in a number of situations (eg. Moos, 1968; Nelson et al, 1969).

In the Argyle and Little (1972) study, for example, subjects were asked to imagine that they were in the presence of others who had pre-defined role relationships to the subjects (eg. boyfriend, boss, liked neighbour, less liked coworker), and to rate their behaviour when in the presence of these people on 18 bipolar scales. In this study, person variance (16.1%) was considerably less than the variance attributable to situations (43.6%) or to P x S interaction (40.2%). Another important finding in this study was that the bipolar scales themselves varied in the percentage variance attributable to the three sources. For example, 'discuss personal problems ... not at all' showed 12.1% variance due to persons, 52.7% to situations, and 35.1% due to P x S interaction, whereas 'take great care with personal appearance etc. ... less' showed 37.8% variance due to persons, 19.2% to situations and 43.0% to P x S interaction.

Moos (1968) asked 30 patients with psychiatric disorders in a Veterans Administration Hospital, and 10 members of staff, to describe their reactions on 33 bipolar scales in a variety of different situations - a total of 9 were analysed (such as getting up in the morning, having lunch, small group therapy). The interesting finding emerged that, for staff, the percentage variance attributable to persons (mean = 6.78%)
was much less than it was for the patients (mean = 29.36%), whereas the P x S interaction variance was greater for the staff (mean = 25.40%) than for the patients (mean = 16.95%). For the staff, the situations variance was also generally less than the persons variance, but this was not so for the patients. However, as in the Argyle and Little (1972) study, these averages conceal large variations in individual scores. These findings are also consistent with those of Rausch et al (1959a, 1959b, 1960) who found that both person and setting variance contributed significantly to overall behavioural variance, but that normal children showed greater differences between situations than did disturbed children. These results have sometimes been taken to support the notion (eg. Mischel, 1973; 1977) that inability to differentiate between situations in terms of the responses required in them may be a sign of psychological dysfunctioning. This may sound reasonable, but it needs the support of more than the above studies. The results of Moos (1968), for example, could merely reflect the greater variability of situations experienced by the staff (who were not always on the wards) compared to the institutionalised patients.

Though these studies have considerable theoretical importance for demonstrating that situations as well as persons should be studied, the analysis of variance approach is not without its limitations. The first is that the P x S interaction variance is often confounded with error variance, because there is usually only one data point in each P x S x Response cell. This means that the P x S interaction variance is nearly always overestimated (eg. Epstein, 1977, Golding, 1975; Argyle and Little, 1972). Secondly, the analysis of variance paradigm only reflects the degree to which individuals differ in absolute 'amounts' across situations. It does not indicate whether, for example, the rank order of individuals stays the same across situations. It is therefore possible to have a high percentage of variance due to, say, situations but for the rank order
of every person to stay exactly the same in each one. This is obviously a type of consistency which would support the trait model, and yet high situational variance is usually taken as evidence against it. Golding (1975) argues that the coefficient of generalisability (Cronbach et al, 1972) is a more appropriate index of consistency, for it indicates the extent to which one can generalise from an individual's observed scores in a set of situations, drawn randomly (which they often are not) from a universe of situations, to his mean score across all situations in the universe. This index therefore provides a different interpretative framework from the analysis of variance paradigm, and one which could, in principle, result in very different conclusions. (In practice, however, the conclusions do not seem to differ markedly – see Golding, 1975). Its main disadvantage is that it is horrendously complex to compute – Golding, for example, does not report the coefficients for the interactions, for this reason.

If both differences in 'amount' (obtainable via the analysis of variance approach), and of rank order (obtainable by the simple correlation coefficient) are required, the information could be obtained by using both the ANOVA and the correlational approach together.

Another problem with the analysis of variance paradigm (though by no means unique to it) is that virtually any point of view can be supported by a judicious selection of variables. Sarason, Smith, and Diener (1975), in a review of over 300 studies, found different results from those of Bowers (1973). The variance in this latter survey due to P x S interaction was a mere 4.7% (as compared to the Bowers figure of 20.77%). As they point out 'the figures presented in this article are a composite representing studies in which the independent variables vary widely along a continuum of theoretical meaningfulness' (p.203) and suggest that this is one reason why their figure differs so markedly from that of Bowers. But if percentages of variance differ so much depending upon the exact
variables studied, what is needed is replication of studies to find out how stable the findings are. Alas, this has not often been done. It also suggests that person-situation interaction may be exceedingly complex, perhaps even that no general rules can be drawn across situational domains.

This relates to the final problem (discussed here) of the analysis of variance approach: it is merely descriptive. It does not explain how person factors and situational factors interact in producing behaviour. As Nelson (1977) puts it 'the findings do not provide a basis for predicting or explaining which personality factors and situational factors are relevant or how they interact' (p.109).

In spite of the above powerful problems, the analysis of variance technique remains useful for demonstrating the point that situational and P x S interaction variance, as well as person variance, may have significant effects upon behaviour. In conjunction with the correlational approach, it can be used to examine the 'consistency' of behaviour across situations. This examination, it appears to me, is vitally necessary in the area of self-esteem - particularly in the light of the assumption (noted by Wylie, 1974) held by most investigators in this field that the self-esteem of an individual does remain constant across situations.

If the notion of a trait as being an unvarying characteristic of a person does not hold up under scrutiny of the analysis of variance approach, why should it be so common for individuals to use trait terms when describing themselves and others? Mischel (1968) goes so far as to say that traits are attributes of the perceiver rather than of the perceived, a conclusion Jones and Nisbett (1971) endorse. They cite evidence that actors and observers have differing perceptions of the causes of behaviour: actors tend to attribute their behaviour to situational factors, whereas observers tend to attribute the actor's behaviour to dispositional factors (of course, psychologists are well known observers of people). Why should
this difference occur?

As various authors note (e.g. Mischel, 1968; 1973; Jones and Nisbett, 1971, Argyle and Little, 1972) we normally only meet other people in a very limited set of situations—e.g. at home, at work, or in the pub. The limited information we gain about others from these restricted sets of circumstances may be used to form personality impressions of the others, in which we generalise to traits. Another reason is that we ourselves may tend to evoke certain responses in others; another is that human beings can only handle a certain amount of information at once, and traits may therefore be adopted as summary labels, despite their lack of heuristic value. Bem and Allen (1974) suggest that our implicit personality theories may cause us to perceive relationships which are not there, and to fill in the gaps of our knowledge about others with consistent data we have manufactured ourselves. Gergen (1968) also points out that there is a cultural ethic which endorses consistency, but sanctions inconsistency. This may lead us to expect others to behave predictably, and to infer that they are doing so when we really do not know. We may also find validating evidence for consistency in cognitive variables such as intelligence and hence infer consistency in non-cognitive variables such as personality traits. Another possibility, noted by Snyder and Swann (1978), is that an observer may bring about the behaviour he expects in another person, i.e. behaviour which confirms his ideas about the characteristics of that person. If he expects the person to be aggressive and extroverted, he may only present opportunities for that person to appear aggressive and extroverted. In a series of studies these authors demonstrated convincingly that subjects asked to find out whether or not another was extroverted asked the other questions which would allow him to appear extroverted (and subsequently rated the other as being extroverted), whereas subjects asked to find out whether or not a person was introverted asked mainly questions which would allow the person to demonstrate his introverted
nature (and hence subjects later rated the person as being introverted). Finally, our language is rich in trait terms, but has only poor facilities for discussing situations.

But do traits really lack heuristic value? Should we discard traits altogether, and reconceptualise personality? It certainly seems evident from the research discussed above that, at the very least, we should pay much more attention to situational variables, and the way these interact with personality factors. But a number of investigators have reported studies which suggest that traits do have adequate construct validity, and are useful predictors of human behaviour (Epstein, 1977; Bem and Allen, 1974; Allport, 1966; Gormly and Edelberg, 1974; McGowan and Gormly, 1976). Others have presented theoretical critiques of the 'anti-trait' position, along with supporting evidence (e.g. Hogan, De Soto and Solano, 1977; Stagner, 1976; Block, 1976; 1977). It is all-important to bear in mind, throughout this debate, that to a large extent the utility of traits will depend upon the exact traits being studied. All too often it is not made clear what behaviours the investigator would wish to predict.

Several authors (e.g. McGowan and Gormly, 1976; Epstein, 1977) argue that personality traits are not predictive of highly specific acts encountered in a single situation, but are predictive of a person's behaviour averaged over several inter-correlated acts. They both present data to support this view. For example, Epstein (1977) asked subjects to keep a record of their most positive and negative emotional experience each day for 24-34 days. Subjects described the incidents in their own words, and rated their emotions and behaviours on separate lists of adjectives. These lists were grouped into scales according to a factor analysis. The mean scores of the subjects on the scales were then correlated for the first two days, yielding reliability coefficients of around 0.20. Epstein then went on to correlate mean scores for all the odd days with those for
all the even days. He found that the reliability coefficients rose dramatically as the number of observations increased - reaching a mean correlation for the positive emotions of 0.88. It is worth noting here, however, that these coefficients may be overestimated because they are based on means. Nevertheless, his demonstration is useful for it shows that the more repeated observations there are, the more reliable traits become. Epstein also provides evidence that some personality variables are intrinsically more reliable than others, and that within-subject personality patterning becomes increasingly more reliable the more repeated observations there are. These findings would lead us to expect that, given enough observations, we would be able to predict an individual's averaged behaviour very well, though not necessarily a particular behaviour in any one given situation. This is possibly what people do when describing themselves and others - behaviour varies, perhaps, around an overall mean which is different from person to person. As Epstein (1977) himself points out, information about averages alone will not be sufficient to predict behaviour in a specific situation.

The study by McGowan and Gormly (1976) provides evidence that energetic behaviour is reasonably consistent across a variety of situations. They examined such behaviours as speed of going upstairs, errors on complex mazes, and the longest distance run by subjects in a day, and obtained correlations of more than 0.70 between the averaged behaviours and the trait score. Correlations between specific behaviours and the trait score were much lower (around 0.40), as was the agreement between the experimenter's judgements of energetic behaviour and peers' ratings.

In another study (Gormly and Edelberg, 1974), the validity of the trait of aggression is examined. Peer ratings were used to judge the aggressiveness of individuals, and pairs of individuals (one of whom was high and the other low on aggression according to peer ratings) were observed by non-professional judges who rated which member of the pair
was more aggressive on each of three occasions. Though the peer-ratings and the judges showed an average of around 70% agreement, the self-ratings and judges' assessments correlated only 0.40. The authors conclude 'the results of this study provide strong evidence for the position that social aggressiveness can accurately be considered a personality trait.' (p. 192). However, their results must be interpreted with caution, because the judges were only required to make very simple, temporally close judgements regarding a pair of people. Their study would have been a much more convincing demonstration of trait validity had the judges to rank, say 5 or 6 people on aggressiveness, and predicted their behaviour in a future situation which was then examined by another group of judges. The fact that they found relatively low agreement between peer-ratings and self-ratings is interesting; it could be taken as further support for the actor-observer discrepancy.

Bem and Allen (1974) reject the nomothetic viewpoint altogether, saying it is a fallacy to believe that 'a particular trait dimension or set of trait dimensions is universally applicable to all persons and that individual differences are to be identified with different locations on those dimensions' (p. 509). Instead, they adopt Allport's (1937) view that individuals differ in the ways in which traits are related to one another, and they also differ in which traits are even relevant. They posit that inconsistency has to be seen from the individual's point of view: 'It is only when we fail to discover a set of rationally-scaled equivalence classes which conform to the individual's behaviour that a judgement of "inconsistency" is finally rendered' (p. 510).

They present data to illustrate this point. Individuals rated themselves on various traits such as friendliness and conscientiousness, and also indicated how consistent they were on these traits. Ratings of the subjects were then obtained, on the above two traits, from the subject's mother, father, a peer, and a judge who observed him in a group discussion
and a waiting room. Altogether there were six ratings of the individual—presumably these referred to different sets of situations. For friendliness, subjects who said they were consistent were rated more similarly by the raters than were subjects who said they were inconsistent. For conscientiousness, however, self-ratings did not produce this result; but a partition of subjects on the basis of a questionnaire measuring consistency of conscientiousness did—inconsistent subjects were rated less similarly than consistent subjects. Bem and Allen suggest that this result was because subjects' conceptions of conscientiousness differed from that of the investigator's, whereas for friendliness they agreed. But if this is so, why should the raters of the subjects agree with the investigators? They conclude 'It is not possible, in principle, to do any better than predicting some of the people some of the time' (p. 517).

Snyder and Monson (1975) and Snyder and Tanke (1976) have also demonstrated that some people are more consistent than others. An interesting study which also makes this point is by Magnusson and Ekehammar (1975). They found, using Latent Profile Analysis (Mardberg, 1972; 1973), that subjects could be divided into homogeneous subgroups on the basis of their anxiety profiles. Three types of profile emerged, which differed slightly for males and females. For example, for females, the first profile was characterized by a low anxiety level, high transitiuational consistency, and high response consistency (59% of the subjects); the second by high anxiety level, high transitiuational consistency, and low response consistency (17% of the subjects); and the third by a moderate anxiety level, low transitiuational consistency, and low response consistency (24% of the subjects). This is a promising approach, and goes some way towards examining the patterns of person-situation interaction which may occur.

It is apparent from the foregoing discussion that the status of traits in general is still an open question (see Magnusson and Endler, 1977; Krauskopf, 1978; Endler and Magnusson, 1978), though the debate is
not new (Ekehammer, 1974). Current issues concern the ways in which situations should be conceptualised and classified (eg. Pervin, 1975(a), 1975(b); Forgas, 1976; 1979; Magnusson, 1971; 1974; Frederiksen, 1972; Krause, 1970; Sells, 1963; amongst others); and whether personality needs to be reconceptualised (eg. the social learning approach of Mischel, 1973; the generative rules model of Argyle, 1975). So far two main conclusions seem to have emerged. The first is that personality variables vary in the extent to which they are cross-situationally consistent. The second is that people also vary in this way. This implies that some understanding of the ways in which people interact with situations will increase our predictive powers concerning human behaviour.
The work briefly reviewed in the preceding section has major implications for the study of SE, providing the assumption is made that SE is an important determiner of behaviour. In this section, such implications will be discussed. As pointed out earlier, the assumption that SE does not vary across situations characterises the great bulk of empirical research in this area (as opposed to theorising about the self). As Gergen (1977) puts it: 'it is commonly held, within the mental tradition and by the public more generally, that the individual develops over time (and typically within the first year of life) an essential mode of self-evaluation. He comes to attach a basic value to his identity, which may vary from extremely positive ... to extremely negative ... It is generally held that the individual carries such evaluative sets with him across time and circumstance.' (p.145-146).

Such an assumption means, as Wylie (1974) points out, that neither the situation nor the reference group are specified when research is being carried out on SE. Yet SE may vary in many ways from one situation to another, including content, level, salience, and consistency. If the situation is unspecified, then different subjects may imagine different ones when rating their SE, and this could have effects on the dependent variables. In order to make such possible effects clear, let us consider the following imaginary experiment.

An experimenter is interested in examining the effects of SE level upon reactions to evaluative feedback. In order to do this, he decides to use an HSE group and an LSE group. He distributes a questionnaire to a large group of subjects and, on the basis of each person's rank-order of SE score, he divides the subjects up into two groups: HSE's and LSE's. Each subject comes along individually to a laboratory and performs an intelligence test. He is then given either positive or negative evaluative
feedback, and his expectations about future performance are measured. Such expectations constitute the dependent variable of this hypothetical experiment.

To simplify matters considerably, let us suppose that half of the subjects rated their SE for a social situation (e.g., being at a party), and the other half rated their SE for an intellectual situation (e.g., doing an exam). Obviously, the experimenter would not know this because he neither specified the situation(s) subjects should have in mind when rating their SE, nor did he ask subjects to indicate the situations they had in mind. Let us suppose further that all those who have LSE in intellectual situations have HSE in social situations, and vice versa. What effects would such an occurrence have on the results of this experiment?

It is clear that the experimenter's samples of 'HSE's' and 'LSE's' will be grossly inadequate for a testing of his hypothesis. This is because only half the 'HSE' group will have HSE in the intellectual situation of doing an intelligence test. Likewise, only half the 'LSE' group will have LSE for his experimental situation.

Therefore, if the experimenter found a difference between his two SE groups in their expectations about future performance after positive or negative feedback, such a difference could not be due to differences in SE level. Yet he would probably conclude it was. On the other hand, if the experimenter did not find a difference, it would be totally inappropriate for him to conclude (as he probably would) that SE level had no effect upon the dependent variable.

Obviously, such an extreme example is unlikely to occur. But it nevertheless serves to graphically illustrate the possible consequences of (a) variability in SE level across situations accompanied by (b) lack of experimenter awareness that this is the case. Totally inappropriate conclusions may be reached concerning the role of SE in behaviour.

A similar type of distortion may occur if subjects use different
comparison-others when making their self-ratings, and if this is not recognised by the experimenter. For example, some people may have LSE when they compare themselves with, say, their family, but HSE when they compare themselves with their peers at work. For others, this relationship between SE level and comparison-others may be reversed. Such a circumstance could have similar effects upon the relationships found between SE and the dependent variables as the variability of SE across situations.

Now, it is not the intention of this thesis to provide either a theory of how SE will vary across situations nor to provide a classification of situations. Instead, it is intended to see whether, in fact, SE does vary across situations in level and content; whether comparison-others vary intraindividually across situations; and if they do, to suggest that investigators in this area pay attention to such findings.

Gergen and his coworkers have already found that situational variables may have dramatic effects upon what they refer to as 'self-presentation' and 'SE' (though in terms of the definition of SE arrived at earlier, they in fact only measure self-assessments because they ask subjects to rate themselves on sets of characteristics and do not ask them how satisfied or dissatisfied they are with these self-assessments). They found that 'SE' could be so easily manipulated, without subjects feeling insincere or dishonest, that Gergen concluded that the prospects of a science of the self are dim.

Is this rather pessimistic conclusion justified? It all depends what is meant by 'exact' and what is exact enough depends upon the purposes of the investigator. Even if the self-concept is extremely ephemeral, and can only ever have as its basis the introspective reports of subjects, it may still be possible to formulate nomothetic 'laws' governing self-processes, and to construct models about them. Granted, the most parsimonious model of 'no change across situations' may not be applicable, but
this does not mean we should abandon the search for another model, or that a science of the self is limited.

For example, even if a person's SE fluctuates continuously across situations, there may well be limits to this fluctuation. SE could vary around an average level, which was different for each individual. Gergen himself (1977) implies this when he argues that the self-esteem change which occurs due to therapy may not be permanent. In other words, the individual reverts back to his 'usual' level when outside the influence of the therapist. People may also vary in the extent to which they vary around their average level. These 'laws' may not be very exact, but they are surely a step in the right direction.

Gergen's approach has been to vary the situation and to examine subsequent changes in 'SE'. Though this approach can determine situational variables which are potent enough to have effects on SE, it does not allow us to examine person-situation interaction nor to investigate the possibility that these changes in SE level are fluctuations around an average which characterizes them in their everyday lives. Suppose, for example, we know a person's average level of SE, a host of situational variables which might affect it, and also his range of fluctuation, we may well be in a position to make fairly detailed predictions regarding his feelings of identity in a variety of situations, and regarding how he should behave.

What variables have been shown to affect SE?

Gergen and Wishnov (1965) found that the characteristics of the other were important. The subject was given a self-descriptive note purportedly from another person, and the note either emphasised the good points of the other, or the others' bad points. Subjects were then asked to send a self-descriptive note to the other. Compared with previous self-ratings, subjects who received a note from the egotistical other used more positive terms and fewer negative terms to describe themselves, but subjects who received a note from the self-critical other used fewer
positive terms and the same number of negative terms to describe themselves.

