CHAPTER 2

REVIEW OF RELATED LITERATURE

Concept of empathy

Empathy has been defined differently by different investigators. Flevell et al. (1968) define it as the ability to infer accurately the feelings of other people. Thus, empathy becomes primarily a cognitive skill. For others, it is also the tendency to experience vicariously the emotions others are feeling (Hoffman, 1975; Feshbach, 1978). This latter definition accords a more important role to empathy in prompting altruism since it implies that an empathic person would be motivated to help another person in need. Traditionally, when the question of empathy has been raised, it has been raised in the context of counselling and psychotherapy. (Rogers, 1957; Rogers and Traux, 1967).

One of the earlier references to the concept of empathy appeared in the writings of the German psychologist, Theodore Lipps (1907) who labeled the phenomenon "Einflulbing", the literal translation of which is "feeling into another" (Katz, 1963). The affective dimension of empathy is evident in the writings and conceptions of McDougall (1908), Freud (1950), and
Sullivan (1953). Empathy is more demanding and difficult requiring the person to empathize with people and groups who are basically different, whose commonalities and similarities are minimal. An early conception of empathy, in predominantly cognitive terms, appeared in the writings of George Mead (1934). A critical theorizing element in his theory was that empathy was acquired or learned through role taking and initiation, a proposition that is represented in contemporary cognitive approaches.

During more recent years, a renewed interest in the term has developed, particularly in developmental and clinical psychology. There is evidence (Hoffman, 1970) that the development of empathy, as a part of the process of socialization, is an important prerequisite to the development of morals. The role of empathy in counselling and therapy is important, because it allows the therapist to experience the client's anger, fear or confession as if it were the therapist's own. The therapist can thereby communicate a "clear understanding of these conditions to the client (Rogers, 1961)".

Recently, the importance of empathy has been studied as an element in such diverse situations as the mother-child relationships (Schachtal, 1959), teacher-student relationships (Gage, 1968), accurate
perception of others (Steiner, 1955) and employee-employer relationships (Haire, 1955).

The importance of the empathic process in reducing inter-group tensions (Cottrell and Dymond, 1949), social prejudice (Tock et al., 1968) as well as in cross-cultural perception (Bagby, 1968) has also been exposed. In a number of investigations, differences in empathy have been explored. Maccoby and Jacklin (1974), Block (1976), Hoffman (1977), Davis (1983), and Eisenberg et al. (1983) reviewed that there are individual differences in having empathetic tendency. Their investigation also explored that females are more empathic than the males.

Measures of emotional empathic response have typically relied on self-report and physiological indicators. Stotland. (1969) reported a lack of consistency among various self-report measures of empathy, palmar sweating and vasoconstriction. The low validity of the self-report is due to their confounding with response style (Jackson and Messick, 1958) and the inadequacy of the physiological measures is in part due to their failure to differentiate various facts of emotional experiences. Mehrabian and Epstein (1972) have developed a generalized measure of empathy. In two experiments, scores on this
scale correlated with independent measures of aggression and altruism. In the first study, highly empathic subjects aggressed less towards another person than did the less empathic subjects. The situation employed was the teacher-learner shock paradigm. In a second study, empathic subjects were more likely to help another subject in the experiment than were those low in empathy.

**MEHRABIAN AND EPSTEIN'S (1972) MEASURE OF EMOTIONAL EMPATHY (1972).**

The object of the Mehrabian and Epstein's (1972) study was to develop an adequate measure of emotional empathy and to test its validity in two settings where such a personality attribute is relevant. In the first section, experimenters investigated the correlates of empathy in aggressive situations; in the second section, an experiment investigated its relation to helping. Based on those two experiments, Mehrabian and Epstein developed a questionnaire measure of emotional empathy which has already been discussed in Chapter 1.