Another experiment (Morse and Gergen, 1970) again showed that characteristics of the other affect self-ratings, this time physical characteristics. Subjects were applicants for two (real) jobs. On arriving for their preliminary assessments they filled in some self-rating forms and were then exposed to 'Mr. Clean' (an efficient, smart, business-like fellow) or to 'Mr. Dirty' (a scruffy, unkempt, inefficient fellow) who was obviously applying for the same job (or another one, but this made no difference to the results). Subjects then rerated themselves. Compared to their previous ratings, subjects exposed to Mr. Clean showed a decrease in 'SE' whereas those exposed to Mr. Dirty showed an increase in 'SE'. It may be surmised that this change in 'SE' was because subjects felt their chances of obtaining the job decreased and increased respectively - but unfortunately this was not measured.

These two experiments, however, show that the assumption that the reference group is unimportant as far as 'SE' is concerned is false. If the reference group is unspecified, the validity of the SE measure may be reduced. There is also the possibility that the reference group (i.e. comparison-others) changes according to the situation, and hence SE in these different situations may vary. As Gergen (1977) puts it 'What is true about the self depends upon those available for comparison' (p.153). Who subjects compare themselves with, and whether this varies across situations was investigated in the first study to be reported here.

Other variables which have been shown to affect SE include success and failure experiences (see Shrauger, 1975, for a review); listening to one's own voice versus not doing so (Ickes, Wicklund, and Ferris, 1973); being confronted with a mirror versus not being (Ickes et al, 1973); the goal of the interaction (eg. Gergen, 1965; Gergen and Taylor, 1969); and the subjects' motivations (eg. Jones, Gergen and Davis, 1962); amongst
others. These experiments are all extremely useful, helpful in destroying the myth that SE is non-varying. However, they generally examine mean differences in SE level across the experimental conditions, and do not look at other types of consistency – for example rank-order similarity as measured by correlation coefficients. As mentioned in Section 1.5(a), rank-order variability across situations is what should be measured in order to see if SE is a personality trait. Also, these studies use non-routine experimental situations. It would be useful and interesting to study SE in ordinary, everyday situations for two main reasons.

First, Rosenberg (1968) points out that people can be highly selective in a variety of ways in order to expose themselves to information which is not too damaging to their SE. For example, they can choose situations (such as occupations), friends, and values, and because of the inherent ambiguity which exists in social interaction they can impose virtually any interpretation upon any event. Finally, people can select from their memory store evidence to support any image of themselves they may wish to present. Given this psychological selectivity, it is possible that people's SE (at least in adulthood, when more choices are available, and the individual has more control over his life) is less varying in normal everyday situations than might be supposed from the available research. It may not be, but the matter is a worthy one for investigation. In a related vein, Duval and Wicklund's theory of objective self-awareness (Duval and Wicklund, 1972; Wicklund, 1975) has as a central concept the notion of avoidance of thinking about oneself. People are hypothesized to avoid self-confrontation when this would lead to an awareness of a negative discrepancy between their selves and their ideals on a given dimension. Again, this avoidance could mean that SE is fairly consistent in everyday situations.

Second, as pointed out earlier, SE may vary in several ways, two of which are content (the self-assessments may vary) and level. It is
possible that both vary in indiscriminate ways across situations, thus not being amenable to scientific analysis. Whilst level of SE has been studied (though, as mentioned above, only in relation to means) content of SE has remained largely unexplored. It is possible, for example, that the self-assessments (from which SE ultimately derives) vary in predictable ways across situations, depending upon what abilities and personal qualities the person perceives as being necessary in them. There may or may not be consensus among people regarding the relevance of particular abilities and qualities for particular situations, but it appears a topic worthy of investigation, because it is crucial to an understanding of how SE level affects behaviour.

It is proposed here that there is a strong connection between the level of a person's SE within a situation, and the activities performed in the situation. In order to take part in the situation, certain activities must be performed. The performance of such activities will lead to certain aspects of the self-concept being brought into awareness. SE level on these characteristics should remain fairly constant across situations providing the comparison-others remain constant. Thus, if two situations engage similar aspects of the self, they will also be associated with similar SE levels.

Such a proposition carries with it the assumption that aspects of the self-concept which are not engaged by a particular situation will not contribute to SE within that situation. This, however, is extremely difficult to test empirically. On the one hand, merely asking someone to rate their SE on irrelevant aspects of the self-concept will inevitably bring that aspect of the self-concept into awareness, along with its associated SE level. On the other hand, a comparison of aspects of the self-concept which are judged irrelevant to a particular situation with those that are judged relevant may not yield any significant differences. This could be because the subjects have HSE - therefore, by definition,
SE level on the relevant and irrelevant aspects of the self-concept should be the same (ie. there should be no difference between their self-assessments and their ideals). Thus this more fundamental proposition is not tested. Instead, it is assumed to be supported if situations engaging similar aspects of the self-concept are associated with more similar levels of SE than are situations engaging dissimilar aspects of the self-concept.

Unfortunately, content of SE has hardly ever been directly investigated. Gordon (1968) carried out a content analysis of responses to the 'Who am I' test (Bugental and Zelen, 1950; Kuhn and McPartland, 1954), but did not specify the situations (nor, indeed, the inter-rater reliability of his coding scheme). Whether the content of SE varies across situations and whether, as hypothesized, situations similar in SE content are also similar in SE level, is investigated in two studies of this thesis.

One study which attempted to investigate the variability of self-rankings across situations was conducted by Borgatta (1960). He first collected self-assessments from his subjects and then, several weeks later, they took part in discussions with four other subjects (all strangers to each other). The group discussed four topics, each for twenty minutes. Then subjects took part in a three-man discussion with different strangers and discussed three different topics, again for twenty minutes each. Amongst other things, subjects were asked to rank themselves and the other group members on forty items, which Borgatta (on the basis of factor-analytic work) grouped into three scores: assertiveness (A), sociability (S), and emotionality (E). When the self-rankings were compared in these two, what Borgatta calls 'independent' situations, moderate correlations were found: for A, $r = 0.42$; for S, $r = 0.51$, for E, $r = 0.50$. He concluded that there was a high stability in the self-rankings even when the subject takes part in different situations with different comparison-others.
However, several points may be raised which throw doubt on this conclusion. First, how independent were the situations? In terms of the abilities and skills required for effective functioning in the situations, they would appear very similar: both involved discussions, with strangers, on a variety of topics. So the parts of the self-concept brought into awareness may have been very highly related, leading to similar self-rankings. Second, mean differences were not examined, and they may (though I suspect this is unlikely) have been substantial.

Third, the self-ranking scores which were correlated were composite scores consisting of several items each. This would serve to reduce variability between the two situations in the self-rankings. In conclusion, Borgatta's study, whilst an important contribution, may underestimate variability in self-assessments across situations.

In sum, level of SE across situations remains to be investigated in the absence of potent experimental manipulations. In addition, there has been a paucity of research into the content of SE, and a lack of consideration to how situations might be classified in terms of SE level. Such omissions in research may have important implications both for theory and research within this area.
Wylie (1974) agrees with the present author concerning the over-abundance of definitions and constructs which are only loosely (if at all) articulated into a theoretical framework from which testable hypotheses can be derived. Not only are there too many of these constructs, but many of them are ill-defined and ambiguous. Because of this ambiguity, it is difficult to know what kind of operational definition the particular construct should have, and which kinds would be inappropriate. For example, Rogers (1951) says that the self-concept consists only of those things which are admissable to awareness, and that behaviour depends upon the self-concept. Yet later on, he states that unconscious processes may affect behaviour as well. He neglects a proper analysis of the problems posed by his possibly conflicting statements, such as whether a measure of the self-concept should try to assess unconscious processes as well as conscious ones, and whether what a person says about himself should be taken at face value.

Wylie's thorough book makes it clear that there are several areas of weakness in research concerning self-evaluation. She concludes that, in general, the methodology is inadequate and does not take into account the knowledge available concerning how to construct tests, how to test and increase their validity, and so on. Some of her main criticisms are:

(1) The method used is often only vaguely indicated, making interpretation of results, further analysis, and replication impossible.

(2) In many studies, there is no mention of construct validity or, at best, there is only an inadequate exploration of it. She points out, for example, that the multitrait-multimethod method of Campbell and Fiske (1959) is geared towards the assessment of construct validity and yet very few such studies have been carried out in this area. An exception to this is Hamilton (1971) who used five measures of self-esteem, four
measures of dominance and four measures of dogmatism. The method is basically extremely simple, and involves using at least two measures of a particular trait (say SE) to measure the trait and at least one other. The resulting scores are then intercorrelated, and the correlation matrix is examined for evidence of convergent and discriminant validity. Convergent validity occurs when correlations between different measures of the same trait are high, and discriminant validity occurs when the correlations indicate that the measures lack variance due to irrelevant factors (such as the method used). For example, the convergent validity correlations should be higher than the correlations between different traits measured by the same method. Hamilton (1971) found in his matrix that there was reasonable convergent validity for three of the self-esteem measures (the Janis-Field inadequacy Scale, 1959; self-ratings; and two scales from the California Psychological Inventory: Gough, 1957) but rather less evidence of discriminant validity. Obviously, much more work is required along these lines, however.

Researchers in this area have also tended to neglect an analysis of variables, other than the one ostensibly being assessed, which might exert systematic influences upon results and hence provide alternative explanations for them. One of these is social-desirability, or the tendency to present oneself in a favourable light regardless of one's real opinion about oneself. However, after examining the literature carefully herself, Wylie (1974) concludes that this may not be too much of a problem providing reasonable precautions are taken (which they are often not) such as establishing rapport, and making it desirable for subjects to be honest. The main bases for this rather encouraging conclusion (which holds pending further research) are that several assumptions upon which the social-desirability explanation of results rest are not substantiated by relevant studies in the literature (see Wylie, 1974, pp.52-66 for a more complete discussion of this issue). For example,

1 NOTE: These are not measures of self-esteem as defined here.
one assumption is that all subjects agree on what is socially desirable, but this has been found not to be the case (e.g., Borislow, 1958; Messick and Jackson, 1961). In any case, even if subjects do display the socially desirable responses, they may well be being totally honest—merely exhibiting the signs of adequate socialisation. It is, of course, virtually impossible ever to know whether or not the subject is lying. Most research (not only in the area of self-evaluation) assumes subjects are being honest.

The most common type of validity of a measure assessed in the literature is predictive validity: that is, a prediction is made on the basis of theory and tested using a particular SE or self-concept measure. However, this provides only ambiguous information regarding the validity of the instrument because if the prediction does not hold up when empirically tested it could either be because the theory is wrong, or because the measure is inadequate. On the other hand, if the prediction does hold up it could either be because the theory is correct or because there are method factors which systematically influence the results in the direction predicted by the theory. Therefore, at best, only partial support for the theory is forthcoming. To be more sure the theory is correct would require replication using different instruments—then at least method factors could virtually be ruled out of the set of alternative explanations for the results found (though never absolutely completely, of course).

It is rather more difficult to explain away the predicted result of an experiment in terms of method factors, however, when the design of the experiment is anything other than the simplest. It would then become necessary to postulate that method factors interact in specific ways with the experimental variables. To illustrate what is meant here, suppose a three-by-three analysis of variance type design is used, with different subjects in each of the nine cells—let us say three levels of self-esteem, and three levels of positive reinforcement being the independent variables,
with effort expended in doing an experimental task being the dependent variable. In order for method factors to account for the results obtained from a study of this type, they would not only have to explain any main effects which might occur, but all of the interactions as well. I would suggest that this becomes increasingly implausible the more complex the experimental design.

(3) Sometimes necessary control groups are lacking, or subjects are inappropriately chosen, or the basis for their assignment to groups is not made clear with no information being provided about randomization procedures or matching.

(4) Findings of marginal statistical significance may be used to infer psychological processes of great importance. So often results in the predicted direction which do not reach conventional statistical significance are treated as if they do, and in many studies such a large number of statistical tests are made that some 'significant' results would be expected purely by chance. Another common failing is that groups of matched 'predictees' are not used when the investigator is attempting to establish a 'chance' baseline.

(5) Often studies are not replicated and therefore the stability of the findings is open to question.

(6) In many studies, artefacts may occur because of overlapping instruments, failure to use blind judgements, effects of common response sets (which may inflate the relationship between the independent and dependent variables) and the use of verbal reports to measure both the self-concept and behaviours allegedly related to the self-concept.

It is obviously unforgivable for an investigator to present his study in an ambiguous way, so that replication is impossible; and it is also unforgivable to use inadequate designs and statistical analyses. However, Wylie's justifiable criticisms neglect one point which is a
very important one: the necessity for exploration of certain hypotheses. Although using hundreds of statistical tests within one study may produce artefacts, it may also reveal relationships among variables which might never otherwise come to light, and these relationships may be further studied in later work. Nevertheless Wylie's criticisms are well taken. There are so many people working in the area of self-conception and SE that it seems essential to pay more attention to methodological factors of the sort she mentions, and soon, or else the rather poor returns obtained in this area will continue.

Wylie (1961, 1974) reviews in great detail the methods used to measure SE and the self-concept by well over 1000 authors. Her aim is to critically evaluate methodology of recent research, and she says less about the theorizing behind the experimental hypotheses. Burns (1979) also has a useful chapter on extant measuring instruments. As there seems little point in merely repeating what Wylie says, only the issues most relevant to the research carried out here will be discussed.

As already pointed out in sections 1.2 and 1.3, there is an important distinction between self-assessments and SE. All too often, however, these two different constructs are treated as if they mean the same thing. Many measuring instruments in fact measure self-assessments rather than SE. Insofar as theories refer to SE, therefore, such confusion in measuring instruments will not help those who wish to empirically test their theories.

Another problem with extant measuring instruments is their probable cultural specificity. To use an instrument in a culture outside that in which it is developed will probably reduce its validity. Most SE questionnaires originate in the U.S.A., and to use them in Great Britain can only exacerbate the twin problems of meaning and relevance mentioned earlier.

Central to the definition of SE supplied in section 1.3 is the notion of a comparison: the person compares what he thinks he is like (his self-assessments) with what he would like to be like (his 'ideals').
Insofar as his self-assessments match his 'ideals', in a particular set of attributes, the person may be said to have high self-esteem (HSE). Insofar as the self-assessments are a long way from the 'ideals', the person may be said to have low self-esteem (LSE), for that set of attributes.

A very direct way of measuring SE springs to mind immediately the above points are considered, namely: to use a discrepancy measure. That is, for a particular set of items, a person rates two concepts. First, the concept 'what I am really like' (providing his self-assessments) and, secondly, the concept 'what I would ideally like to be like' (providing his 'ideals'). For each item, the discrepancy between his self-assessment and his 'ideal' can be measured, and combined in some way with such discrepancies from the other attributes to give an overall SE score. Discrepancy measures are not uncommon in the literature. For example, the Q-sort procedure (Stephenson, 1953) lends itself easily to such a use. Butler and Haigh (1954) developed a set of 100 items which the subject has to sort twice into 9 piles such that there is a quasi-normal distribution of items in each pile. The subject sorts the items the first time according to the degree they are 'like me' (ie. the self-assessments), and the second time according to the degree 'I would most like within myself to be' (ie. the ideals). The correlation between the two sorts provides an index of SE. Such a measure has been used extensively by Carl Rogers and his coworkers in their client-centred therapy (eg. Rogers and Dymond, 1954). The correlation between 'self' and 'ideal' may be regarded as a discrepancy measure of SE. The lower the correlation, the larger the difference between the self and the ideal.

Another form of discrepancy measure has a modified semantic-differential format (Osgood, Suci, and Tannenbaum, 1957). The semantic-differential may be regarded in two ways, First, as an instrument designed to test Osgood et al's theory of meaning. As such it is probably inappropriate
as a measure of SE because the items in the questionnaire were not
chosen with measuring SE in mind. The second way of regarding the
semantic-differential, however, is as a Likert-type rating scale which
is a measuring technique applicable to any attitudinal concept. As such,
it may be a highly appropriate way of measuring SE, providing the criteria
for item selection are made clear.

The problem of transferring a set of items from one culture to
another has already been mentioned, as have the issues of meaning and
relevance. Given these problems, it seemed highly desirable to derive
a new measure of SE which is relevant and meaningful to the subjects on
which it is to be used. The semantic-differential type was chosen rather
than the Q-sort for the following reasons:

(a) It is convenient to administer to a group of people at once.
(b) Such a questionnaire does not take a long time to complete
(depending, of course, on the number of items).
(c) Heise (1972) presents evidence that, virtually irrespective of
the items used, the semantic-differential has adequate reliability.

Given that a modified version of the semantic-differential is to be
used, it is highly appropriate to consider some of the criticisms which
have been levied against it as a measure of SE, principally by Wylie (1974)
and Burns (1979).

A major criticism of Wylie is that the semantic-differential was
constructed in order to test a specific theory of meaning. Thus the use
of items for measuring SE which are simply chosen from the Osgood et al
(1957) item-pool may be totally inappropriate. This criticism obviously
refers to the issue of item-selection. Whilst it is accepted that item-
selection is of crucial importance, such a criticism does not refer to
the semantic-differential format as a technique of measurement.

A more serious criticism concerns the assumption, mentioned previously,
that such discrepancy measures make. This assumption is that there is a
monotonic relationship between the size of the discrepancy, and the degree of distress it produces in subjects. As Wylie (1974) points out, it may be that discrepancies of the same size produce different degrees of distress in subjects because the items on which they were obtained differ in importance. Though this assumption has not, to my knowledge, been tested, it is very easy to examine empirically. This is done in Chapter 4.

Another criticism concerns the possibility that, for some subjects, 'self' scores may be less than their 'ideal' scores, whereas for others 'self' scores are greater than their 'ideals'. Wylie feels that such scores may not mean the same thing, though they are treated as equivalent in the computation of the overall SE score. This criticism is rejected here. The whole basis for measuring SE using a discrepancy score is the assumption that it is the discrepancy per se which has its own importance to the subject, in addition to actual 'self' or 'ideal' score. Now, it may be the case that self-assessments and ideals exert their own influence upon behaviour. The level-of-aspiration literature (eg. Lewin et al, 1944) for example, clearly demonstrates the effects of goal setting and of different self-assessments. In addition, people whose 'ideals' are such as to make them behave contrary to societal norms will probably experience difficulties independently of their self-ideal discrepancies. Nevertheless, it is contended here that the size of the discrepancy per se has motivating effects on behaviour. Also, of course, where 'self' scores and 'ideal' scores lie will be largely determined by the wording of the item. However, it still seems worthwhile to investigate how often 'self' scores are greater than 'ideal' scores, and data relevant to this is presented in Chapter 4.

Kaplan (1972) draws attention to another problem of the semantic-differential. This is the problem of ambivalence which arises when questionnaires are used where item-pairs are presented as bi-polar opposites. In a traditional semantic-differential format, the middle-point
has a dual meaning. It could either mean that the subject feels the item is irrelevant to the concept, or that the subject feels both poles apply equally. If the latter, there is no way of telling how intensely the subject feels this ambivalence. Such dual meaning generalises to all the points on the scale. Insofar as ambivalence exists, therefore, using bi-polar scales is highly misleading. Kaplan (1972) suggests that a simple way round this problem is to present each pole of each item separately. However, this has the disadvantage of doubling the length of the questionnaire.

Now, ambivalence is not of central importance here. Nevertheless, the problem is recognised. Therefore, in some of the questionnaires used here (where it does not matter if they are twice as long), both poles of an item are separately measured. Where time is at a premium, only one pole is measured. This removes the confusion inherent in bi-polar measures, but does not assess ambivalence. Considerable data has been amassed by the present author on ambivalence in SE, but this has been left out of the present thesis due to restrictions of space.

Finally, as in all Likert scales, a particular SE score can be obtained in a whole variety of ways. For example, a person may have an LSE score because he has small discrepancies on a large number of items, or because he has large discrepancies on only a few items. The assumption is made, therefore, that the discomfort produced is equivalent for all subjects with the same SE score. This may or may not be a reasonable assumption. Insofar as the questionnaire has good construct validity, it is reasonable to suppose that the making of this assumption does not produce too much measurement error.

In Chapter 3, an SE questionnaire is used where the items are selected with schoolchildren in mind. In Chapter 4, an SE questionnaire is presented for use on university and polytechnic students. Considerable attention is given to item-selection, and an empirical test is made of
the criticisms discussed here. In addition, some evidence is provided relating to its construct validity. The questionnaire, on the basis of the evidence provided in Chapter 4, appears to be a good measure of SE as defined here. It is therefore used in all the later experiments of this thesis. As will be seen, its construct validity remains high.
1.7: Summary and Proposed Investigations

In the present chapter it is argued that, first, there is an important theoretical distinction to be made between self-assessments and SE. Self-assessments involve an individual's evaluation of himself with respect to particular attributes. SE involves the individual comparing his self-assessment with an ideal state and may be conceptualized as the resultant 'self-ideal' discrepancies. It is pointed out that this distinction is very often not made in the literature, leading to conceptual confusion and poor progress in our attempts to understand how SE affects behaviour.

Secondly, it is argued that several assumptions about SE are made in the literature, and that these assumptions have not been adequately tested. If shown to be empirically unjustified, then to make these assumptions will further muddy the waters of our knowledge about SE. These assumptions are: (a) that SE is a personality trait which does not vary across situations in either level or content; and (b) that the comparison-others do not vary across subjects or situations.

Thirdly, it is suggested that a good way to operationalise the concept of SE is to measure it using a discrepancy measure. In this, the individual rates his self-assessments for a set of attributes and, in addition, rates his 'ideal' scores for each attribute. It is argued that a modified version of the semantic-differential questionnaire format is appropriate. Several criticisms of discrepancy measures when used to assess SE are presented and discussed. It is pointed out that many of these criticisms apply to all questionnaires (eg. lack of clear criteria for item selection), and can be dealt with by using satisfactory techniques for development of the questionnaire. Several of the criticisms depend upon whether, or not, the assumptions made by discrepancy questionnaires when used as a measure of SE, are empirically justified. Such assumptions
may easily be put to empirical test.

It is therefore proposed to investigate:

(a) Whether or not the assumption is justified that content of SE does not vary across situations;

(b) Whether or not the assumption is justified that comparison-others do not vary across situations;

(c) Whether or not the assumption is justified that level of SE does not vary across situations. In this chapter it is suggested that lack of attention by SE researchers to the social situation is a characteristic shared by many other social psychologists. It is pointed out that only recently has attention been paid to the conceptualization and classification of situations. It is proposed that different situations will engage different aspects of the self-concept, depending upon the activities needed to accomplish the goals of the participants in the situation. Insofar as comparison-others are stable across situations, it is hypothesized that situations engaging similar aspects of the self will also be similar in SE level, and that situations engaging different aspects of the self will be less similar in SE level. This hypothesis is also investigated in this thesis.

(d) It is proposed to develop a questionnaire to measure SE, and to test empirically the assumptions made by the questionnaire. In particular, the assumption is tested that there is a monotonic relationship between the size of the discrepancy and the degree of distress it produces in the individual.

(e) Once these very basic assumptions have been examined, it is proposed to investigate the role of SE in determining subjects' reactions to self-relevant feedback information. A full discussion of the theoretical frameworks and empirical research in this area is postponed till Chapter 5. Chapters 6–8 present a series of experiments designed to resolve a major controversy in the area. Chapter 9 contains the conclusions of the thesis as a whole.
CHAPTER 2

Content of Self-Esteem and Comparison-Others
at Home and at School
ABSTRACT

An exploratory study is presented which examined two assumptions permeating the literature on the self-concept and SE. The first is that the content of SE does not vary across situations, and the second is that the people with whom subjects compare themselves (i.e. comparison-others) do not vary across situations. A modified version of the Kelly (1955) repertory grid technique was used to elicit both the comparison-others and the constructs used to describe the self in two environments: home and school. It was found that a substantial proportion of the constructs were different at home and at school, and that these constructs were closely related to the activities performed in the two environments. Also, the comparison-others varied dramatically in the two environments: at school they were virtually all of the same sex and age whereas at home they were not. Thus both assumptions are shown to be false. Implications for both theory and measurement are discussed.
As mentioned in Chapter 1, most research into the self-concept and self-esteem (SE) has rested upon two implicit assumptions. The first is that the SE of a person is a general personality characteristic which does not vary across situations. It was pointed out that there are several extremely important ways in which SE might vary intraindividually from one situation to another. Two of these are:

(a) in its content: the aspects of the self which are brought into conscious awareness in one situation may not be the same as the aspects of the self which are focussed upon in another situation; and

(b) in its level: a person's self-ideal discrepancies across the set of aspects of the self which are relevant in one situation may be different from the self-ideal discrepancies associated with the aspects of the self which are relevant in another situation.

Now, if SE does affect behaviour, it is important to examine how SE might vary across situations. Accordingly, the study reported in the present chapter investigates whether content of SE varies across situations. Later studies look at both content and level of SE across situations.

Variability of SE across situations has implications for measurement as well as for theory. Firstly, if certain aspects of the self are more relevant in some situations than in others, the validity of SE scores will be reduced if they are obtained without specification of the situation(s) to which they should apply. Secondly, if the situation is not specified subjects may imagine different ones when making their ratings, and this would also lower the validity of SE scores.

The second implicit assumption embedded in the literature is that the people with whom subjects compare themselves (i.e. comparison-others) when reaching assessments of their abilities and skills, do not need to
be taken into account when carrying out research on SE or the self-concept. Again, this assumption should be put to the test because it has implications for both theory and measurement. It is perfectly possible that comparison-others do vary intrinsically across situations, and also that different people use different comparison-others depending upon the groups to which they feel they belong. For example, the comparison-others used by males may be different from those used by females. Comparison-others provide the standards against which people judge themselves. If the comparison-others vary across situations, it is also possible that the self-assessments may vary as a direct result. For example, a person who compares herself with schoolchildren at school and with her family at home could well reach different self-assessments of her intelligence at school and at home.

There are at least four possible consequences of intrinsically variations in comparison-others across situations for particular aspects of the self;

(a) The different comparison-others do not produce changes in either self-assessments or values ('ideals');

(b) The different comparison-others produce changes in self-assessments, but values remain relatively unvarying;

(c) The different comparison-others produce changes in values, but not in self-assessments;

(d) The different comparison-others produce changes in both values and in self-assessments.

Thus, for possibilities (b); (c); and (d); changes in level of SE would occur as a direct consequence of variation in comparison-others across situations. Even if situations differ totally in the aspects of the self focussed upon within them, the use of different comparison-others may still have consequences for SE: for example, in the goals which are set.
It is therefore important for a theory of the self to consider (a) whether comparison-others do vary across situations for an individual; (b) whether individuals differ from each other in the characteristics (eg. age, sex) of the comparison-others they use; and (c) the consequences of such variations in comparison-others.

There has been a great deal of research following on from Festinger's (1954) theory of social comparison processes. Most of this work has involved presenting subjects with different comparison-others and measuring changes in self-assessments of particular skills and opinions. It has been shown that different comparison-others do, indeed, affect a person's self-assessments (see, for example, Latané, 1966, for a review of early studies). More recently, the theory has been invoked to explain inter-group comparison processes (eg. Turner, 1975, Billig, 1976; Suls and Miller, 1977; Moscovici and Paicheler, 1978; Tajfel, 1978; Turner, 1978). This is not the place to review such research. Though the importance of comparison-others for theorising about the self has been recognised (eg. Harré and Secord, 1972; Wylie, 1974), little empirical work appears to have been performed.

An exception to this general rule is a study by Morse and Gergen (1970), which shows how important comparison-others are for self-assessments. In their study, students who had applied for a job came along to a laboratory to complete forms relevant to their application. When they arrived, they found another person already there. This other was either smartly dressed and efficient (Mr. Clean), or was sloppily dressed and inefficient (Mr. Dirty). The self-assessments of students who had Mr. Clean as a comparison-other dropped compared to those who had Mr. Dirty as a comparison-other. But this study does little to further our understanding of who will be chosen as comparison-others in everyday life.

The exploratory study reported in this chapter aims to critically
evaluate two of the assumptions in the literature on SE. First, that content of SE does not vary across situations. Second, that the comparison-others used do not vary across situations. Sex differences in the latter are also examined.

A modified version of Kelly's repertory grid technique (Kelly, 1955) is used in order to elicit the constructs used to describe the self in two environments: home and school. These constructs form the content of SE within these two environments. Comparison-others were found by asking subjects to name up to five people who were similar to them and up to five people who were different from them in each of the environments. Subjects then provided three characteristics of each comparison-other:

(a) the sex
(b) the age
(c) the relationship he/she had with the subject (whether a family member or otherwise).

It is appropriate to stress here that there were no constraints whatsoever upon subjects' choice of comparison-others. They were perfectly free to choose the same people at home and at school if they so wished.

It should be pointed out that the rationale behind this study rests upon several assumptions. These assumptions are outlined here so that the reader may judge their reasonableness. In the first place, it is assumed that by asking subjects to describe themselves in relation to other people (of their own choosing), their important self-relevant constructs will be elicited. This implies that there are few, if any, self-relevant constructs which are derived in the absence of self-other comparisons. It is assumed further that these self-relevant constructs may be regarded as the content of SE: that is, these constructs are proposed to be the important dimensions along which subjects compare their assessments of what they are like with what they would like to be like. Thirdly, it is assumed that subjects actually do compare themselves
with the people they nominate as comparison-others. It should be emphasized, in support of this assumption, that subjects were not 'forced' to produce five comparison-others in each condition, and that they had a completely free choice in their nominations.
METHOD

Subjects: These are male and female schoolchildren aged 14 and 15 years old. They are from two different schools. There are 32 subjects in all, 16 from each school. Half in each group are males, and half are females. All are white and English-born.

Design: For the content of SE, there is one independent variable with two levels, namely: whether the constructs are obtained at home or at school.

For the comparison-others, there are two independent variables, each with two levels, namely: (i) home vs school

and (ii) sex of subject (male vs female)

Procedure: Subjects were interviewed individually in order to allow them maximum freedom in response time, and to ensure their privacy. The interview schedule was designed so that the constructs used to describe the self, and the comparison-others, could be elicited for the two different environments: home and school. These two environments were chosen because they applied to all subjects. For half the subjects of each sex at each school, the home environment was dealt with first. For the other half, the school environment was dealt with first. Thus any order-effects (such as taking time to 'warm up' in the interview) would not affect the dependent variables. Also, two female interviewers were used, whose interviewees were randomly chosen. Thus the results are not due solely to experimenter effects associated with one interviewer.

Upon arrival at the interviewing room, a subject was asked to sit down opposite the interviewer who immediately introduced herself. She explained that she was interested in what people thought about themselves, and asked the subject if he/she minded the interview being taped. Once
assured of its confidentiality, all said they did not.

The actual procedure used was a modification of the Kelly repertoary grid technique (Kelly, 1955), as follows (it is described here for the 'at home' environment. Apart from changes in the wording, it is identical for the 'at school' environment):

Each subject was asked to imagine that he/she was at home, and to describe the things that he/she did there. This was to try and ensure that the subject was thinking carefully about being at home.

The subject was then given five pieces of card and asked to write down the names of up to five people who were similar to him at home, and up to five people who were different from him at home. These are the comparison-others. On the relevant card, the subjects also wrote down the age and sex of each comparison-other and specified the relationship he/she had with the comparison-other (eg. mother, father, girlfriend etc.).

Next, the subject was asked to describe, for the similar comparison-others, in what ways each was similar to him/her; and for the dissimilar comparison-others, in what ways each was dissimilar to him/her. Each comparison-other was considered in turn. In this way, the constructs used to describe the self at home were elicited. This provided the content of SE at home.

The procedure was repeated in exactly the same way for the environment of being in a lesson at school.

Thus subjects were able to provide up to 20 comparison-others, but they were not forced to produce more than they wished.

Care was taken in the interviews not to provide subjects with constructs. Instead questions such as 'why?' or 'can you give a reason for that?' were used as a means of following leads that subjects gave. Rapport between interviewer and interviewee was generally high. The interviews were from 25 minutes to 1½ hours in length depending upon how quickly subjects answered the questions.
RESULTS

The results may be divided into two sections:

(1) the constructs used to describe the self in the two environments; and

(2) the comparison-others used in the two environments.

(1) Constructs used to describe the self

(a) Derivation of the constructs

The tape-recordings of the interviews were transcribed, and each construct used was noted on a card. The constructs were not at all difficult to pick out because the subject would generally use a sentence of the form 'I am X but he is Y' for the different comparison-others; or of the form 'we are both X' for the similar comparison-others.

The pole of the construct used to describe the self was noted on a card. Sometimes subjects would use a particular construct more than once - both within a self/other comparison and in different self/other comparisons. For each subject, the number of self/other comparisons in which a construct was used was counted.

Over 300 constructs were produced altogether. Often, however, different subjects appeared to be using the same construct, but in a slightly different form. For example, one subject might say 'I mess around, but he does not'; whereas another might say 'I fool about but he does not'. These different expressions of the same construct would originally both have been put on different cards.

But in order to allow some form of statistical analysis of the data, and because it did not seem sensible to call two obviously very similar constructs different, it was decided to group the similar constructs together. This was done initially by the experimenter, and the original
number of constructs was reduced to 140. Now, this is a risky business: the grouping of similar constructs by one person may not be the same as the grouping of similar constructs by another. Accordingly, two checks were made on the reliability of the procedure. These were as follows:

Check I
Method: A fifth of the constructs (ie. 28) were randomly chosen from the reduced number, with the provisos that each construct was used by four or more subjects (ie. one-eighth of the total) and consisted of four or more different expressions. Two independent raters were then asked to rate how sure they were that a particular expression used by a subject was an instance of the construct, on a 0 (it is definitely not an instance) to 10 (I am 100% sure it is an instance) scale. They were told to include as instances of the construct expressions which contained its opposite.

The 28 constructs had associated with them 221 expressions.
Results: The mean rating for the first rater was 9.79 (SD=0.88), and for the second was 9.85 (SD = 0.62). For the first rater, 92.76% of the expressions were given a rating of 10; and for the second rater 92.3% of the expressions were given a rating of 10.

Thus there is good evidence that the expressions are regarded as instances of the construct by both raters as well as the experimenter.

Check II (using the same two independent raters as for Check I)
Method: Five construct categories which had not been used in the first check were randomly chosen such that each was used by four or more subjects and each had four or more expressions.

Each rater was given five cards, each with a different construct category named on it. They were then read in a random order expressions which had been put into the five construct categories by the first
experimenter. Their task was to decide into which construct-category to place the expression. There were 46 expressions altogether.

Results: In effect there were three raters, including the experimenter. The measure of agreement was the percentage of the expressions put into the same category by all three raters. It should be noted that this is an extremely strict measure of agreement.

Agreement was 91.3% and disagreement was 8.7%. There were 42 agreements and four disagreements. This difference is highly significant ($\chi^2 = 29.76, p < 0.001$).

As a result of these two checks it may be concluded that the grouping procedure has a high degree of reliability, at least for the commonly-used constructs. Instances of a construct could readily be recognised and distinguished from instances of other constructs.

(b) Does construct usage differ from home to school?

There are 46 constructs which were used by nine or more subjects (ie. by over one quarter of the sample). This was regarded as being sufficiently frequent to allow a meaningful statistical comparison of the differential usage of these constructs at home and at school.

For each of these 46 constructs it was counted in how many self/other comparisons a subject used it at home, and in how many self/other comparisons he used it at school.

Using this data, each construct could be analysed using the Wilcoxon test to see if it was used significantly ($p < 0.05$) more often at home than at school. Of the 46 constructs, 13 were used significantly more often at home; 14 were used significantly more often at school; and the usage of the remaining 19 did not differ significantly from home to school. The results of these analyses are shown in Table 2.1.
### TABLE 2.1

WILCOXON RESULTS (2-tailed) UPON CONSTRUCTS USED BY NINE OR MORE SUBJECTS

(a) Constructs Used More at Home

<table>
<thead>
<tr>
<th>Construct</th>
<th>No. Self/Other Comparisons in which construct used at:</th>
<th>Wilcoxon 'T'</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Home</td>
<td>School</td>
<td>N</td>
</tr>
<tr>
<td>Relations with parents</td>
<td>38</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>Going out/staying in</td>
<td>52</td>
<td>12</td>
<td>22</td>
</tr>
<tr>
<td>Arguing and temper</td>
<td>32</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>Dress</td>
<td>41</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>Looks/appearance</td>
<td>38</td>
<td>17</td>
<td>19</td>
</tr>
<tr>
<td>Liking of sport</td>
<td>23</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>Relations with siblings</td>
<td>26</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>Age</td>
<td>19</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Friends at Home</td>
<td>20</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Helping parents</td>
<td>56</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Neatness</td>
<td>15</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>Working on a job</td>
<td>20</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Being bothered</td>
<td>12</td>
<td>3</td>
<td>9</td>
</tr>
</tbody>
</table>

Continued ...
TABLE 2.1: continued

(b) Constructs Used More at School

| Construct                      | No. Self/Other Comparisons in which construct used at: | Wilcoxon 'T' | p ≤  
|-------------------------------|--------------------------------------------------------|--------------|------
|                               | Home | School | N |       |       |
| Talkativeness                 | 12   | 53     | 25 | 5.5   | 0.01  |
| Fooling around                | 8    | 46     | 22 | 0.0   | 0.01  |
| Academic ability              | 13   | 52     | 22 | 9.0   | 0.01  |
| Attitude to schoolwork        | 7    | 58     | 21 | 7.0   | 0.01  |
| Humour                        | 11   | 26     | 19 | 19.5  | 0.01  |
| Effort in schoolwork          | 4    | 68     | 17 | 0.0   | 0.01  |
| Listening                     | 1    | 13     | 11 | 4.5   | 0.01  |
| Obedience                     | 2    | 14     | 11 | 0.0   | 0.01  |
| Academic achievement          | 2    | 10     | 10 | 8.0   | 0.05  |
| Shyness                       | 5    | 11     | 10 | 2.5   | 0.06  |
| Answering or asking questions | 1    | 9      | 9  | 4.0   | 0.05  |
| Being interrupted and distracted | 2  | 15     | 9  | 6.0   | 0.02  |
| Cheekiness                    | 3    | 11     | 9  | 5.0   | 0.07  |
| Where people sit              | 0    | 23     | 9  | 0.0   | 0.07  |

Continued...
### TABLE 2.1: continued

(c) Constructs whose usage does not differ significantly from home to school

<table>
<thead>
<tr>
<th>Construct</th>
<th>No. Self/Other Comparisons in which construct used at:</th>
<th>Wilcoxon 'T'</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Home</td>
<td>School</td>
<td>$N^1$</td>
</tr>
<tr>
<td>Friendship</td>
<td>27</td>
<td>33</td>
<td>24</td>
</tr>
<tr>
<td>Speaking to people</td>
<td>19</td>
<td>33</td>
<td>21</td>
</tr>
<tr>
<td>Quietness</td>
<td>14</td>
<td>24</td>
<td>16</td>
</tr>
<tr>
<td>Rowdy</td>
<td>9</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>Being told off</td>
<td>8</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Aggression (physical)</td>
<td>12</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>Being right (or wrong)</td>
<td>9</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Mixing with others</td>
<td>9</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td>Going round with</td>
<td>7</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Status</td>
<td>7</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Topics of conversation</td>
<td>7</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>Homework</td>
<td>17</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Showing off</td>
<td>8</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Helping peers</td>
<td>9</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Nice</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Relations with opposite sex</td>
<td>10</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Satisfaction with self</td>
<td>6</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Speed of work</td>
<td>2</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Standing up for self or other</td>
<td>5</td>
<td>5</td>
<td>9</td>
</tr>
</tbody>
</table>

**NB:** $N^1$ shows the number of subjects using the construct.
Table 2.2 shows a further seven constructs which were used exclusively at home or at school but only by between four and eight subjects (i.e. by less than an eighth but by more than a quarter of the sample).