The response to each item of the measure of empathic tendency is on a +4 (very strong agreement) to -4 (very strong disagreement) scale and the (+) and (-) signs preceding each item indicate the direction of scoring. To compute a total empathy score, the signs of a subject's
responses on the negative (-) items are changed and then an algebraic sum of all 33 responses to the scale is obtained. For the sample which includes an equal number of males and females, mean = 33 and standard deviation = 24. However, males and females differ significantly in their empathic tendency. For males, M = 23, SD = 22; for females, M = 44 and SD = 21.

The relation of emotional empathy to aggressive behaviour

The two experiments in this section were specifically designed to test the empathy-nonaggression and immediacy-non-aggression hypotheses. The inclusion of both factors within the same experiment was motivated by the following hypothesis: the decrement in aggression in a more immediate feedback condition is especially greater for more empathic persons. The latter hypothesis was suggested in part by Buss's (1961) findings which showed differential responsiveness of subjects to painful emotional feedback from victims.

Ninety-one undergraduates of both sexes were the subjects. The independent variable included in the experiment was empathic tendency, two levels of immediacy of the victim (immediate versus removed) and sex of the subject. The experiment consisted of two ostensibly
distinct parts, separated by an interval of one week. In the first session, the empathy questionnaire was administered to groups of ten subjects.

The second study was described as an experiment concerned with "personality and learning", and it was of only twenty minutes duration. Subjects were each paired with an experimental confederate of the same sex. The experimenter "randomly" selected the subject to act as a teacher and the confederate to act as a pupil. Having read a character sketch of a third person, the pupil confederate was to try to make predictions about his personality. Pupil, in reality, answered according to a predetermined schedule, which corresponded to the teacher's answer key on 15 of 30 trials (i.e. he made 15 errors). The teacher subject was instructed to use the controls on half of the device to indicate to the pupil when he had answered correctly and to punish him each time he made an error. To punish the pupil for each error, the teacher had a choice of administering any one of the seven intensities of electric shock. The pupil had an electrode attached to one hand, which supposedly delivered the electric shocks. In reality, he never received any shock. The differential reaction of the confederate as a
function of shock intensity was designed to approximate real life situations and consisted of increasing degrees of facial grimaces, gasping and jerking of the arm to which the electrode was attached.

Prior to the teaching session, to convince the teacher that he would be shocking the pupil, the electrode was placed on the teacher's hand, and real shocks were delivered. The teacher received the lowest and middle levels of the seven intensities of shock while being shown the corresponding scale values. This procedure served to inform him of the level of pain associated with the shock scale. The lowest level was barely supraliminal.

In the immediate condition, the pupil was seated 8 feet from the teacher and was fully visible. In the nonimmediate condition, the pupil was seated in another room, out of the view of the teacher, but could be heard. This teacher-pupil shock paradigm was adopted from Milgram's (1965) experiment.

An overall aggression score was computed for each subject. A stepwise multiple regression method described by Cohen (1968) was used to explore the significant main and interactive effects of the following factors on the degree of aggression: sex, empathic
tendency, and immediacy of the victim. Mean aggression of males was 2.55, and 2.05 for females. Retesting was done on 104 subjects from the same population. The results for average aggression obtained from this replication showed only the single significant .05 level effect of immediacy x empathic tendency ($B = -0.023$).

The results from both experiments showed that although empathy itself was not a sufficient condition for inhibiting aggression, differences in empathy did make a difference when combined with differences in immediacy of the victim. Whereas low empathy subjects aggressed with equal intensity against an immediate Vs a nonimmediate victim, highly empathic person aggressed less when the victim was more immediate.

In the first experiment, males aggressed significantly more than females, but in the second one, the effect was similar, but nonsignificant. Past experimental data and current observations also indicate that the suffering of another person may evoke a strong aversive reaction in the onlooker. In the research work on obedience carried out by Milgram (1963), 35% of the Ss refused to apply strong shocks to their partners, reacting with violent emotions to their (pretended) sufferings.
Those who submitted to orders and applied maximum shocks were, on the whole, greatly perturbed.