As can be seen from these tables, the content of SE does appear to vary in these two environments. Most of the constructs used more frequently at school are concerned with academic abilities and attitudes towards schoolwork. At home, however, constructs are concerned with how people look, and how they get on with their family. The constructs which are not used significantly more often at home or at school are concerned with more general social skills, such as 'friendship' or with how the subjects react to people in authority - a common element to both home and school.

Given that the content of SE does appear to vary from home to school, it is perfectly possible that level of SE also varies between these two environments. Level of SE is not measured in this study, however (though it is in the next two).

(c) Subsidiary Analyses

(i) Are some constructs used more often by males than by females?

The constructs used by nine or more subjects were analysed, using Mann Whitney tests, with this question in mind. None of the constructs were used more often by one sex than by the other.

(ii) Are some constructs used to describe one sex more often than another?

Again, this analysis was performed upon all constructs used by nine or more subjects. Once again, no significant differences were found for any construct.
TABLE 2.2
Constructs used exclusively at home or at school by < 9 but by ≥ 4 subjects

<table>
<thead>
<tr>
<th>(a) Used Exclusively at Home: Construct</th>
<th>No. Self/Other Comparisons N in which used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health</td>
<td>10</td>
</tr>
<tr>
<td>Religiousness</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(b) Used Exclusively at School: Construct</th>
<th>No. Self/Other Comparisons N in which used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessments of work</td>
<td>12</td>
</tr>
<tr>
<td>Creep</td>
<td>7</td>
</tr>
<tr>
<td>Relations with teachers</td>
<td>13</td>
</tr>
<tr>
<td>Truancy</td>
<td>16</td>
</tr>
<tr>
<td>Writing</td>
<td>7</td>
</tr>
</tbody>
</table>

NB: N<sub>1</sub> shows the number of subjects using the construct
(d) **The activities performed at home and at school**

Subjects were asked to imagine they were at home (or at school) and to say what they did there. The data thus generated were examined. 46 different activities were mentioned, but only 11 of these were mentioned by nine or more subjects. Upon examination of these 46, it seemed to the present author that they could be classified along two dimensions at home: (a) leisure - helping and (b) social - unsocial. At school, two dimensions also appear to serve as a classificatory heuristic, namely: (a) proschool - antischool and (b) social - unsocial. This classification, which is derived from a post-hoc examination of the data, is presented in order to facilitate understanding of the constructs data. The 46 activities are shown classified, along the above mentioned dimensions, in Table 2.3.

As can be seen from Table 2.3, the activities are very closely related to the aspects of the self focussed upon in each environment. This does aid interpretation of the constructs: many of them at home do appear to be concerned with the leisure/non-leisure distinction in the activities; those at school with the proschool/antischool distinction; and those used about equally often at home and school with the social/non-social dimension.
### TABLE 2.3

**Classification of activities performed at home and at school**

(a) **Home Activities**

<table>
<thead>
<tr>
<th>(1) High Leisure; Social</th>
<th>(2) High Leisure; Non-social</th>
</tr>
</thead>
<tbody>
<tr>
<td>Go out with friends</td>
<td>Watch T.V.</td>
</tr>
<tr>
<td>Play games with family</td>
<td>Play records alone</td>
</tr>
<tr>
<td>Fight with siblings</td>
<td>Play musical instrument</td>
</tr>
<tr>
<td>Do hobbies with family</td>
<td>Crochet</td>
</tr>
<tr>
<td>Play records with friends</td>
<td>Go into own room</td>
</tr>
<tr>
<td>Have friends round</td>
<td>Go for walks alone</td>
</tr>
<tr>
<td>Play football</td>
<td>Read</td>
</tr>
<tr>
<td>Mess around with others</td>
<td>Write</td>
</tr>
<tr>
<td></td>
<td>Draw</td>
</tr>
<tr>
<td></td>
<td>Make things</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(3) Low Leisure; Social</th>
<th>(4) Low Leisure; Non-social</th>
</tr>
</thead>
<tbody>
<tr>
<td>Help get brothers and</td>
<td>Help round house</td>
</tr>
<tr>
<td>sisters ready</td>
<td>Cook meals</td>
</tr>
<tr>
<td>Look after people</td>
<td>Look after pets</td>
</tr>
<tr>
<td>Do things for others</td>
<td>Gardening</td>
</tr>
<tr>
<td></td>
<td>Babysit</td>
</tr>
<tr>
<td></td>
<td>Washing</td>
</tr>
<tr>
<td></td>
<td>Go to shop</td>
</tr>
<tr>
<td></td>
<td>Prepare Sunday-School lesson</td>
</tr>
<tr>
<td></td>
<td>Do paper-round</td>
</tr>
<tr>
<td></td>
<td>Decorating</td>
</tr>
</tbody>
</table>

Continued ...
(b) **School Activities**

<table>
<thead>
<tr>
<th>(1) <strong>Pro-school; Social</strong></th>
<th>(2) <strong>Pro-school; Non-social</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ask and answer questions</td>
<td>Listen to teachers</td>
</tr>
<tr>
<td>Sit next to friend</td>
<td>Do own work</td>
</tr>
<tr>
<td>Lend things</td>
<td>Do what told</td>
</tr>
<tr>
<td></td>
<td>Read</td>
</tr>
<tr>
<td></td>
<td>Write</td>
</tr>
<tr>
<td></td>
<td>Draw</td>
</tr>
<tr>
<td></td>
<td>Make things</td>
</tr>
<tr>
<td></td>
<td>Do homework</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(3) <strong>Anti-school; Social</strong></th>
<th>(4) <strong>Anti-school; Non-social</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Talk</td>
<td>Doodle</td>
</tr>
<tr>
<td>Have a laugh</td>
<td>Look out window</td>
</tr>
<tr>
<td>Annoy teachers</td>
<td>Do not pay attention</td>
</tr>
<tr>
<td>Interrupt</td>
<td>Ignore teachers</td>
</tr>
<tr>
<td>Mess around</td>
<td>Make noises</td>
</tr>
<tr>
<td></td>
<td>Eat</td>
</tr>
</tbody>
</table>
(2) Comparison-others at home and at school

The characteristics of each comparison-other produced in the four conditions (home-similar; home-different; school-similar; school-different) were noted for each subject. The three characteristics examined are (a) age; (b) sex; and (c) relationship with the subject.

Though the data for each of these characteristics is frequency data, the \( \chi^2 \) test is not appropriate for its analysis, for two reasons. First, subjects were constrained in the number of comparison-others they could produce (no more than 20). Second, subjects could produce more than one comparison-other for each of the response categories.

However, ANOVA tests could be performed on the data, using a 2 (male subjects vs female subjects) x 2 (home vs school) x 2 (similar comparison-others vs different comparison-others) design with number of comparison-others as the dependent variable. The mean number of comparison-others produced in each condition is shown in Table 2.4. It is worth pointing out that the means are considerably less than 5, which suggests that there is not a ceiling effect on the number of comparison-others produced. Because the frequencies in each condition vary slightly, the raw data is converted into percentages for the purposes of analysis.

TABLE 2.4

Mean number of comparison-others produced in each condition

<table>
<thead>
<tr>
<th></th>
<th>Home</th>
<th></th>
<th>School</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Similar</td>
<td>Different</td>
<td>Similar</td>
</tr>
<tr>
<td>Male subjects</td>
<td>3.44</td>
<td>3.75</td>
<td>2.94</td>
</tr>
<tr>
<td>Female subjects</td>
<td>3.56</td>
<td>3.69</td>
<td>3.19</td>
</tr>
</tbody>
</table>
(a) Age of comparison-others

For the purposes of analysis, comparison-others are divided into two age-categories:

(a) those who were within one year of age of the subject are placed into a 'same age' category.

(b) those who were more than one year older or more than one year younger than the subject are placed into a 'different age' category.

Obviously, the percentage of comparison-others who are a different age from the subject, and of those who are the same age as the subject, must add up to 100. Therefore an analysis is performed on the same age comparison-others only. The mean percentage of same-age comparison-others chosen by each subject in each condition is shown in Table 2.5(a), and the ANOVA results are shown in Table 2.5(b).

<table>
<thead>
<tr>
<th></th>
<th>Home</th>
<th>School</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Similar</td>
<td>Different</td>
</tr>
<tr>
<td>Male subjects</td>
<td>50.42</td>
<td>45.52</td>
</tr>
<tr>
<td>Female subjects</td>
<td>35.00</td>
<td>31.98</td>
</tr>
</tbody>
</table>

As is shown clearly by Table 2.5(a), there is a huge effect for home vs school ($F; 1, 30 = 123.28, p < 0.01$) such that nearly all the comparison-others at school are of the same age whereas at home a substantial proportion are of a different age. None of the other effects even approach significance. This supports the hypothesis that the comparison-others at school are different from those at home.
### TABLE 2.5(b)
Percentage of same-age comparison-others: ANOVA results

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex of Subject (A)</td>
<td>1862.94</td>
<td>1</td>
<td>1862.94</td>
<td>2.02</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error A</td>
<td>27672.00</td>
<td>30</td>
<td>922.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home vs School (B)</td>
<td>110940.52</td>
<td>1</td>
<td>110940.52</td>
<td>123.28</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>AB</td>
<td>1500.97</td>
<td>1</td>
<td>1500.97</td>
<td>1.67</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error B; AB</td>
<td>26997.30</td>
<td>30</td>
<td>899.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Similar vs Different (C)</td>
<td>179.74</td>
<td>1</td>
<td>179.74</td>
<td>0.46</td>
<td>n.s.</td>
</tr>
<tr>
<td>AC</td>
<td>0.20</td>
<td>1</td>
<td>0.20</td>
<td>0.00</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error C; AC</td>
<td>11725.20</td>
<td>30</td>
<td>390.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BC</td>
<td>80.77</td>
<td>1</td>
<td>80.77</td>
<td>0.22</td>
<td>n.s.</td>
</tr>
<tr>
<td>ABC</td>
<td>23.65</td>
<td>1</td>
<td>23.65</td>
<td>0.07</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error BC; ABC</td>
<td>10800.00</td>
<td>30</td>
<td>360.02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Sex of comparison-others

The percentage of comparison-others who are the same sex as the subject is analysed. The mean scores are shown in Table 2.6(a), and the ANOVA results in Table 2.6(b).

### TABLE 2.6(a)
Percentage of same-sex comparison-others: Mean scores

<table>
<thead>
<tr>
<th></th>
<th>Home</th>
<th></th>
<th>School</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Similar</td>
<td>Different</td>
<td>Similar</td>
<td>Different</td>
</tr>
<tr>
<td>Male subjects</td>
<td>83.75</td>
<td>60.00</td>
<td>91.43</td>
<td>90.67</td>
</tr>
<tr>
<td>Female subjects</td>
<td>73.02</td>
<td>64.17</td>
<td>91.67</td>
<td>88.33</td>
</tr>
</tbody>
</table>
As can be seen in Table 2.6(a), more of the comparison-others at school are of the same sex than are those at home. This effect is significant \((F; 1, 30 = 25.95, p < 0.01)\). However, well over half the comparison-others in all conditions are of the same sex. More of the similar comparison-others are of the same sex than of the different ones \((F; 1, 30 = 5.45, p < 0.05)\), but this tendency is somewhat greater at home than at school as shown by the almost-significant interaction \((F; 1, 30 = 3.41, p < 0.10)\). As for the age distribution of comparison-others, there is only a small non-significant effect due to the sex of the subject, and the largest effect by far is due to the home vs school dichotomy.

**TABLE 2.6(b)**

**Percentage of same-sex comparison-others: ANOVA results**

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex of Subject (A)</td>
<td>149.91</td>
<td>1</td>
<td>149.91</td>
<td>0.18</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error A</td>
<td>25587.00</td>
<td>30</td>
<td>852.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home vs School (B)</td>
<td>13173.08</td>
<td>1</td>
<td>13173.08</td>
<td>25.97</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>AB</td>
<td>39.90</td>
<td>1</td>
<td>39.90</td>
<td>0.08</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error B; AB</td>
<td>15215.70</td>
<td>30</td>
<td>507.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Similar vs Different (c)</td>
<td>2693.78</td>
<td>1</td>
<td>2693.78</td>
<td>5.45</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>AC</td>
<td>303.75</td>
<td>1</td>
<td>303.75</td>
<td>0.61</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error C; AC</td>
<td>14836.50</td>
<td>30</td>
<td>494.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BC</td>
<td>1625.50</td>
<td>1</td>
<td>1625.50</td>
<td>3.41</td>
<td>&lt;0.10</td>
</tr>
<tr>
<td>ABC</td>
<td>610.23</td>
<td>1</td>
<td>610.23</td>
<td>1.28</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error BC; ABC</td>
<td>14286.00</td>
<td>30</td>
<td>476.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(c) Relationship of comparison-others to subjects

Here a simple division is made between whether or not the comparison-other is a member of a subject's family. Table 2.7(a) shows the mean percentage of comparison-others who are members of the subjects' families, for each condition; and Table 2.7(b) shows the ANOVA results. Again it is extremely clear that the largest effect

<table>
<thead>
<tr>
<th></th>
<th>Home Similar</th>
<th>Home Different</th>
<th>School Similar</th>
<th>School Different</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male subjects</td>
<td>47.50</td>
<td>47.40</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Female subjects</td>
<td>61.04</td>
<td>58.65</td>
<td>2.22</td>
<td>1.25</td>
</tr>
</tbody>
</table>

is for the home vs school dichotomy ($F; 1, 30 = 73.11, p < 0.01$). None of the other terms even approach significance. Most of the comparison others who are not members of a subject's family are friends of the subject.
TABLE 2.7(b)
Percentage of comparison-others who are members of the subjects' families: ANOVA results

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex of subject (A)</td>
<td>1597.61</td>
<td>1</td>
<td>1597.61</td>
<td>1.35</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error A</td>
<td>356.33</td>
<td>30</td>
<td>1187.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home vs School (B)</td>
<td>8913.45</td>
<td>1</td>
<td>8913.45</td>
<td>73.11</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>AB</td>
<td>909.03</td>
<td>1</td>
<td>909.03</td>
<td>0.75</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error B; AB</td>
<td>3657.24</td>
<td>30</td>
<td>1219.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Similar vs Different (C)</td>
<td>24.09</td>
<td>1</td>
<td>24.09</td>
<td>0.05</td>
<td>n.s.</td>
</tr>
<tr>
<td>AC</td>
<td>21.30</td>
<td>1</td>
<td>21.30</td>
<td>0.05</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error C; AC</td>
<td>1372.00</td>
<td>30</td>
<td>457.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BC</td>
<td>4.66</td>
<td>1</td>
<td>4.66</td>
<td>0.01</td>
<td>n.s.</td>
</tr>
<tr>
<td>ABC</td>
<td>3.48</td>
<td>1</td>
<td>3.48</td>
<td>0.01</td>
<td>n.s.</td>
</tr>
<tr>
<td>Error BC; ABC</td>
<td>12352.20</td>
<td>30</td>
<td>411.74</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(d) Summary of results; Comparison-others

Taking the three dichotomous characteristics of comparison-others together, there is considerable evidence to suggest that the comparison-others used at home are different from those used at school. At school, nearly all the comparison-others are of the same sex, the same age, and are not related to the subjects. At home, however, many more comparison-others are of a different age and sex from the subjects, and around half of them are members of the subjects' families. Sex-differences between subjects are virtually non-existent, except, of course, that males mainly
choose male comparison-others at school, and females choose female ones. Differences between similar and different comparison-others are also very small though there is a tendency for similar comparison-others to be more often of the same sex than are different comparison others.

Thus the evidence overwhelmingly supports the suggestion that subjects use different comparison-others in different environments. It may well be the case, therefore, that the very criteria upon which self-assessments (and SE) are based change from one situation to another.
DISCUSSION

In the present study, two assumptions which permeate the literature on SE are found to be false. The first such assumption is that the content of SE does not vary from one environment to another. One manifestation of this assumption in the literature is the failure of experimenters to specify the situation to which subjects should address themselves when filling in SE questionnaires. In the study reported here, 58.70% of the constructs used by a quarter or more of the subjects are used significantly more often either at home or at school. Now it may be the case that such variation in content of SE has no implications for level of SE: people may have highly similar levels of SE in different situations, even though the relevant dimensions of the self concept vary from one situation to another. On the other hand, it may be that SE level does vary across situations partly as a result of variation in SE content. A limitation of the present study is that SE level is not assessed. The next two chapters address the issue of variations in SE level across situations.

The second assumption shown to be false in this study is that the people with whom a subject compares himself remain constant across situations. Again, this assumption is manifest in the literature by the failure of experimenters to specify the others with whom subjects should compare themselves when making their self-ratings (despite the exhortations of Wylie, 1974). The comparison-others data show that comparison-others vary dramatically from home to school, and also between subjects. At school, virtually all of the comparison-others are of the same sex as the subject; are of the same age; and are not members of the subjects' families. At home, however, approximately half the comparison-others belong to the subjects' families and are of a different age; and many more than at school are of the different sex.

It has been suggested that, by comparing himself with others, a person is able to reach conclusions about his own abilities and skills:
whether he regards himself as intelligent or stupid, for example, may largely depend upon whether he compares himself with people who are more intelligent than he is, or with people who are less intelligent than he is. As mentioned in Chapter 1, Festinger (1954) postulates that people have a desire to evaluate themselves; that they do this by comparing themselves with others when (as is often the case) objective information is not readily available; and that they select as comparison-others those whose ability is equal or slightly inferior to their own, in order to maintain and enhance SE. In this respect, it is interesting to note that not a single person compared himself to a teacher. Though much work remains to be done in the area of comparison-others and how, precisely, they affect a person's self-ideal discrepancies (SE), enough has been discovered to show that it is unjustified to assume that comparison-others do not vary across situations. In the remaining studies of this thesis, the comparison-others are always specified, though in fairly general terms.

Thus two recommendations may be made to researchers in this area:

(a) that a situation be specified for subjects to consider whilst they are making their self-ratings; and

(b) that the comparison-others should always be specified.

The constructs used to describe the self appear to be closely related to the activities performed at home and at school, though it must be stressed that the analysis of the activities is both subjective and post-hoc. Nevertheless, such descriptive examinations of data are not without merit, particularly in an exploratory study like this one, because they may lead to the formulation of new hypotheses which may be subsequently more rigorously tested. In the present case, the constructs used more often at school appear to be related to the pro-school vs anti-school activities, and those used more often at home to leisure vs non-leisure activities. At both home and school schoolchildren are surrounded
by other people, and many of the constructs used in both these environments relate to a social vs non-social dimension.

At home, constructs concerned with personal appearance are very salient, including such aspects as a person's looks; whether or not he is tidy; how he dresses; and his age. In addition, arguing is salient as are relationships with parents and siblings. Having friends round, and going out or staying in are also important constructs.

At school, many of the constructs used are very different from those used at home. Not surprisingly, they are largely connected with a person's intellectual and academic abilities: whether a person is clever, works hard, does well at school, and is obedient or whether he fools around, is talkative, does not try hard and is stupid.