Those examined in an analogous experimental situation by Tilker (1970) reacted with greater distress, the clearer the signals of pain from their partners (personally unknown to them). In this experiment, the reactions of persons who witnessed other people's pain at times assumed extreme intensity and were manifested in such acts as snapping wires off the apparatus and physically menacing the experimenter. Benger (1962) and Lazetta (1951) reported that when confronted with signs of emotionality on the part of another person, one may experience either empathic or counterempathic reactions (envy, sadism) oneself. Evidence for the occurrence of such arousal in response to signs of pain on the part of victim has been obtained in several experiments (Baron, 1971; Rule and Leger, 1976).

The findings of such studies indicate that if an aggressor has not been strongly angered by his or her victim and is similar to this person in several ways (age, race and background), signs of pain on the part of the victim can inhibit further aggression (Buss, 1966, Baron, 1974). Baron (1971) found that the stronger the pain of the 'victim', the weaker the aggression, regardless of
whether the subjects were previously attacked by their 'victims'. Feshbach (1979) reported that the signs of discomfort on the part of the victim may fail to induce feelings of empathy. Eliaasz (1975) indicated that the manifestations of suffering may block an aggressive action, but do not invariably do so.

In a number of investigations with children and adults, a consistent inverse relationship between empathy and aggression has been obtained (Feshbach and Feshbach, 1969; Huckabay, 1972; Feshbach, 1976). In addition, there is some preliminary evidence that training in empathy or empathy-related skills such as role playing behaviour and perspective taking, holds considerable promise for reducing aggressive or antisocial behaviours and promoting social adjustment and social perspective skills (Staub, 1971: Chandler, Greenspan and Barenbein, 1974; Spivak and Shure, 1974; Pitkanen, 1974; Lieshout et al. 1978). Feshbach (1978), viewed that the affective component of empathy had a very special relationship to the regulation of aggression. Aggressive behaviour is a social response that has the defining characteristic of inflicting injury upon persons or objects, causing pain and distress. The observation of these aversive behaviours should elicit
distress responses in an empathic observer even if the observer is the instigator of the aggressive act. The painful consequences of an aggressive act, through the vicarious affective response of empathy, may be expected to function as inhibitors of the instigator's aggressive tendencies, thus contributing to an inverse relationship between empathy and aggression.

Ohbuchi and Ken-ichi (1988) reviewed a study on the effects of situational and personality factors of empathy on aggressive behaviour. They viewed that the emotional reactions are determined by the motives involved in aggressive behaviour. Empathy as a personality trait is characterized as being differentially responsive to empathy-arousing stimuli, with an emphasis on the interaction effects on aggression of personality and situational factors in the arousal of empathy.

The relation of emotional empathy to helping behaviour

The second study was designed to explore the combined contribution of situational and personality variables to helping behaviour. In order to introduce variations in liking, each subject was paired with an experimental confederate who was previously presented as
having either similar or dissimilar attitude (Byrne, 1969). It was also recognized that the actual face to face interchanges between subjects and confederates could influence liking.

The personality attitudes measured were: emotional empathic tendency, succorance (Jackson, 1967), affiliative tendency and sensitivity to rejection (Mehrabian, 1970), and approval seeking tendency (Crowne and Marlowe, 1960). In the first section, it has been shown that the measures of emotional empathic tendency are related to subjects' differential aggressiveness toward a victim when they received pain feedback from him. Thus, it was assumed that this personality measure was the most relevant for exploring an observer's responsiveness to someone in need of help. It was hypothesized that altruistic behaviour is a correlate of empathic tendency and of similarity with the person needing help.

Eighty-one female undergraduates at the University of California were selected for this study. Upon arriving at the experiment, each subject was placed in an individual room and asked to fill out an attitude questionnaire similar to the one employed by Byrne (1969). After the collection of this questionnaire the subjects
responded to the measures of emotional empathic tendency, affiliative tendency and sensitivity to rejection (Mehrabian, 1970), succorance (Jackson, 1967), and approval-seeking tendency (Crowne and Marlowe, 1960). A semantic differential questionnaire which measured characteristic emotions in terms of the three feeling factors (pleasure, arousal, and dominance) proposed by Mehrabian and Russell (1972) was also included in this set. The subjects were paired with confederates who were portrayed as holding either similar or dissimilar attitudes. According to a standardized script, the confederates acted emotionally upset about a personal problem and asked the subjects to volunteer time for a class assignment. A postinteraction questionnaire assessed subjects' perceived similarity to, and liking of, the confederates.