Many of the constructs which do not discriminate between home and school are concerned with social interaction - for example, friendship, mixing with others, speaking to people and relations with the opposite sex. In both environments there are usually adults in authority: the parents in one and the teachers in the other. Subjects seemed to feel that one place was as bad as the other as regards being told off or being told what to do, and felt that standing up for themselves was important both at home and at school. In addition, whether or not a person was satisfied with himself was salient at both home and school - though the fact that many of the other constructs were different in the two environments suggests that satisfaction with the self for a subject might vary across them.

In sum, this introductory study has demonstrated that two of the assumptions which permeate the literature on the self and SE are probably unjustified. If both the constructs focussed upon are different in different situations, and the comparison-others are different in different situations, level of SE may well vary across situations as a direct result. If, therefore, the situation and comparison-others are not
specified for subjects when they are filling in self-rating questionnaires, it may well be that they imagine different situations and use comparison-others with different characteristics. Such variation can only lead to, perhaps considerable, measurement error. If level of SE does affect behaviour (for example, affecting how people react to feedback information), such measurement error may lead to serious distortion of results, as pointed out in Chapter 1 (section 1.1).

The next logical step would seem to be to examine whether SE level varies across situations. The next two studies investigate this question. One study is reported in Chapter 3. The second is reported in Chapter 4, after a suitable SE questionnaire has been developed.
CHAPTER 3

Variability of Self-Esteem Across Situations
A passive role-playing experiment is reported which assesses variability in SE level across four situations. It is found that, as hypothesized, variations in SE level are closely related to variations in SE content provided that the comparison-others are held constant. In addition, a hypothesis derived from Duval and Wicklund's (1972) Theory of Objective Self-Awareness is tested and shown to be false: subjects did not experience greater self-ideal discrepancies on those aspects of themselves they thought about a great deal than on those aspects of themselves they did not think about at all. It is concluded: (a) that it is unjustified for researchers measuring SE to assume that it does not vary intraindividually from one situation to another; and (b) that variability in SE level is not totally random.
INTRODUCTION

In Chapter 2 it was shown that the content of SE, and the comparison-others used by schoolchildren differ in the two environments of home and school. If comparison-others and content of SE vary across situations, it is quite possible that level of SE will also vary as a direct result. Variation in the comparison-others could mean that the criteria used for judging self-assessments or 'ideals' vary across situations leading to changes in SE level. Variation in the content of SE could mean that the SE level associated with those aspects of the self focussed upon in one situation is not the same as the SE level associated with those aspects of the self focussed upon in another situation. It may well be, as suggested by Gergen (1977), that there is little possibility of an exact science of the self because the self is continually changing and evidence from memory can be dredged up to support any view of oneself. That is: 'there is little in the way of factual data that lies waiting to be discovered, explored, or assessed' (p.166).

Gergen's view is rejected here. Instead, it is suggested that SE level will be similar in similar situations, and that the self-ideal discrepancies associated with particular self-assessments will be relatively constant across situations - provided, of course, that the comparison-others are held constant. In other words, whenever a particular aspect of the self is focussed upon, it will be accompanied by a particular self-ideal discrepancy. Insofar as certain situations engage similar aspects of the self, therefore, they should have similar SE level. This hypothesis is tested in the present study.

One problem in such research is to select the situations for investigation. Now the classification of situations is a thorny problem which has preoccupied psychologists for many years. It will be appreciated that there are many possible ways of classifying situations depending upon the theoretical predelictions of the investigator. Sells (1963)
and Argyle (1975, 1979), for example, prefer to classify situations in terms of their 'objective' characteristics, though the classificatory systems each ends up with are very different. On the other hand, Forgas (1976), and Magnusson (e.g. 1971) prefer to classify situations in terms of the 'subjective' characteristics through the use of multi-dimensional-scaling procedures (see, for example, Forgas, 1979).

It is not the intention of this thesis to attempt an exhaustive and thorough classification of situations. Nevertheless, some kind of classification needs to be made in order to select situations for the purpose of examining SE level within them. As Mischel (1973) points out, by a judicious selection of situations virtually any view about the variability of a personality characteristic may be supported. To take an obvious example, a situation in which people are continually provoked is more likely to produce aggression than one in which people are not. From this it may be concluded that aggression is not a personality trait because it varies from one situation to another. This would be an inappropriate conclusion: firstly, because rank-order of aggression within a sample of people may be identical across the two situations, though 'amount' of aggression may vary; and secondly because had two situations been chosen which had a similar degree of provocativeness in them, aggression levels may well not have varied across the two situations. Thus contradictory conclusions may be reached depending upon the situations selected for study.

How might this trap be avoided? It is unfortunately impossible to measure SE in every situation of a person's life, so situations must somehow be sampled and in such a way as to avoid misleading conclusions. In the present thesis, two partial solutions to the problem are used. First, two studies are used employing different situations which thus increases the generality of the findings. Second, situations in which the SE of people in general should be low are not contrasted with
situations in which the SE of people in general should be high. For example, if SE is formed through a perceived inability to live up to one's goals (via mechanisms such as heavy criticism from significant others; failure on tasks; and comparing oneself with very successful others), then situations in which people fail should be associated with lower SE than those in which people succeed. Such potent variations in situations are avoided for the purposes of examining variations in SE. Instead, the crucial classificatory principle for the selection of situations for the present purpose is taken to be the activities performed in them. Different activities mean that different personal qualities and skills will be focussed upon (ie. brought into conscious awareness), and therefore it follows that the self-ideal discrepancies themselves may differ when associated with situations involving different activities.

Two sets of activities in particular are examined: (a) social activities and (b) intellectual activities. It is perfectly possible that a person may be an intellectual wizard having HSE in situations where intellectual activities are performed, but be a social incompetent, having LSE in situations where social activities are performed – or vice versa – because in the two types of situations different aspects of the self will be focussed upon.

However, the main hypothesis (a more powerful one) to be tested in the study of this chapter, is that the less similar two situations are in the activities performed in them, the less similar they will be in the SE level associated with them. Thus, it is being suggested that not only does SE level vary across situations, but that it does so in ways which are associated with the content of SE. Variation in SE level per se is, however, examined in this study in two ways. First, by looking at correlations of SE level for each situation-pair (thus examining rank-order variation); and, secondly, by examining mean differences in SE
level (thus examining situational differences).

Comparison—others are not investigated directly in this study. Instead, subjects (schoolchildren) are asked to compare themselves with their peers when filling in the SE questionnaire.

In accordance with the methodological considerations raised in the introduction, a semantic-differential format for the SE questionnaire is used. This specifies each pole of each construct separately in order to remove the ambiguity caused by the dual meaning (intensity and ambivalence) of each point on a bi-polar scale. The 'd-score' of Osgood et al (1957) is used as the measure of SE. The questionnaire was constructed by the author so that many of the common constructs employed by the schoolchildren in the previous study could be used. The subjects in both studies are approximately the same age. The degree to which each aspect of the self is focussed upon in the different situations is measured also, by asking subjects to answer the question 'how much I think about this aspect of myself in this situation' (henceforth referred to as the 'content ratings') for each construct on the questionnaire.

The design of the present study allows another hypothesis to be tested. According to Duval and Wicklund (1972) people in general tend to have negative feelings about themselves and thus will have LSE when in situations where they think about themselves. They present some evidence for this proposition. If this proposition is true, it follows that the self-ideal discrepancies associated with aspects of the self which are thought about a great deal should be larger than those which are thought about less. This hypothesis is tested here. However, it must be stressed that this hypothesis is expected to be false, because not everyone should have LSE on aspects of themselves which are often in their conscious awareness. Though Wicklund (1975) does admit that people may have HSE he does so more as a theoretical possibility than as an empirical probability. Thus the hypothesis requires to be tested.
In sum, the present study is designed to test the hypothesis that the more similar two situations are in the aspects of the self focussed upon within them, the more similar will be the SE level generated by them. In addition, a direct test is made of the proposition, derived from Duval and Wicklund's (1972) Theory of Objective Self-Awareness, that aspects of the self thought about often are more likely to show large self-ideal discrepancies than are aspects of the self not focussed upon so often.
**METHOD**

**Subjects:** Subjects are male and female schoolchildren, aged approximately 14 years, from two classes at a Leicestershire school. Data was collected from all members of each class present on a particular day, which is 53 in all. However, two subjects failed to complete the questionnaire and so are eliminated from the analyses.

**Design:** Situations constitute the single independent variable. Four different situations are used, two of them being situations where social skills may be expected to predominate; and two being situations where intellectual skills may be expected to predominate. As well as involving different skills, the situations were chosen so that they would be familiar to the subjects, and were relevant to the home-school distinction used in the previous study.

The four situations, as described to the subjects, are:

(a) You are in your favourite lesson at school.
(b) You are in the school hall, sitting an important examination.
(c) You are at home, having an argument with your brother (or sister) about some things he (or she) has borrowed without asking and not yet returned.
(d) It is Friday evening, and you have gone round to your friend's house for a while. Several other of your friends are also there.

There are two dependent variables, as follows:

(1) Level of SE in each situation.

(2) How much each aspect of the self is focussed upon in each situation (relevance ratings).

Appendix A (p. 462) shows the actual questionnaire used in this study, along with the instructions given to subjects.
**Procedure:** The two classes were not forewarned of the experimenter's coming. When the class was assembled, the experimenter introduced herself as 'a member of the university who is carrying out some research and who would be grateful for your help'. Subjects were told that they did not have to participate if they did not wish to, but they all seemed keen to take part. After these preliminaries, questionnaires and the instructions were distributed. Each subject was given his own booklet which contained the instructions on the first page followed by four copies of the questionnaire each with a different situation specified at the top. The order of the situations within these booklets was systematically rotated.

When the questionnaires had been distributed, the experimenter read out the instructions as the subjects followed them through. The subjects were encouraged to ask questions if they did not understand the instructions. These were answered. Once subjects seemed clear about the task, they began to fill in the questionnaires at their own pace. Once all four questionnaires were completed by all subjects, the booklets were collected by the experimenter, and a discussion about the research took place.
RESULTS

(1) Split-half reliability of the questionnaire

Because the questionnaire used in this study is constructed by the author, it is important to examine its reliability. The split-half reliability, averaged across all four situations, is .72. This level of reliability is acceptably high for the purpose of basic research. As Nunnally (1967) argues 'What a satisfactory level of reliability is depends on how a measure is being used ... for basic research, it can be argued that increasing reliabilities beyond .80 is often wasteful. At that level, correlations are attenuated very little by measurement error ... . In basic research the concern is with the size of the correlations and with differences in means for different experimental treatments, for which purpose a reliability coefficient of .80 for the different measures is adequate ... . (But) in those applied settings where important decisions are made with respect to specific test scores, a reliability of .90 is the minimum that should be tolerated, and a reliability of .95 should be considered as the desirable standard' (p.226).

Though on the low side, the reliability of the questionnaire may be regarded as adequate for the present purposes.

(2) Variations in SE level across situations

There are two ways in which the consistency of SE across situations may be examined. Firstly, the extent to which an individual's rank-order of SE shifts from one situation to another. This may be measured using the correlation-coefficient. Secondly, the extent to which the mean scores for the group as a whole vary across situations. This may be measured using the ANOVA test. Both types of variation are examined here.
(a) Rank-order variation

This is examined by correlating the SE scores for each pair of situations. Obviously a correlation of 1.00 would indicated that subjects' rank order of SE remains constant. The intercorrelations of SE level across situations are shown in Table 3.1.

<table>
<thead>
<tr>
<th>Situation pairs</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Favourite lesson x Exam</td>
<td>.58</td>
</tr>
<tr>
<td>Favourite lesson x Argument with sib</td>
<td>.45</td>
</tr>
<tr>
<td>Favourite lesson x At friend's</td>
<td>.62</td>
</tr>
<tr>
<td>Exam x Argument with sib</td>
<td>.48</td>
</tr>
<tr>
<td>Exam x At friend's</td>
<td>.60</td>
</tr>
<tr>
<td>Argument with sib x At friend's</td>
<td>.49</td>
</tr>
</tbody>
</table>

From Table 3.1 it can be seen that the intercorrelations are all positive and moderately high (mean intercorrelation = .54). These results suggest that there is some cross-situational variability in an individual's rank-order of SE level. Surprisingly, however, the correlation between the two situations classified a-priori as involving mainly social activities (ie. 'at friend's' and 'argument with sib') is at the lower end of the range (r = .49), and the SE level within the situation 'at friend's' correlated highly with that of being in a favourite lesson at school (r = .62), and that of being in an important examination (r = .60). Possible explanations for this result will be presented in the discussion.
TABLE 3.2(a)
Variations in SE level across situations: Mean scores

<table>
<thead>
<tr>
<th></th>
<th>Favourite lesson</th>
<th>In exam</th>
<th>Argument with sib</th>
<th>At friend's house</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean SE level</td>
<td>6.71</td>
<td>7.67</td>
<td>8.57</td>
<td>7.12</td>
</tr>
</tbody>
</table>

1 NB: The higher the SE score, the lower the SE.

TABLE 3.2(b)
Variations in SE level across situations: ANOVA results

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Situations</td>
<td>99.48</td>
<td>3</td>
<td>33.16</td>
<td>5.70</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Error (subjects x sits)</td>
<td>873.10</td>
<td>150</td>
<td>5.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2401.97</td>
<td>203</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The situation 'having argument with sib' has the lowest correlations of SE with the other situations.

(b) Differences in mean SE level across situations

Though the correlational findings in the previous section show that a subject's rank-order of SE is not especially stable across situations, such results tell us little about whether mean SE levels for the group as a whole vary from one situation to another. In order to find out whether this occurs, an ANOVA test is performed on the data. Table 3.2(a) shows the mean SE scores, and Table 3.2(b) shows the ANOVA results.
There is a significant effect due to situations ($F; 3, 150 = 5.82, p < 0.01$). Multiple comparisons (using Tukey's 'honestly-significant-difference' technique; see Winer, 1971, p.198) reveal that the mean SE level in the situation 'argument with sib' is significantly lower than mean SE level in the situation 'in favourite lesson at school' ($q = 5.47, p < 0.01$), and the mean SE level in the situation 'at friend's house' ($q = 4.26, p < 0.05$). All other multiple comparisons are non-significant.

Thus, not only does the situation 'argument with sib' have the lowest correlations with the other situations, the subjects as a group have a lower SE level in this situation than they do in two of the others.

These results show that the rank-order of individuals in their SE level changes from one situation to another, and that some situations generate lower SE than do others. These findings, whilst hardly surprising, show that investigators cannot assume that SE level remains constant across situations when carrying out research in this area.

(3) Similarity of situations in terms of content and level of SE

The main hypothesis being tested in the present study is that situations which are similar in the content of the self they engage will be more similar in SE level than will those situations which are dissimilar in the content of the self they engage. This hypothesis rests upon the assumption that if the comparison-others are held constant across situations, then the self-ideal discrepancies associated with a particular quality also will not vary. In the present study, comparison-others are held constant across situations by asking subjects to compare themselves with the other members of their class.

In order to test this hypothesis, the correlations of SE for each situation-pair are compared with the median correlation of the content-ratings for each situation-pair. Both sets of correlations, for each
situations in terms of their content of SE are less than those in terms of their SE level, there is a very close correspondence between the two sets of correlations. Indeed, the correlation between them is .75 ($t = 2.27$, $p < 0.01$), which is significant despite having so few degrees of freedom.

Thus there is strong support for the main hypothesis being tested here, namely: that the more similar two situations are in the content of the self focussed upon, the more similar they will be in their SE level.

(4) The Duval — Wicklund hypothesis

This hypothesis is that when people think about themselves they will experience dissatisfaction because, in general, people do not live up to their goals. This hypothesis is rejected here. Instead, it is
proposed that the self-ideal discrepancies associated with particular aspects of the self may be small or large, even when the aspect of the self is thought about a great deal.

The Duval — Wicklund hypothesis may be tested by comparing the self-ideal discrepancies associated with aspects of the self which are thought about a great deal with those which are not thought about at all in a particular situation. The mean self-ideal discrepancy associated with content -ratings of '0' is compared with the mean self-ideal discrepancy associated with the content -ratings of '5', for each subject, by means of a correlated t-test. There are 37 subjects for whom this comparison could be made. The rest either did not have any content -ratings of '0' or did not have any content -ratings of '5'. The mean self-ideal discrepancy associated with content -ratings of '0' (across all 37 subjects) = 1.04. The mean self-ideal discrepancy associated with the content -ratings of '5' (across all 37 subjects) = 1.08. This difference is not significant (t = 0.08, n.s.). Thus the Duval — Wicklund hypothesis is not given any support in the present study.
DISCUSSION

From the results of this study it is clear that level of SE does not remain stable across several commonly-occurring situations. However, this variation in SE level is not random and unpredictable. Instead, it is closely related to the similarity of the situations in terms of the aspects of the self focussed upon within them - at least, when the comparison-others are held constant.

This result has important implications for research on SE. In the first place, it strongly suggests that measurement error may be reduced by specifying the situation when measuring SE level. By doing this, the investigator will be more sure that the SE scores are truly comparable across subjects. The consequences (perhaps disastrous) of not doing so are spelled out in Chapter 1 (pp. 49 - 50).

Little support for the hypothesis derived from Duval and Wicklund's (1972) Theory of Objective Self-Awareness is found in this study. The subjects did not have larger self-ideal discrepancies associated with aspects of themselves they thought about a great deal than they did for aspects of themselves they did not think about at all. This result is entirely what the present author expected because, by definition, people with HSE have small self-ideal discrepancies. Nevertheless, the Theory of Objective Self-Awareness has provided a framework which more recent investigators have used to study the effects of task-focussed versus self-focussed attention upon the performance of HSE's versus LSE's (eg. Brockner and Hulton, 1978). In such studies, it has generally been found that when LSE's have self-focussed attention their performance is worse than when they have task-focussed attention, but that this difference does not occur to such a great extent for HSE's.

Though SE level is found in the present study to vary as the aspects of the self focussed upon vary, the classificatory scheme proposed in
the introduction is not strongly supported. The suggestion was that two of the situations ('in fav. lesson', and 'in an examination') could be regarded as involving predominantly intellectual skills and that the other two ('at friend's house' and 'argument with sibling') would involve primarily social skills. Thus the correlations in SE level between the two 'social' situations should have been high, as should the correlation between the two 'intellectual' situations. This occurred for the 'intellectual' situations ($r = 0.58$) but not for the 'social' situations ($r = 0.49$). Several post-hoc explanations of this result spring to mind: (a) The a-priori classification of the situations made by the investigator does not match the classification of the situations made by the subjects. If this has occurred, it would lend support to the arguments of Forgas (1979) who regards classification of situations in terms of the subjective criteria employed by the participants within them as the best approach. He points out that the subjective dimensions employed by people for discriminating between situations may differ. (b) The results may be misleading because subjects are in a 'passive' as opposed to an 'active' role-playing situation. That is: they merely imagined the situations rather than actually took part in them. Ginsburg (1979) argues that 'passive' role-playing is more likely to produce misleading results than is 'active' role-playing. (c) SE level within these situations may well be determined by factors additional to the aspects of the self focussed upon.

Thus the main hypothesis has been supported by the present study (namely that SE level in two situations is more similar the more similar the two situations are in the content of SE they engage, given specified comparison-others). However, the classificatory scheme proposed by the investigator is not given strong support here. Instead, schoolchildren appear to have similar SE levels in the situations: being in their favourite lesson at school; being in an examination; and being at a
friend's house. Before the classification based upon an experimenter-
determined a-priori analysis of the abilities and skills needed in
particular situations is abandoned altogether, however, another study
is performed. This time, active role-playing is used (see Chapter 4).

Finally, some discussion is warranted of the SE questionnaire
used in the present study. Its reliability may be regarded as satisfac-
tory for the present purposes (Nunnally, 1967). Nevertheless it
must be admitted that the questionnaire did not include items chosen
specifically for their relevance to the social-intellectual classifi-
cation of the situations. This could be another reason for the lack
of support such a classification received in the present study. The major
part of this thesis (Chapters 5 - 8) is devoted to an examination of
the reactions of HSE's and LSE's to feedback. Therefore, an adequate
SE questionnaire is required which is not only empirically justified
(i.e. is reliable and valid) but is also theoretically satisfying.
Chapter 4 is devoted to the development of such an SE questionnaire.
In addition, another study examining the relationship between SE level
and content is presented as well as data pertaining to testing the
assumptions made by discrepancy-measures when used to assess SE.

In sum, the present study has demonstrated that SE variability
in level occurs across situations, both in terms of mean SE level for
a group of subjects and in terms of the rank-order of individuals' SE
scores within a group. In addition, this variability is shown to be
closely related to the similarity of situations in terms of the content
of SE they engage. These results therefore suggest that the assumption
of lack of variability in SE level across situations, made by many
theorists, is unjustified. But they also suggest that such variability
in SE level is not totally random, is predictable, and may therefore be
taken into account when carrying out research in this area. Finally,
the Duval — Wicklund hypothesis that self-ideal discrepancies asso-
ciated with aspects of the self focussed upon a great deal would be
larger than those associated with aspects of the self not focussed upon at all, is shown to be false.
CHAPTER 4

Development of a Self-Esteem Questionnaire
ABSTRACT

In this chapter, an SE questionnaire is developed which measures SE as defined in Chapter 1. Particular attention is given to the meaningfulness and relevance of the items included in the questionnaire for the subject populations on which it is to be used. In addition, the reliability and validity of the questionnaire is examined and found to be highly satisfactory for the purposes of basic research. Several criticisms of discrepancy measures of SE are discussed and, where possible, are tested empirically. It is found that ideal scores do contribute significantly to overall SE scores on the questionnaire. This finding is in marked contrast to the speculations of Burns (1979) who says that ideal scores are a 'relatively homogeneous constant' (p.85). It is also found that, as expected, there is a high correlation between the size of a self-ideal discrepancy and the degree of distress it is associated with. The second half of this chapter consists of an active role-playing study examining variability in SE across situations. As in Chapter 3, there is evidence both for considerable intraindividual variation in SE level across situations, and for the suggestion that SE level and content are closely related. Once again, subjects do not appear to have any lower SE for aspects of themselves focussed upon a great deal than for those aspects of themselves not focussed upon at all, within a particular situation. This is in contrast to what would be predicted on the basis of Duval and Wicklund's (1972) Theory of Objective Self-Awareness.
INTRODUCTION

In the study reported in Chapter 3, subjects were asked to imagine that they were taking part in specified situations. Their SE level, and the content of SE, was assessed in each of four situations. It was found that the variation in SE level across situations was closely related to how similar the situations were in the content of SE they engaged. Thus, providing that comparison-others are held constant, there is some evidence that the relationship between 'self' scores and 'ideal' scores remains fairly stable across situations for specific abilities and skills.

However, these results were found in what may be described as a 'passive' role-playing situation (eg. Mixon, 1976). In other words, the situations were merely described for the subjects, who then imagined what being in such a situation would be like. It is possible that these 'imaginings' bear little relationship to what subjects think when they actually take part in such situations. Though this is an important problem, it has largely been ignored by investigators working in the area of perception of situations (eg. Forgas, 1976; Argyle and Little, 1972; Pervin, 1976). As Mixon (1979) points out, the use of passive role-playing assumes that certain situations may be adequately modelled by the procedure. Yet little work has been performed to test the adequacy of such a model.

A way of increasing the generality of the findings of the previous study would therefore be to use an 'active' role-playing procedure whereby subjects actually take part in simulations of the situations described by the experimenter. In the study reported in this chapter, such active role-playing is used. Subjects actually took part in four situations, and after each one rated their SE level and content. As in Chapter 3, two main hypotheses are tested. These are:

(a) That the more similar two situations are in the content of the
self they engage, the more similar they will be in the level of SE
associated with them (given that the comparison-others are specified and
held constant across the situations); and

(b) The Duval — Wicklund (1972; 1975) hypothesis that self-
ideal discrepancies associated with aspects of the self focussed upon a
great deal will be greater than the self-ideal discrepancies associated
with those aspects of the self not focussed upon at all. This hypothesis
was not supported in the previous chapter, and is not expected to be
supported here, either.

The first part of this chapter is, however, devoted to the develop-
ment of a reliable and valid measure of SE which is in accordance with
the theoretical definition of SE proposed in Chapter 1. That is: the
questionnaire should measure both self-assessments and ideals for each
self-relevant construct. In addition, the items included in the question-
naire should be relevant and meaningful to the subjects. An attempt was
made to ensure relevance by obtaining lists of the relevant abilities
and personal qualities in a specified number of situations (situations
which would be similar to those used in later studies) from a large sample
of the population upon which the questionnaire would later be used.
Meaningfulness was assumed once steps were taken to ensure that the wor-
ding of the items corresponded to the phrases used by the sample.

Wylie (1974) and Burns (1979) appear to regard discrepancy measures
as unnecessary and of low validity when used as measures of SE. Conse-
quently, some attention is paid to their criticisms, and empirical evi-
dence is presented which shows them to be unjustified.
1. Development of the SE Questionnaire

INTRODUCTION

The aim behind the development of this questionnaire is to arrive at a theoretically satisfying and empirically justified measure of SE. To be theoretically satisfying, the questionnaire must measure SE as defined in Chapter 1 of this thesis. This means that both 'self-assessments' and 'ideals' should be measured for each aspect of the self, and the discrepancy between them used as a measure of SE. As pointed out in Chapter 1, many measures of SE rest on the assumption that self-assessments are equivalent to SE. This, however, would only be true if everyone held the same values. But why make such an assumption when it is manifestly not true? Rosenberg (1965) shows, for example, that different adolescents value different personal qualities, and Rokeach's (1968; 1973) research on values is premised upon the assumption that people's values differ. Thus the measurement of ideal scores is strongly recommended, in addition to the measurement of self-assessments.

Wylie (1974) and Burns (1979) are highly critical of self-ideal discrepancy indices when used as measures of SE. In the first place, Wylie points out that the theory of meaning embodied in the work of Osgood et al (1957) was never intended to apply, and has never been shown to apply, to SE. This criticism, however, is rejected as a reason for not using the semantic-differential as a measure of SE. The semantic-differential may be regarded, quite simply, as a questionnaire format - a form of rating scale - and as such is independent of Osgood et al's theory of meaning.

Secondly, Wylie criticises researchers for not making clear their criteria for item selection, and she questions the appropriateness of using a standard semantic-differential set of items developed from the
factor analyses Osgood et al performed on their data. This argument is wholly accepted here. But it does not mean that self-ideal discrepancies per se are a poor measure of SE. All it means is that care must be taken in the selection of items, and this obviously applies to the construction of any questionnaire. Efforts are made in the present chapter to make very clear the criteria used for selection of items, and to convince the reader that the items finally included are relevant and meaningful for the subject populations upon which the questionnaire is based.

A third criticism of both Wylie (1974) and Burns (1979) is that sometimes the 'self' score may be 'greater than' the 'ideal' score, and that this does not have the same meaning as a discrepancy where the 'self' score is 'less than' the 'ideal' score. This argument, taken one stage further would state that two self-ideal discrepancies of equal magnitude but at different points on a scale mean different things. And, of course, that two discrepancies for the same item but for different subjects mean different things. All these points are undoubtedly true. I cannot understand why Wylie singles out the first as being a weakness, but not the other two. They are all major problems in attitude measurement, and undoubtedly make a substantial contribution to measurement error. All that can be said in defence of these criticisms, as far as SE is concerned, is that if the predicted theoretical results are obtained using a particular SE measure, then the measurement error is unlikely to be completely drowning the 'true score' variance. According to the theoretical notions presented in Chapter 1, it is the size of the discrepancy which is important in SE measurement - not whereabouts upon the scale the 'self' and 'ideal' scores lie. It is the chronic 'sense of failure', associated with large self-ideal discrepancies, which motivates LSE's to behave in certain ways in certain circumstances. For HSE's, it is the chronic 'sense of success', associated with small self-ideal discrepancies, which is the motivating factor. Under other circumstances, SE level may not be a factor governing behaviour. Instead,
other forces may dominate - such as the 'actual' goals pursued or the 'actual' self-assessments. In case these counter-arguments should fail to satisfy, the percentage of 'self' scores 'greater than' the 'ideal' scores will be examined. It should be remembered that the only real counter-argument to this criticism is the construct validity of the questionnaire. This will be demonstrated by the experimental results presented in later chapters.

A fourth criticism put forward by Wylie is that self-ideal discrepancy measures assume there is a monotonic relationship between the size of a discrepancy and the degree of discomfort associated with it. In other words, that the larger the discrepancy the greater the degree of discomfort associated with it, no matter which aspect of the self is involved. But it may be the case that some aspects of the self are regarded as much more important to a person than are other aspects, and that a discrepancy of a particular size on an important aspect is associated with much greater distress than a discrepancy of the same size associated with a relatively unimportant aspect of the self. Wylie points out that the monotonicity assumption may be tested very easily. All that is needed is to ask subjects to indicate how distressed they are for the self-ideal discrepancy on each item. The Degree-of-Distress scores, and the self-ideal discrepancy scores can then be correlated. A high correlation would support the assumption of a monotonic relationship.

Fifth, both Burns and Wylie suggest that the discrepancy score may be a poor measure of SE because there is little intersubject variability in ideal scores. Burns (1979), for example, says 'But there is some question whether the discrepancy is satisfactory from a measurement point of view and whether it is really all that different from the topic of positive and negative self-concepts. Firstly, it is not known how much variance is contributed by each part to the variance in scores on this dual index. Secondly, there is considerable evidence ... that there is
very little intersubject variability on ideal self-concept ratings.' (p.85). These two issues are examined empirically once the SE questionnaire has been developed. The correlations of the self scores and the ideal scores with the SE scores will be measured, in order to discover how much variance in the latter is accounted for by each of the former two. Intersubject variability on ideal scores needs to be re-examined as there is no guarantee results found for different measures will be identical. Thus intersubject variability in the ideal scores will be compared with intersubject variability in the self scores.

Two points need to be borne in mind here, however. Firstly, even if intersubject variability in ideal scores is low, the fact that subjects have been asked to articulate a standard against which their self-scores are assessed, may exert a considerable influence on the resultant SE scores. Secondly, if the ideal scores do not contribute a large proportion of variance to the overall SE scores, but the self-scores do, this does not mean that the self-ideal discrepancy index is a poor measure of SE. It could equally well mean that self-scores per se are as good at measuring SE as are self-ideal discrepancies.

In sum, the reasons given by Wylie (1974) and Burns (1979) for not using self-ideal discrepancies as a measure of SE are either unfounded on theoretical grounds, or are open to empirical test. Empirical tests of these authors' criticisms are performed in a later section of this chapter.

A prime consideration is to ensure that the items selected for inclusion in the questionnaire are meaningful and relevant: not only for the subjects upon which the questionnaire is to be used, but also for the experimental situations in which subjects are placed. This is achieved by performing a content-analysis upon the free-descriptions given by individuals of the abilities and personal qualities they thought would be useful in certain situations. The most frequently-mentioned
characteristics are then incorporated into a questionnaire of moderate length. Only one pole of each construct is used in order to remove the ambiguity created by ambivalence (see Kaplan, 1972). Thus the final format is a modified semantic-differential type Likert scale. The use of both poles, separately rated, would double the length of the questionnaire and make it unwieldy. Finally, the instructions accompanying the questionnaire are devised so that both the comparison-others and the situations may be specified.
1. The free-descriptions

Subjects: There are 78 subjects altogether. 58 of these are university students who were recruited in their halls of residence or in coffee bars around the university. The remaining 20 are student nurses who lived in a hall of residence not far from the university. These are included because it is anticipated that in a later experiment student nurses will be used as subjects. Half the subjects were male and half were female.

Design: Six contexts were drawn up with respect to which subjects are asked to provide free descriptions of the abilities and personal qualities which would be useful in these contexts. Each subject is asked to provide descriptions for two of the contexts, such that there were an equal number of descriptions provided for each of the six contexts; and for each context half the descriptions were provided by males and half by females. Order of presentation of the contexts was carefully rotated in order to cancel out possible order-effects such as the mentioning of particular characteristics in the second context by a subject merely because he/she mentioned it in the first context.

So there are 13 males and 13 females providing descriptions for each of the six contexts.

Procedure: A student was approached and asked if he or she had 5 minutes or so to spare to help in some psychological research. Several people refused, all of them on the grounds that they were just on their way to a lecture or seminar. The rest provided descriptions for two contexts and everything they said was written down. The subject was then thanked and any questions he or she had about the research were
answered. The exact instructions to the subject were as follows:

'Excuse me, but I'm trying to develop a questionnaire concerning what people think about various everyday situations. I wonder if you could spare me 5 minutes or so, to answer a couple of questions?'

If the person agreed, he or she was asked to imagine one of the following six contexts:

(1) Imagine you are writing, as part of your course-work, a synopsis (including an interpretation) of an important journal article. Your synopsis will be marked, and will count towards your final mark in the examinations.

(2) Imagine that as part of your course work you've been asked to form a group of 4 people with other students in your year, and to write a comprehensive synopsis (with interpretation) between you of an important area in your subject. As a group, you have to present the main arguments at a seminar and this will count towards your marks in the exams. It is therefore important that you all agree on what to say. At the moment you are all together discussing what to put in the synopsis.

(3) Imagine you are waiting outside a lecture room, or seminar room, and you are having an enjoyable chat with one of your friends (of the same sex).

(4) Imagine you are getting to know someone (of the same sex). For example, whilst waiting to go in for a lecture, or over dinner in hall, or at a social occasion.
(5) Could you describe in detail the qualities a person would have if he/she were extremely socially skilled (i.e. is very good at getting on with other people) - for example, is very good at parties, at chatting to people during coffee breaks, and so on?

(6) Could you describe in detail the qualities a person would have if he/she were extremely intellectually skilled - for example, is very good at his/her course work, does very well in exams and so on?

With the exception of numbers (5) and (6) which are direct questions in themselves, the subject was asked to answer the questions:

(1) 'Could you describe in detail what you are really like when you are taking part in this situation?'

and

(2) 'What abilities and personal qualities do you think would be useful in this situation?'

The first question was an attempt to get the subject thinking about the situation. The qualities and abilities were derived from the second question. Everything the subject said was written down. Then he/she was asked to repeat the performance for another scenario.

2. The content-analysis

The aim of the content-analysis is to provide a list of the most frequently occurring abilities and qualities in each of the six contexts, thus allowing a rational choice to be made of items for inclusion into the final questionnaire. Such items should therefore be both relevant and meaningful, provided that the content-analysis is reliable. The
following method was employed:

(a) First of all, each phrase generated by the subjects was written onto a separate index card. A record was kept on each card of the contexts which generated every occurrence of the phrase, and by which subject. If a phrase was used more than once by a particular subject with reference to a context, it was counted as having occurred once only. Therefore, because there are 26 subjects describing each context, the maximum number of recorded occurrences for each phrase is 26.

(b) The procedure was checked for reliability by having a second judge read through a random sample of 25% of the free-descriptions. The second judge also kept a record of the number of contexts per subject in which each phrase occurs. Inter-judge reliability may be assessed by correlating the number of each of the items derived by each judge from the sample of descriptions.

(c) The next stage consisted of grouping the phrases so that those with a similar meaning are grouped together. For example, 'extroverted' and 'outgoing', originally put on separate cards, would now be grouped together. By combining items into categories it is possible to represent more of them in a questionnaire of moderate length, and to avoid duplication of items with very similar meaning but which are both very frequently mentioned. Obviously, such a grouping procedure must be checked for its reliability. Two possible methods were considered, as follows:

(i) Having an independent rater also sort the items into categories of his/her own choice. Brennan and Light (1974) provide a measure which shows whether agreement between two raters using this method is greater than chance. Such a procedure has disadvantages, however. Firstly, agreement may be reduced because raters use different numbers of categories. Judges are known to differ in the number of categories they use to describe a particular set of data. Secondly, they may differ in the criteria used to assign items to a particular category and thus end up with different
categories. Yet the judges may well be aware that different criteria for categorisation are possible, and merely have chosen different ones, thus apparently completely disagreeing but in reality showing some agreement. For these two reasons, this method of checking the grouping procedure was rejected.

(ii) Providing a second rater with a list of predefined categories and asking him/her to sort the items into these categories. This procedure is more structured than the previous one because differences between the judges in their degree of differentiation are to some extent ruled out. Consequently, the obtained measure of agreement will be a more realistic reflection of the actual agreement between the judges. This method is therefore chosen here. The appropriate measure of agreement to use for this task is Cohen's Kappa coefficient (Cohen, 1960).

In order to define the categories, two judges discussed the data in some detail. The items were then given to two other judges who were asked to sort the items into the agreed-upon categories. The criterion for inclusion into the questionnaire is as follows: all categories are to be included where the two independent judges agree that more than a third of the 26 subjects (i.e. 9 or more) mention it in a particular context. The most common phrase in each category is the one used to represent that category in the final questionnaire.
RESULTS

1. Extraction of items from the free-descriptions

Agreement at this stage of the content analysis is extremely high. The task proved to be easy and unambiguous. Two independent judges extracted items from the same random sample of 25% of the free descriptions (ie. 39). The correlation between the two raters for the number of items extracted from each of the free descriptions is 0.94 (p < 0.01).

Altogether, 346 items were noted by the experimenter. The mean number of items per description is 6.13.

2. Categorization of the items

The discussion between two raters (one being the author) led to 39 categories plus another labelled miscellaneous. The 39 meaningful categories accounted for 293 of the 346 items ie. 84.7% of them.

Two independent judges then separately sorted the 346 items into the 40 categories. Agreement between the judges was found using Cohen's Kappa (1960). The Kappa Coefficient is 0.68, which is highly significant (Z = 60, p < 0.001). This means that there is considerable above chance agreement between the two judges in the partitioning of the items into the categories. It should be noted that Kappa and not Weighted Kappa (Cohen, 1968, 1972) is used. The former provides the lower bound of agreement between the two judges for it counts as disagreement every case where one judge places an item into a different category from the other judge. It does not take into account the possibility that some of the categories may be related. If two categories are related, a judge may place an item into one of them but may also wish to place it into the other. Therefore he would agree with another judge who placed it into the other, related, category. The .68 agreement obtained here is regarded as satisfactory, and hence Weighted Kappa is not assessed - it could only
show greater agreement!

The 39 meaningful categories account for 90.5% of the items for the first judge, and for 95.4% of the items for the other. Most of the data is therefore accounted for by the category system employed.

3. Selection of items for the questionnaire

Categories are included in the questionnaire if they satisfy the a-priori criterion that the two judges agree that more than a third of the subjects (≥9) used the category in at least one context.

Table 4.1 gives details of the categories used and their frequency of usage by different subjects according to the two judges. As can be seen, the two judges have very high agreement on the importance (frequency of usage) of a particular category relative to the others. Indeed, the averaged correlation for the frequency of use of the categories by the two judges is 0.87.