Results showed that in a regression analysis, helping behaviour was a significant correlate of empathic tendency ($B = 0.31$), with both variables normalized. The attitude similarity manipulation was found to be highly correlated with perceived similarity ($r = 0.79$, $p < .01$), but it was unrelated to liking or helping ($r = .18$, $p<.10$). A separate regression analysis (Mehrabian and Russell, 1972)
showed that the primary emotional component of empathic tendency was heightened arousal ($B = 0.33$).

Thus, persons who are characterized as possessing higher empathic tendency tend to be more aroused by others' emotional experiences of both positive and negative quality. The measure of emotional empathy (Mehrabian and Russell, 1970) showed validity in quite distinct settings and therefore has potential applications in the corresponding situations.

Several more studies have been conducted to study the relationship between empathic tendency and helping behaviour. In a number of investigations of prosocial behaviour as altruism, sharing and generosity, empathy is postulated to be a key mediating mechanism (Staub, 1971; Hoffaman, 1975). Krebs (1975) reported that the people who are characteristically high in empathy are more apt to provide help to someone experiencing distress than those low in empathy. The phenomenon goes even further; even when there is no real distress or strong emotion involved, highly empathic people are more likely to offer help. Eisenberg-Berg and Mussen (1978) found that male subjects who helped an experimenter by acting as subjects in a long tedious experiment, had significantly
higher empathy scores, as measured by a written personality assessment, than non-volunteers did.

Some theories of empathy suggest that helping behaviour may be caused by the anticipation of the victims' well-being and happiness following the termination of distress. The feeling of empathy for a victim will enhance the probability of helping, but the exact mechanism by which empathy causes helping is not entirely understood (Rushton, 1980). In a separate investigation, Batson et al. (1981) reinforced the idea that empathy leads to altruistic rather than to egoistic motivation to help. Toi and Batson (1982) reported that helping behaviour can be motivated by people's observation of distress of a victim, because observers begin to put themselves in the place of a victim, feeling as if it were they who were suffering. The results of their study indicated that subjects in the low empathy condition helped less when avoidance was easy than when it was difficult. But subjects in the high empathy condition displayed a high rate of helping even when avoidance was easy. Thus altruism, fed by empathy, triumphed over the self-oriented goal of avoiding an unpleasant situation. The relationship between altruism and empathy was also
found by different investigators such as Kalliopuska (1983), Morgan and Sharon (1984), Barnett (1985), Eisenberg et al (1987), Jackson (1987), Stiff et al. (1988), Schaller et al. (1988), Schroeder et al. (1988) and Eisenberg et al. (1989).

Mehrabian and Epstein's (1972) scale has been extensively used in a number of research settings. Davis (1983) has described the development of an individual difference measure of empathy in adults by using Mehrabian and Epstein's (1972) empathy scale and Hogan's (1969) empathy scale. The correlation between empathic concern for these two scales and Interpersonal Reactivity Index Subscale was computed separately for males and females as shown below:

<table>
<thead>
<tr>
<th></th>
<th>Hogan Empathy Scale</th>
<th>Mehrabian and Epstein Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>Empathic concern</td>
<td>.11</td>
<td>.25</td>
</tr>
</tbody>
</table>

Kalliopuska (1983) assessed verbal components of emotional empathy, perceptual and motor skills using the 33 item empathy scale (Mehrabian and Epstein, 1972). He also used the same scale for measuring empathy in children and social class, between empathy and birth order, and between empathy in adults and social class. Barnett (1983)
used the said scale in studying the effects of anticipated victim, responsiveness and empathy upon volunteering. Burleson et al. (1983) examined the collective contributions of a social cognitive variable (interpersonal construct abstractness) and a motivational variable (emotional empathy) to the production of sensitive comforting strategies among the subjects. They also preferred Mehrabian and Epstein (1972) empathic scale. The aforesaid scale was also preferred by Morgan et al. (1984) and Esienbert et al. (1987).