The frequency criterion results in the choice of 23 of the 39 categories. For two further categories (no. 11: 'make other feel at ease'; and no. 37 'feel good') at least one of the judges thought that 9 or more subjects used it in a particular situation. However, there was considerable confusion between these two categories and it was decided to use only one of them, the one which was used most (ie. no.11). The final questionnaire, therefore, consists of 24 of the most frequently-used categories out of the 39 possible ones, each category being represented by the most commonly-occurring phrase within it.
### TABLE 4.1

Content-Analysis Categories and their Frequency of usage (by different subjects) According to Two Independent Judges

<table>
<thead>
<tr>
<th>Category</th>
<th>1*</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A**</td>
<td>B**</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>1. Able to think logically</td>
<td>13</td>
<td>14</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2. Able to use background knowledge</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3. Work hard</td>
<td>8</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4. Tactful</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5. Understanding of issues related to work</td>
<td>26</td>
<td>26</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6. Pick up cues given by others (sensitive to others)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>7. Interested in people</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>8. Many outside interests (to talk about)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>9. Interested in work (enthusiastic about work)</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>10. Humourous</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>11. Make others feel at ease</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>12. Good memory</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13. Extrovert</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>14. Express self well</td>
<td>17</td>
<td>17</td>
<td>15</td>
<td>19</td>
<td>6</td>
<td>11</td>
</tr>
</tbody>
</table>

**Notes:**
- A** and B** indicate the frequency of usage by Subject A and Subject B, respectively.
- Each category's usage values are presented in 1 or 2 columns, depending on the number of independent judges involved.
Table 4.1 (continued)

<table>
<thead>
<tr>
<th>Category</th>
<th>Usage 1</th>
<th>Usage 2</th>
<th>Usage 3</th>
<th>Usage 4</th>
<th>Usage 5</th>
<th>Usage 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>15. Concentrate</td>
<td>8 8 0 0</td>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
<td>9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Compromise (accept new ideas)</td>
<td>2 2 26 26</td>
<td>6 8 8 10</td>
<td>10 14 14</td>
<td>8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Friendly</td>
<td>0 0 1 1</td>
<td>9 11 16 17</td>
<td>11 12 0 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Not reluctant to present own views</td>
<td>3 3 25 23</td>
<td>5 7 1 0</td>
<td>2 2 6 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Find out what other is like</td>
<td>0 1 0 1</td>
<td>0 0 13 14</td>
<td>0 1 0 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Intelligent</td>
<td>5 8 4 4</td>
<td>2 3 0 0</td>
<td>2 2 10 11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Confident</td>
<td>2 2 3 6</td>
<td>1 2 5 7</td>
<td>22 20 6 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Listen to others</td>
<td>0 0 10 9</td>
<td>13 14 10 10</td>
<td>6 7 1 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Stay calm (don't panic)</td>
<td>14 20 1 5</td>
<td>2 2 2 4</td>
<td>0 2 5 11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. Talk of things other is interested in</td>
<td>0 0 0 0</td>
<td>11 12 18 14</td>
<td>6 3 0 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. Introverted</td>
<td>0 0 0 0</td>
<td>0 0 6 7</td>
<td>0 0 4 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. Careful</td>
<td>3 5 1 0</td>
<td>0 0 0 0</td>
<td>0 0 1 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Honest</td>
<td>1 1 1 1</td>
<td>1 1 2 2</td>
<td>1 1 0 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. Helpful</td>
<td>0 0 0 0</td>
<td>4 4 1 3</td>
<td>1 3 0 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. Tense</td>
<td>1 3 2 2</td>
<td>0 0 1 1</td>
<td>0 0 0 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. Stamina</td>
<td>3 3 0 0</td>
<td>0 0 0 0</td>
<td>0 1 2 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31. Dominant</td>
<td>0 0 4 4</td>
<td>0 0 2 2</td>
<td>4 4 1 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32. Not condescending</td>
<td>0 0 2 2</td>
<td>0 0 2 2</td>
<td>8 8 0 0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4.1 (continued)

<table>
<thead>
<tr>
<th>Category</th>
<th>1 A</th>
<th>1 B</th>
<th>2 A</th>
<th>2 B</th>
<th>3 A</th>
<th>3 B</th>
<th>4 A</th>
<th>4 B</th>
<th>5 A</th>
<th>5 B</th>
<th>6 A</th>
<th>6 B</th>
</tr>
</thead>
<tbody>
<tr>
<td>33. Resilient</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>34. Leadership</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>35. Strong character</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>36. Unemotional</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>37. Feel good (emotionally)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>14</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>38. Co-operative</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>39. Physical appearance</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40. Miscellaneous</td>
<td>14</td>
<td>5</td>
<td>8</td>
<td>4</td>
<td>17</td>
<td>5</td>
<td>16</td>
<td>7</td>
<td>18</td>
<td>15</td>
<td>18</td>
<td>8</td>
</tr>
</tbody>
</table>

* Nos. 1-6 correspond to the contexts 1-6 on pp. 130 - 131.

** A = 1st judge; B = 2nd judge

A copy of the questionnaire is provided in the appendix on p. 469. The first 24 items in Table 4.1 are those included in the questionnaire. They are arranged in the order in which they appear in the questionnaire (ie. randomly).

Instructions to accompany the SE questionnaire

These were developed in order to allow the specification of the comparison-others and of the situations. Subjects are asked to rate three concepts on the same page:

(a) 'What I am really like' on a scale from 0 ('This item does not describe what I am really like at all') to 6 ('This item describes what I am really like extremely well indeed');

(b) 'What I would ideally like to be like' on a scale from 0 ('This
item does not describe what I would ideally like to be like at all') to 6 ('This item describes what I would ideally like to be like extremely well indeed');

(c) Finally, subjects are asked either to rate the concept 'How much this difference between what I am really like and what I would ideally like to be like bothers me' (on a 0: 'Does not bother me at all' to 6: 'Bothers me very much indeed' scale) or to rate the concept 'How much I think about this aspect of myself in this situation' (again, on a 0: 'I do not think about this aspect of myself at all' to 6 'I think about this aspect of myself a great deal' scale).

Subjects are asked to compare themselves with a specified group (usually their peers) when making their ratings. In addition, the situations about which they should be thinking when making the ratings are specified. Subjects are asked to proceed fairly quickly. The 'self' and 'ideal' scores are distinguished by the way in which they are rated: one circle is put around the relevant number to indicate the 'self' score, and two circles around the relevant number indicate the 'ideal' score. Subjects are told that their 'self' and 'ideal' scores could be the same number for a particular item. The third rating is placed in a column on the right-hand-side of the page, opposite the item to which it applies.

The instructions are presented in Appendix B pp. 467 - 468.
The purpose of this section is to examine the reliability and validity of the SE questionnaire constructed by the author. Validity will be assessed mainly in relation to the criticisms of the semantic-differential as a measure of SE advanced by Wylie (1974) and Burns (1979). Construct validity is left for the reader to judge after a perusal of Chapters 6-8, which report experiments carried out using the questionnaire. The present author is convinced that the questionnaire has excellent construct validity. The convergent and discriminant validity (see Campbell and Fiske, 1959) of the questionnaire is left unexamined, because it depends totally upon (a) how well other constructs are measured and upon (b) how well other measures of SE do so. There is little reason to assume that other measures of SE have high validity given both the methodological criticisms of Wylie (1974) and the low intercorrelations of other SE measures (eg. Ziller, 1973). Therefore, insofar as the questionnaire results suggest that Wylie's criticisms are empirically unjustified, and the experimental results are in line with theoretical predictions, the questionnaire will be taken as a reasonably valid indicator of SE. The assumption is made that the questionnaire will be meaningful to the subjects because it incorporates phrases taken verbatim from a sample of the subject population. The relevance of the items is dealt with in a separate study reported later in this chapter.

1. Reliability

A group of 60 student nurses were asked to fill in the questionnaire twice, with a week long interval between the two sessions. They were not told at the first session that there would be another session. The nurses were asked to compare themselves with the other student nurses in their group, and to refer to intellectual situations such as 'doing a
written test as part of your course' whilst making their ratings.

The test-retest correlation is 0.82. This is perfectly satisfactory for the purposes of basic research, which is what the questionnaire is to be used for (see Nunnally, 1967, p.226 and Chapter 3 of this thesis). Such a reliability coefficient means that correlations and differences between means will be very little attenuated by measurement error.

2. Validity

(a) Method

In order to examine validity-related issues, the questionnaire was distributed to two large samples of students: a sample of 73 first-year university students studying psychology, and a sample of 79 first-year polytechnic students studying applied social studies. In both samples the students were asked to compare themselves to the other students in their year and class, and to apply the ratings to intellectual situations such as writing an essay. In addition to their 'self' and 'ideal' scores, the students also rated the concept 'How much this difference between what I am really like and what I would ideally like to be like bothers me'.

(b) Results

(i) The percentage of 'ideal scores which are 'less than' the 'self' scores.

For the university students, data from 68 of the subjects could be used. The rest failed to fully complete the questionnaire. Only 1.35% of the 'ideal' scores are 'less than' the 'self' scores (ie. 0.32 items per subject).

For the polytechnic students, data from 69 of the subjects could be used, the rest having failed to fully complete the questionnaire. Only 3.8% of the 'ideal' scores are 'less than' the 'self' scores (ie. 0.91 items per subject).
Thus there is little evidence here that Wylie's criticism is empirically justified: the lack of discrimination in discrepancy measures between 'positive' and 'negative' discrepancies does not pose a problem because the discrepancies are overwhelmingly of the same sign. Obviously, such a conclusion only applies to the particular item-set used in the questionnaire.

(ii) Is there a monotonic relationship between the size of a discrepancy and the degree of distress it produces?

If there is such a monotonic relationship, there should be a high positive correlation between the size of the discrepancies and the rating 'How much this difference between what I am really like and what I would ideally like to be like bothers me'.

The correlation was obtained for each subject in both of the samples of students (ie. for 137 subjects). The mean correlation for the combined sample is 0.73. This is positive, significant (p < 0.01), and moderately high.

Figure 4.1 shows the frequency distribution of the correlations for the combined sample. As can be seen, none of the correlations are negative, and for 89.05% of the subjects the correlation is significant at the p <= 0.01 level (all correlations above 0.48 are significant at this level).

These results suggest that the assumption made by this discrepancy measure that there is a monotonic relationship between the size of the discrepancy and the degree of distress associated with it, holds good. Thus another of Wylie's criticisms is found to be empirically unjustified.

(iii) Is there 'significant' intersubject variability in the 'ideal' scores?

The only meaningful way to answer this question is to compare the variance of the 'self' scores for each item with the variance of the 'ideal' scores for each item, and these variances are shown, for the 68
Figure 4.1

Frequency distribution for 137 subjects of correlations between size of the self-ideal discrepancies and the rating 'how much this difference bothers me'.
university students, in Table 4.2. As can be seen, the variances of
the 'ideal' scores are always less than the corresponding variances of
the 'self' scores, but are by no means negligible. The 'ideal' scores
variances are generally about a third to half of the 'self' scores
variances.

(iv) How much variance in the SE scores do the 'self' scores and
the 'ideal' scores each account for?

(a) Contribution of the 'self' scores:

This is found by correlating the sum of the self scores with the
SE scores across subjects. It was thought that the sample of university
students alone would be large enough to provide a reliable indication of
the size of this correlation.

The correlation between the sum of the 'self' scores and the SE
scores is $-0.47$ ($p < 0.01$).

Thus the 'self' scores account for 22.09% of the variance in the
SE scores.

(b) Contribution of the 'ideal' scores:

This is found by correlating the sum of the 'ideal' scores with
the SE scores across subjects.

The correlation between the sum of the 'ideal' scores and the SE
scores is $+0.26$ ($p < 0.025$).

Thus the 'ideal' scores account for 6.75% of the variance in the
SE scores.

It is clear from these results that the 'self' scores contribute
a larger proportion of the variance in the SE scores than do the 'ideal'
scores. Two points should be borne in mind when interpreting these
results, however. In the first place, a large proportion of the variance
in the SE scores remains unaccounted for, even assuming that the contri-
bution of the 'self' scores is completely independent of the 'ideal'
scores. This is probably largely because the SE scores are not a linear
<table>
<thead>
<tr>
<th>Item</th>
<th>Variance of the 'Self' Score</th>
<th>Variance of the 'Ideal' Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Think logically</td>
<td>1.26</td>
<td>0.47</td>
</tr>
<tr>
<td>2. Use background knowledge</td>
<td>0.97</td>
<td>0.51</td>
</tr>
<tr>
<td>3. Work hard</td>
<td>2.06</td>
<td>0.88</td>
</tr>
<tr>
<td>4. Tactful</td>
<td>2.28</td>
<td>0.90</td>
</tr>
<tr>
<td>5. Understand work</td>
<td>1.16</td>
<td>0.43</td>
</tr>
<tr>
<td>6. Pick up cues</td>
<td>1.23</td>
<td>0.84</td>
</tr>
<tr>
<td>7. Interested in people</td>
<td>1.71</td>
<td>1.31</td>
</tr>
<tr>
<td>8. Many outside interests</td>
<td>2.00</td>
<td>1.13</td>
</tr>
<tr>
<td>9. Interested in work</td>
<td>1.19</td>
<td>0.54</td>
</tr>
<tr>
<td>10. Humorous</td>
<td>1.31</td>
<td>0.73</td>
</tr>
<tr>
<td>11. Make others feel at ease</td>
<td>1.44</td>
<td>0.64</td>
</tr>
<tr>
<td>12. Good memory</td>
<td>1.98</td>
<td>0.43</td>
</tr>
<tr>
<td>13. Extrovert</td>
<td>2.18</td>
<td>1.24</td>
</tr>
<tr>
<td>14. Express self well</td>
<td>1.77</td>
<td>0.43</td>
</tr>
<tr>
<td>15. Can concentrate</td>
<td>1.69</td>
<td>0.31</td>
</tr>
<tr>
<td>16. Able to compromise</td>
<td>1.46</td>
<td>0.91</td>
</tr>
<tr>
<td>17. Friendly</td>
<td>1.45</td>
<td>0.56</td>
</tr>
<tr>
<td>18. Not reluctant to present own views</td>
<td>1.86</td>
<td>1.18</td>
</tr>
<tr>
<td>19. Try to find out what other is like</td>
<td>1.03</td>
<td>0.91</td>
</tr>
<tr>
<td>20. Intelligent</td>
<td>1.55</td>
<td>0.67</td>
</tr>
<tr>
<td>21. Confident</td>
<td>1.94</td>
<td>0.60</td>
</tr>
<tr>
<td>22. Listen to others</td>
<td>1.20</td>
<td>0.58</td>
</tr>
<tr>
<td>23. Stay calm</td>
<td>2.69</td>
<td>0.88</td>
</tr>
<tr>
<td>24. Talk of things other is interested in</td>
<td>0.84</td>
<td>0.65</td>
</tr>
</tbody>
</table>
function of either the 'self' or the 'ideal' scores. Instead, SE scores depend upon the squares of the 'self-ideal' discrepancies. In the second place, the contribution of the 'ideal' scores is not negligible and cannot therefore be ignored.

Thus the argument for discarding the 'ideal' scores is empirically unjustified, both on the grounds of intersubject variability and on the grounds of the percentage variance in the SE scores contributed by the 'ideal' scores. Self-assessments are, therefore, not equivalent to SE level.

(c) Conclusions: reliability and validity of the questionnaire

The questionnaire appears to have adequate reliability for basic research purposes. In addition, it may be concluded that the criticisms of Wylie (1974) and Burns (1979), against self-ideal discrepancy questionnaires as measures of SE, are empirically unjustified - at least as far as this questionnaire is concerned. In the first place, very few responses occur (on average, around 2%) where the 'ideal' scores are 'less than' the 'self' scores. Thus, even if the assumption of Wylie is correct that discrepancies where 'ideal' scores are 'less than' 'self' scores have a different meaning from discrepancies where 'self' scores are 'less than' 'ideal' scores, this is hardly likely to significantly affect overall SE scores. This is because responses are overwhelmingly of the kind where 'self' scores are 'less than' 'ideal' scores.

Secondly, there is strong support for the assumption, made by self-ideal discrepancy measures of SE, that there is a positive monotonic relationship between the size of the discrepancy and the degree of distress it is associated with. It must be pointed out, however, that the items in the present questionnaire were carefully selected so as to be important to the subjects. Whether this assumption holds when items are unimportant to subjects remains to be tested. But presumably no-one would wish to knowingly employ such items in an SE questionnaire. Thus the
assumption has been shown to hold good under the conditions which matter - ie. where the items are important.

Thirdly, it appears that the 'ideal' scores are not totally redundant. They do make a significant contribution to the variance in the SE scores, and they do exhibit substantial intersubject variability. Therefore the conclusion of Burns (1979) in regard to self-ideal discrepancy scores, namely that 'All that appears to be happening is the subtracting of a wide range of self-concept scores from a relatively homogeneous constant.' (p.85) must be rejected. When appropriate care is taken over item-selection, self-assessments are by no means equivalent to self-ideal discrepancies (ie. SE).

In conclusion, the questionnaire so far appears to have adequate reliability and validity.
An active role-playing study of SE level and content across situations

INTRODUCTION

The previous part of this chapter has been devoted to the development of a good SE questionnaire. This involved an empirical examination of the assumptions underlying a self-ideal discrepancy questionnaire. The evidence is that the criticisms of Wylie (1975) and Burns (1979) are unjustified.

The purpose of the investigation reported here is to examine the consistency of SE across situations when subjects have actually just participated in those situations. This is known as 'active' role-playing. Various authors (eg. Ginsburg, 1979) have noted that subjects involved in such active role-playing exhibit much of the spontaneity of the 'real thing'.

Four situations are used in this study, all of which are 'analogues' of those verbally described to the sample of students who provided the free-descriptions from which the questionnaire was developed. None of the subjects in the present investigation were members of the previous sample. The four situations are: (a) having a chat with a friend; (b) getting acquainted with a member of the same sex; (c) writing a summary of a passage alone; (d) writing a summary of a passage in a group (a same-sex group, consisting of four members).

The study has two main aims. First of all, to extend the results of the previous investigation by testing two hypotheses:

(a) That the situations most similar in their content of SE would also be those most similar in their level of SE. The more detailed prediction is that the two social situations will be highly similar; that the situation of writing a summary in a group will be moderately
similar both to that of writing a summary alone and to the two social situations; and that the situation of writing a summary alone will be unrelated to the social situations.

and (b) The Duval — Wicklund hypothesis that those aspects of the self occupying people's awareness a great deal in a particular situation will be associated with greater self-ideal discrepancies than those aspects of the self which occupy people's awareness much less. This hypothesis is expected to be false. By definition, aspects of the self occupying the awareness of HSE's should, of course, not be associated with LSE.

The second main aim of the present study is to provide a form of construct validation of the questionnaire. This may be done by comparing the similarity in content of SE for each situation according to frequency analysis involved in its development, and according to the relevance-ratings received by each item in each situation from the subjects in the present study. The mean frequency (for the two judges) of the items in each situation (determined during the course of the development of the questionnaire) may be correlated with the mean relevance rating each item receives in each situation in the present study. If such correlations are high, evidence for the construct validity of the questionnaire will have been found.

In addition, the consistency of SE level per se is examined, as before, in two ways. Firstly, the rank-order variation (correlations) are computed between all situation-pairs. Secondly, a one-way repeated-measures ANOVA is performed to see if mean SE level varies across these four situations.
METHOD

Overview: Subjects consisted of the entire class of second-year psychology students. In the first session, friends and relative strangers were found for each student so that the two social situations could take place. In the second session, the students participated in each of the four situations, with the other students involved in each one carefully specified. Immediately after taking part in each situation, subjects completed the SE questionnaire. The class was divided into two groups, each of whom took part in the situations in a different order.

Subjects: The subjects are the entire class of second-year psychology students at Leicester University, of which there were 54. 41 of these are female and 13 are male. The experiment was conducted during a laboratory class. Because of the restrictions imposed (each subject had to take part in four specified situations), data from 35 subjects was obtained. 5 students did not attend the laboratory session, and 14 only took part in some of the situations (mainly because the 5 who did not attend prevented the others from taking part). A further 2 subjects did not complete all the questionnaires. Hence the results are based upon the remaining 33 (25 females and 8 males).

Design: There is one independent variable with four levels, namely the situations in which the subjects took part. Thus the design is a repeated measures one. The dependent variables are: (a) SE level: measured using the d score of Osgood et al (1957) on the self-ideal discrepancies and (b) SE content: measured by asking subjects to rate on a '0' to '6' scale how much they thought about the particular aspect of themselves signified by the item, in each situation (the relevance-ratings). On this scale, '0' is defined as 'I do not think about this aspect of myself
at all' and '6' as 'I think about this aspect of myself a great deal'.

Procedure:

(i) First session: This took place at the beginning of a lecture, several days before the first group's laboratory class was scheduled. The subjects were all given a questionnaire which attempted to find friends and relative strangers, for each subject. Strangers were found by providing each subject with a list of six randomly-drawn names of same-sex students who belonged to the same laboratory group as the subject, and asking the subjects to rate each name on a scale from zero (labelled 'I have never met this person') to ten (labelled 'I know this person extremely well'). This lists were drawn up so that:

(a) every same-sex name within each laboratory group had an equal probability of being chosen for any one subject (except for the subject's own, of course), and

(b) every name was used an equal number of times.

It was hoped that every subject would be relatively unacquainted with at least one person represented on the list. Students whose names were given a score of 3 or less on the 0-10 scale were judged to be strangers to the subject.

Friends were found by asking subjects the following question:

'Could you now name as many friends as you can, in this psychology class, who are the same sex as yourself? Please only name people you would rate 8, 9, or 10 on the scale (above). If you can't think of many, it does not matter - just name those people you know very well.'

Subjects were provided with room to list six names, and to add on others if they so wished. It was hoped that every subject would name at least one person in the same group in the laboratory class.

Naturally, the confidential nature of the questionnaire was stressed
to the subjects.

From these data it proved to be possible to assign subjects to partners they regarded as friends for the situation 'chatting to a friend', and to partners they considered as relative strangers for the situation 'getting acquainted'. The 'writing synopsis in group' situation was constructed so that subjects were not in the group with anyone named as a friend; so that all group members were of the same sex; and so that there were four members in each group.

(ii) Second session: The second session took place in the psychology laboratory. Once all the subjects were assembled, the questionnaire booklets were distributed (each subject received a personal one) and the following instructions were read out. These were also printed on the first page of the booklet.

This afternoon we are going to attempt a social psychology experiment. As you are aware, social behaviour depends upon many things, some very subtle, and therefore it is vitally important that you follow the instructions to the letter.

You are going to participate in four different situations, each one for 10 minutes. Some of the situations involve several people - the demonstrator will read out the names of the people who do a situation together. When everyone is settled down into the correct group (or pair) the demonstrator will tell you to go ahead with the situation. After 10 minutes you will be asked to stop - please stop what you are doing and fill in the relevant form for each situation as described below. During each situation please do not talk or signal to anyone except those involved in your particular group (if any). It is extremely important that you do not communicate with anyone outside your group.

When the 10 minutes is up for each situation, fill in the questionnaire as follows:
1) Rate the concept 'what I am really like'. That is, your opinion of yourself when you took part in the situation (which is described at the top of the page). For each phrase you will see there is a scale of numbers going from 0 to 6. Please put a circle round one of these numbers so that it indicates how well the phrase described you when you took part in the situation. A circle round the '0' means the phrase doesn't describe you at all, and a circle round the '6' means the phrase describes you extremely well. Circle the intermediate numbers as appropriate. When you are doing this, please compare yourself only to other members of this psychology class.

2) This time rate the concept 'what I would most like to be like'. Again, use the scale as before (a '0' means the phrase doesn't describe what you'd like to be like at all in the situation and a '6' means the phrase describes what you'd like to be like extremely well), but this time put two circles round the number (so that I can distinguish between the two ratings). Again, use the intermediate numbers as you need to. The number can be the same as before (in which case it will have three circles round it) or a different one. Again, when making these ratings, please compare yourself with other members of this psychology class.

3) Lastly, you will see on the questionnaire that there is a column headed 'How relevant this ability or aspect of myself is to this situation'. Put in a number from 0 (meaning the ability or aspect of yourself is not at all relevant) to 6 (meaning it is extremely relevant), for each phrase. Again, use the intermediate numbers as you need to (ie. as often as you like).

Please try to be as quick as you can when filling in the questionnaire - don't stop to ponder on each phrase; it is your immediate impressions and reactions to the situation that I would like you to indicate. Remember, when making the ratings, to compare yourself with other members of the psychology class. Thank-you.
Once these instructions had been read out, any questions subjects had were answered. Subjects then took part in the four situations in either the order 1, 3, 4, 2 or the order 3, 2, 4, 1. After participating in each situation for 10 minutes, the subjects filled in the SE questionnaire in accordance with the instructions.

At the end of the experiment, subjects were given a handout detailing the hypotheses behind the investigation, and a discussion on the nature of personality traits took place.
RESULTS

1. Consistency of SE level across situations

(a) Rank-order consistency

It was predicted that the two social situations ('getting acquainted' and 'having a chat with a friend') would be the most similar in SE level; that 'writing a synopsis in a group' would be moderately similar both to the two social situations and to the situation of 'writing a synopsis alone'; and that the latter would be quite different from the two social situations. Table 4.3 shows the predicted and obtained correlations between each pair of situations.

<table>
<thead>
<tr>
<th>Situation Pair</th>
<th>Predicted Correlation</th>
<th>Obtained Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>'Chat with friend' x 'getting acquainted'</td>
<td>high</td>
<td>.66**</td>
</tr>
<tr>
<td>'Synopsis in group' x 'chat with friend'</td>
<td>moderate</td>
<td>.59**</td>
</tr>
<tr>
<td>'Synopsis in group' x 'getting acquainted'</td>
<td>moderate</td>
<td>.76**</td>
</tr>
<tr>
<td>'Synopsis in group' x 'synopsis alone'</td>
<td>moderate</td>
<td>.31*</td>
</tr>
<tr>
<td>'Synopsis alone' x 'chat with friend'</td>
<td>n.s.</td>
<td>.01</td>
</tr>
<tr>
<td>'Synopsis alone' x 'getting acquainted'</td>
<td>n.s.</td>
<td>.27</td>
</tr>
</tbody>
</table>

NB: * denotes the correlation is significant at $p < 0.05$.

** denotes the correlation is significant at $p < 0.01$.

As can be seen, there is an excellent fit between the predicted
and obtained correlations, with one exception: the correlation between 'getting acquainted' and 'synopsis in group' is higher than was predicted. This suggests that subjects viewed the situation of writing a synopsis in a group more as a social situation involving abilities and skills similar to those involved in getting to know someone, rather than as an intellectual situation involving abilities and skills similar to those involved in writing a synopsis alone.

(b) Variability in mean level of SE

A one-way repeated-measures ANOVA was performed on the data to see if mean SE level was lower in some situations than in others. The ANOVA results are shown in Table 4.4(a) and the means in Table 4.4(b). Multiple

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Situations</td>
<td>88.36</td>
<td>3</td>
<td>29.45</td>
<td>5.85</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Error (sub. x sits.)</td>
<td>482.46</td>
<td>96</td>
<td>5.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1112.33</td>
<td>131</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Situation</th>
<th>'Synopsis alone'</th>
<th>'Synopsis in group'</th>
<th>'Getting acquainted'</th>
<th>'Chat with friend'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean SE score</td>
<td>8.94</td>
<td>7.54</td>
<td>6.68</td>
<td>7.40</td>
</tr>
</tbody>
</table>
comparisons (Tukey's 'honestly-significant-difference' test) revealed that the only significant differences in mean SE level are between the situation of writing a synopsis alone and the rest ($q = 3.59, p < 0.05$ for the comparison with 'synopsis in group'; $q = 5.79, p < 0.01$ for the comparison with 'getting acquainted'; and $q = 3.95, p < 0.05$ for the comparison with 'chat with friend').

It therefore appears that the situation of writing a synopsis alone is different from the other situations in both level of SE for the group as a whole, and in the rank-order of subjects with respect to SE. This latter result is particularly important because it demonstrates yet again that measuring SE without specifying the situation may well reduce the validity of research designs where SE is used as an independent variable.

2. Similarity of the situations in content and level of SE

The main hypothesis to be tested by the present study is that situations which are similar in their content of SE (ie. in the aspects of the self focussed upon within them) will also be similar in their level of SE. Of course, the converse is also predicted: the more dissimilar two situations are in their content of SE, the more dissimilar they should be in their level of SE. This hypothesis is tested in the same way here as it was in Chapter 3. This involves computing the correlation between the correlations of each situation-pair in its level of SE and the correlations of each situation-pair in its content -ratings. Table 4.5 shows the two sets of correlations. It is immediately obvious that the content -rating correlations are lower, in general, than are those for the SE level. However, there is quite a good fit between the two sets of correlations ($r = .70, p < 0.07$). The correlation just misses being significant at the conventional $p < 0.05$ level due to the small degrees of freedom. But it nevertheless provides some support for the hypothesis.
TABLE 4.5
Correlations between situation-pairs in their level of SE and in their relevance-ratings

<table>
<thead>
<tr>
<th>Situation Pair</th>
<th>Correlation in SE level</th>
<th>Median 'content' correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>'Synopsis alone' x 'synopsis in group'</td>
<td>.31</td>
<td>.34*</td>
</tr>
<tr>
<td>'Synopsis alone' x 'getting acquainted'</td>
<td>.27</td>
<td>.21</td>
</tr>
<tr>
<td>'Synopsis alone' x 'chat with friend'</td>
<td>.01</td>
<td>.19</td>
</tr>
<tr>
<td>'Synopsis in group' x 'getting acquainted'</td>
<td>.76**</td>
<td>.32*</td>
</tr>
<tr>
<td>'Synopsis in group' x 'chat with friend'</td>
<td>.59**</td>
<td>.26</td>
</tr>
<tr>
<td>'Getting acquainted' x 'chat with friend'</td>
<td>.66**</td>
<td>.38*</td>
</tr>
</tbody>
</table>

NB: * shows the correlation is significant at the p < 0.05 level.
** shows the correlation is significant at the p < 0.01 level.

The more similar two situations are in their content of SE, the more similar they tend to be in their SE level.

3. The Duval — Wicklund hypothesis

This hypothesis is exactly the same as that tested in Chapter 3. It is that when people think about themselves they will experience dissatisfaction because, in general, people do not live up to their goals. This hypothesis is rejected here. According to Duval and Wicklund, therefore, items given high content —ratings should 'generate' larger self-ideal discrepancies than should items given low content —ratings.

A strong test of this hypothesis may be applied by comparing the
mean self-ideal discrepancy associated with content -ratings of 0 with
the mean self-ideal discrepancy associated with content -ratings of 6,
for each subject. If the mean self-ideal discrepancies associated with
items given a content -rating of '6' are larger than those given a
content -rating of '0', support for the 'Duval and Wicklund hypothesis
would be found. In order to include as many subjects as possible, the
categories '5 and 6' were combined, and compared with the '0' category.

The mean self-ideal discrepancy associated with content -ratings
of '5' and '6' is 1.29, and for those associated with content -ratings
of '0' is 1.25. This difference is not significant (t = 0.16, n.s.).
Thus, the 'Duval and Wicklund' hypothesis is once again given no support.

4. Validity of the Questionnaire

The content of SE in each situation in the present study may be
compared to the content of SE found in the development of the question-
naire, thus providing information relevant to the validity of the question-
naire. The two values compared are: (a) the frequency of usage of each
item as denoted by the two judges involved in the development of the
questionnaire and (b) the mean content -rating of each item obtained
from 33 students who actually participated in the four situations.

The frequencies and the mean ratings are correlated for each item
on the questionnaire for each situation. The four correlations are
shown in Table 4.6. The mean correlation is 0.63 (p < 0.01), which is
moderately high. This suggests that the questionnaire has some validity.
The lowest correlation, as can be seen from Table 4.6, is for the situation
'writing a synopsis in a group' (r = 0.46, p < 0.05). The two highest
correlations are for 'writing a synopsis alone' (r = 0.71, p < 0.01)
and for 'getting acquainted' (r = 0.75, p < 0.01). Thus it seems that
these two situations, at least, may be simulated fairly well in the
laboratory setting.

TABLE 4.6
Correlations between content of SE according to frequency of usage, and according to mean content-ratings

<table>
<thead>
<tr>
<th>Situation</th>
<th>Synopsis Alone</th>
<th>Synopsis in Group</th>
<th>Chat with Friend</th>
<th>Getting Acquainted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>.71**</td>
<td>.46*</td>
<td>.59**</td>
<td>.75**</td>
</tr>
</tbody>
</table>

NB: * shows the correlation is significant at the p < 0.05 level.
** shows the correlation is significant at the p < 0.01 level.
DISCUSSION:

Active Role-Playing Investigation

In this investigation an attempt is made to examine the consistency of SE across four carefully chosen situations: chosen so that they could not be described as extremely anxiety-provoking and so that they were reasonable analogues of situations commonly encountered in the everyday lives of typical university students. A questionnaire was designed so that its items would be both relevant to the situations employed and would contain phrases familiar to most of the subjects.

It was found that the intellectual situation of writing a summary is different from the others in three ways. First of all, its content of SE is different. That is, those aspects of the self focussed upon when writing a summary are different from those focussed upon when engaged in the other situations used here (such as having a chat with a friend). Secondly, the rank-order of a subject's SE score when in the situation of writing a synopsis bears little relationship to its rank-order in the two social situations and only a small relationship to its rank-order in the situation of writing a synopsis in a group. Thirdly, the mean SE level of the group as a whole is lower in this situation than in the others.

These results have important implications for both theory and research in this area, as has been pointed out several times already in this thesis. The fact that there is little rank-order consistency in SE level in situations which differ markedly in SE content demonstrates that researchers need to be extremely careful in two respects: firstly, when generalising their results; and, secondly, when assigning subjects to SE groups for the purpose of examining differences between such groups on dependent variables. Indeed, these findings raise the whole question of whether there is any validity in the measurement of SE without the
specification of situations to which subjects should address themselves when making their ratings. On the basis of the results of the present investigation, it may be strongly recommended that investigators do specify such situations.

There is also support in this investigation for the notion that situations engaging similar aspects of the self will have similar SE levels. The correlation between situational similarity in SE content and SE level is high ($r = 0.70$) but not quite significant ($p < 0.07$). Taken together with the results of the previous study, however, the hypothesis does seem to be supported.

Once again, the Duval — Wicklund hypothesis is not supported by the results. Aspects of the self which subjects think about a great deal are not associated with significantly larger self-ideal discrepancies than are those aspects of the self which subjects do not think about at all. This result does not surprise the present author for, by definition, SE is a continuous variable which ranges from low to high. It does not make sense to talk of a person who has HSE and, at the same time, has large self-ideal discrepancies on those aspects of himself which engage his attention.

Finally, this study demonstrates that the SE questionnaire has good validity. The content-ratings of the items correlate significantly with the frequency of usage of the items in each situation.
SUMMARY AND CONCLUSIONS:

Chapters 2–4

The first half of this thesis has critically examined several assumptions made by researchers in the area of SE and the self-concept. The first experiment examined how the content of SE varied for schoolchildren when they were at home compared to when they were at school. More than half of the frequently-used constructs were used more often either at home or at school. Thus the assumption that SE does not vary in content is shown to be false. In addition, this experiment also examined the comparison-others used by schoolchildren when they were assessing themselves. The comparison-others were different at home and at school, and were different for males as opposed to females. Therefore the assumption that comparison-others do not vary is also shown to be false, though more work is needed on this issue.

The second study demonstrates further that level of SE varies across situations in two ways: in the rank-order of SE scores, and in the mean level. Moreover, support is found in this study for a major hypothesis put forward in Chapter 1. This is that situations which engage similar aspects of the self will also be similar in the rank-order of SE levels associated with them, provided that the comparison-others are held constant. Thus SE level does not vary in a haphazard way across situations. In addition, this experiment found no support for the hypothesis derived from Duval and Wicklund (1972, 1975). This hypothesis states that people should have LSE on those aspects of themselves which engage their attention in particular situations.

The third study also examined SE level and content across situations. The results show, once again, that the assumption of lack of variability in SE level across situations is an unjustified one. The results provide more (but weak) support for the notion that situations engaging similar
aspects of the self will be similar in the rank-order of SE levels associated with them, and once again provide no support for the Duval—Wicklund hypothesis.

It may be concluded from these results that the assumption of lack of variability in SE level across situations is false. Such a conclusion has important implications for both theorising about the self, and for empirical research in the area. In theorising about the self, more attention should be paid to the situational and personal factors which influence SE level within a situation. One situational factor proposed here is the activities performed in the situation. A person's SE level in two situations involving different activities may or may not be the same, whereas a person's SE level in two situations involving similar activities is much more likely to be similar. Nevertheless, such an a-priori classification of situations by the experimenter has its limitations. Subjects just may not perceive the situations in the same way (e.g. Forgas, 1979). Therefore a useful extension to the present research would be the study of individual differences in the perception of situations. This next step was carefully considered by the present author. However, it was thought to be of more importance to understand how SE level affects behaviour within a situation. The studies reported in the remaining chapters of this thesis are a beginning in this direction.

As well as implications for theory, the results of the studies reported so far also have implications for measurement. In particular, it is strongly recommended that further researchers in this area specify two things when measuring SE: (a) the situation(s) in which subjects should imagine themselves when making their self-ratings; and (b) the people with whom subjects should compare themselves when making their self-ratings.

If these are not specified, the consequences for measurement may be
disastrous. Subjects may be imagining different situations and be comparing themselves with comparison-others of widely varying characteristics, when making their ratings. If this occurs, then the relationship between SE and the dependent variables may be grossly distorted (for example, reduced to non-significance).

In this chapter, considerable effort was devoted to the development of a theoretically satisfying and empirically justified questionnaire to measure SE. To be theoretically satisfying, the questionnaire must assess both 'self-assessments' and 'ideals'. A modified semantic-differential format was chosen for a number of reasons such as its ease of administration to a group, and its ability to enable both self-assessments and ideals to be measured at once. The criticisms of Wylie (1974) and Burns (1979), against discrepancy measures of SE, were put to empirical test on two samples of students totalling 137 altogether. All of the criticisms are shown to be unjustified. In particular, the assumption, made by discrepancy measures, that the larger the discrepancy the greater the distress with which it is appreciated, is found to be a reasonable assumption. In addition, it was found that 'self' scores are greater than 'ideal' scores only extremely rarely. This rarity implies that the SE score is unlikely to be significantly affected by their occurrence.

In addition, the questionnaire has good split-half and test-retest reliability (certainly adequate for basic research purposes). Finally, an assessment of its validity was performed, and the questionnaire was found to have adequate validity.

In sum, Chapters 2, 3 and 4 of this thesis have shown:

(a) that two important assumptions generally made by researchers in the area of SE and the self-concept are false. The implications of this for theory and research have been discussed, and

(b) that the assumptions underlying the use of semantic-differential
type discrepancy questionnaires when used as a measure of SE are empirically justified.

The subsequent chapters of this thesis present experiments devoted to an examination of the way in which feedback information is processed by people with LSE and HSE within particular situations. In accordance with the recommendations to investigators made as a result of the studies so far, in later experiments of this thesis where SE is measured the comparison-others and situations are always specified.