The forgoing review makes it clear that Mehrabian and Epstein's scale appears useful in studying the central component of empathy. The review also indicates that emotional empathy has a close relationship to aggression and helping behaviour and also there are individual differences in the empathic level of each person.

**Adaptational procedure - reliability and validity**

Construction and adaptation of a test entails the determination of reliability and validity. The different kinds of reliability and validity are reviewed below, these being relevant to the present study.
Reliability

Reliability refers to the consistency of scores obtained by the same persons when reexamined with the same test on different occasions, or with different sets of equivalent items, or under other variable examining conditions. It indicates the extent to which individual differences in test scores are attributed to "true differences" in the characteristics under consideration as the extent to which they are attributable to chance error.

The reliability of the test with itself is called the "reliability coefficient" of the test. There are many procedures of estimating reliability. Guilford and Fruchter (1978) discuss three general categories of procedures:
1) Internal consistency reliability.
2) Alternate or Parallel-form reliability.
3) Test - retest reliability.

Garrett (1979) describes four procedures in common use for computing the reliability coefficient of a test:
1) Test - retest reliability.
2) Alternate or parallel - form reliability.
3) Spilt - half reliability.

4) Rational equivalence reliability.

(1) Test - retest Reliability:

The most obvious method for finding the reliability of test scores is by repeating the identical test on a second occasion. The correlation is computed between the first and second set of scores. Although the test-retest method is simple and straightforward, it has some serious drawbacks. If the test is repeated immediately, the recall of the answer may tend to increase the scores. Besides, this practice and confidence induced by familiarity with the material would certainly affect scores at the time of retesting. Moreover, the nature of the test itself may change with repetition.

(2) Alternate or Parallel - Form reliability:

One way of avoiding the difficulties encountered in test-retest reliability is through the use of alternate forms of the test. The same person can thus be tested with one form on the first occasion and with another, equivalent form of the second. The correlation between scores obtained on the two forms represents the reliability coefficient of the test. Such reliability
coefficient is a measure of temporal stability and consistency of response.

(3) Split-half reliability:

In the split-half method, the test is first divided into two equivalent 'halves' and the correlation found for these half tests. From the reliability of the half test, self-correlation of the whole test is then estimated by the Spearman-Brown Prophecy formula

\[ r_{II} = \frac{2 \cdot r_{\frac{1}{2} II}}{1 + r_{\frac{1}{2} II}} \]

where \( r_{II} \) = reliability coefficient of the whole test.

\[ r_{\frac{1}{2} II} = \text{reliability coefficient of the half test, found experimentally.} \]

The procedure is to make up two sets of scores by combining alternate items in the test. The first set of scores represents performance on the odd-numbered items and the second set of scores, performance on the even-numbered items. This is one way of making up two half tests. Another way will be comparable in content, difficulty and susceptibility to practice.
Split-half method's main advantage is that all the data for computing reliability are obtained upon one occasion, so that variations brought about by differences between the two testing situations are eliminated.

(4) **Rational Equivalence Reliability:**

The method of rational equivalence represents an attempt to get an estimate of reliability of a test, free from objections raised against the methods outlined above. Two forms of a test are defined as "equivalent" when corresponding items, a, A; b, B; etc. are interchangeable and when the inter-item correlations are the same for both forms. This method stresses the intercorrelations of items in the test and the correlations of items with the test as a whole. In fact, rational equivalence reliability is akin to Guilford's internal consistency reliability and provides an estimate of the dependability of test scores. Here, mention may be made of two formulas of rational equivalence reliability: Kuder Richardson reliability and coefficient alpha. Rather than requiring two half scores, the first technique is based upon an examination of performance on each item. The most widely applicable Kuder-Richardson formula is the following:

\[ r_{tt} = \frac{\left( \frac{n}{(n-1)} \right)}{\frac{SD_t^2}{SD_e^2} - pq} \]
where $SD_t^2$ = variance of the total scores on the test

$pq = \text{proportion of persons who pass (p) and the proportion who do not pass (q) each time.}$

This formula is applicable to tests whose items are scored as right and wrong, or according to some other all or none system. On a personality inventory, the respondent may receive a different numerical score on items, dependent on whether "usually", "sometimes", or "rarely" is checked. For such tests, a generalized formula has been derived, i.e. coefficient alpha. In this formula, pq is replaced by $(SD^2)$, the sum of variance of item scores. The procedure is to find the variance of all individuals' scores for each item and then to add these variances across all items. The complete formula for coefficient alpha is

$$r_{tt} = \frac{(n)}{(n-1)} \times \frac{SD_t^2 - (SD_i^2)}{SD_t^2}$$

There is no one best way of estimating reliability. The method employed will depend upon one's purposes and the meaning one wishes to attach to reliability. A secondary consideration is availability of data in the proper form. Other considerations are testing conditions and the kind of test or other measure.
The correlation coefficient may be computed in various ways, depending upon the nature of data. The most common is the Pearson Product Moment Correlation Coefficient. Such a correlation coefficient takes into account not only the individual's position in the group, but also the amount of his deviation above or below the group mean. If the sum of the cross products is negative, the correlation will be negative. When some products are positive and some negative, the correlation will be close to zero.

Internal Consistency:

The essential characteristic of this method (discussed by Guilford and Fruchter (1978) is that the criterion is none other than the total score on the test itself. Sometimes an adaptation of the contrast group method is used, extreme groups being selected on the basis of total test scores. The performance of the upper criterion group on each test item is then compared with that of the lower criterion group. The items that fail to show a significantly greater proportion of keyed responses in the upper than in the lower criterion group are considered invalid, and are either eliminated or revised. The internal consistency correlations, whether based on items or sub-tests, are essentially measures of homogenity.
Validity

According to Anastasi (1988), the validity of a test concerns what the test measures and how well it does so. It tells us what can be inferred from test scores. A test is valid for a particular purpose, or in a particular situation it is not generally valid. The validity is of three types:

1) Content validity
2) Criterion-related validity
3) Construct validity

(1) **Content validity:**

It involves essentially the systematic examination of the test content to be measured. The content must be broadly defined to include major objectives, such as the application of principles and the interpretation of data, as well as factual knowledge. Moreover, content validity depends on the relevance of the individual's responses to the behavioural area under consideration, rather than on the apparent relevance of item content.

(2) **Criterion-related validity:**

It indicates the effectiveness of a test in predicting an individual's performance in specified activities. For this purpose, performance on the test is checked against a criterion, that is, a direct and
independent measure of that which the test is designed to predict.

(3) **Construct validity:**

The construct validity of a test is the extent to which the test may be said to measure a theoretical construct or trait. Each construct is developed to explain and organize observed response consistencies. It derives from established inter-relationships among behavioural measures. It requires the gradual accumulation of information from a variety of sources. Any data throwing light on the nature of the trait under consideration and the considerations affecting its development and manifestations represent appropriate evidence for this validation.

In order to demonstrate construct validity, we must show not only that a test correlates highly with other variables with which it should theoretically correlate, but also that it does not correlate significantly with variables from which it should differ. Campbell and Fiske (1959) described the former process as convergent validation and the latter as discriminant.

Construct validity, if loosely applied, may open the way for subjective unverified assertions about test
validity. Because construct validity is a broad and complex concept, it has not always been clearly understood by those who employed the term.

However, construct validity is a comprehensive concept, which includes the other types. All the specific techniques for establishing content and criterion-related validity discussed earlier could have been listed again under construct validity.