SCALING THE PERCEPTION OF CRIME SEVERITY USING THURSTONE'S METHOD OF PAIRED COMPARISONS

A thesis presented in partial fulfilment

of the requirements for the degree

of Master of Arts in Psychology

at Massey University

Lucielle Trynes

1993

ACKNOWLEDGMENTS

I would first like to thank my thesis supervisor, John Podd, for his patience, guidance and enthusiasm over my thesis topic. But I would mostly like to thank him for his time that he always gave freely and without hesitation. It is not easy being a lecturer, administrator, supervisor, husband and father, and still find time for one more problem to solve or one more thesis to proof-read.

I would also like to extend my gratitude to the New Zealand Army (Linton Military Camp). Thank you to all those who took part or co-ordinated meetings, places and times. Special thanks to Col. Ottaway for his permission to carry out the study at Linton Military Camp and to recruit army personnel as subjects. Thanks also to Capt. Ian Fenn for liaising between myself and the Army.

Next I would like to offer my appreciation to the psychology office staff at Massey University for always assisting me in in times of trouble and going beyond and above their call of duty. I would also like to thank Harvey Jones for creating the computer programmes that were required in the collection and analysis of data, as well as patiently teaching me the wonders of spread-sheets.

It is a wonder that I still have family and friends left after thesis blues and woes, as well as having little time for anyone over the last few months; but I do. Thank you family (I love you), especially you mama and papa. This thesis is as much a product of your hard work and perseverance as it is mine. Thank you to all my friends for your support, and understanding. Special thanks to Miranda (for use of your most wonderful computer), Gerard (for the time spent on graphics and laser printing) and Martin (for inspiration).

Lastly I would like to thank Nick. Thank you for supplying emotional and financial support - as well as much needed chocolate biscuits on demand. My thesis year would have been much bleaker without you.

ABSTRACT

The present research assessed Thurstone's (1927*a*,*b*) method of paired comparisons as a technique for scaling crimes according to their seriousness as perceived by a New Zealand population. In the first of two experiments, 10 crimes, ranging from murder to possession and use of cannabis, were judged for their seriousness by 78 male and female subjects made up of University students and New Zealand Army personnel. Subjects were given a questionnaire in which each of 10 single-word crime descriptors was paired with every other crime. For each of the possible 45 crime pairs, subjects selected the most serious in the pair. Due to numerous boundary probabilities, Edwards (1957) Case V Incomplete Data Scaling Method was employed to construct crime seriousness scales. The method of paired comparisons produced similar results in the ranking of crime seriousness to a previous New Zealand study (Davis, 1992) that employed magnitude estimation scaling. A high level of relative consensus was found between different community groups based upon occupation and sex. This relative consensus extended to crime severity evaluations obtained from the sample employed and the New Zealand Judiciary and Legislature. A second study was carried out to examine whether the degree of relative consensus could be manipulated by varying the seriousness of the crimes. Subjects (24 males and 27 females) were given a forced-choice, computer-generated questionnaire that presented high and low serious versions of each of the 10 crimes employed in Experiment 1 in a one or two sentence description. Crime seriousness was manipulated by varying the quantity of economic or physical harm inflicted upon the victim. In general, the results showed that the degree of crime seriousness altered the position of any given crime on the scale in a very systematic way. Nonetheless, the relative consensus found was much the same as for the first experiment, suggesting that subjects do not envision a specific crime scenario when evaluating a crime's seriousness.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

OVERVIEW

The perception of crime severity lies at the heart of the criminal justice system. Seriousness of a crime in the eyes of the law is expressed by the amount of punishment prescribed in penal codes. As Forgas (1980) asserted, a nation's criminal code is ultimately a composite expression of the seriousness of a set of behaviours (crimes) as perceived by members of that society. Historically, specific members of society that have been involved in the perception of crime severity have been employees of criminal justice: legislators, judges, the police and juries.

However, as Mulvilhill and Tumin (1969, cited in Forgas, 1980) state, the legal code should be an expression of how crimes are perceived by the population at large. There is growing awareness that if public opinions of crime severity are not taken into consideration in legal perceptions of crime severity, the interests of the public and victims, who pay for and suffer these crimes, may be overlooked (Waller, 1982). For this reason the U.S Supreme Court mandated that the socially perceived severity of a crime be considered in assessing the degree of offence severity for penal statutes.

Thus, there is a need for courts and legislators to have a system of reference against which to evaluate the proportion of the sanction relative to the social perceptions of crime seriousness. Such a system of reference can be provided by crime severity scales. Potential applications for seriousness ratings obtained by such scales have been cited as equating legal proscriptions consistent with public opinion (Kadish, 1963, cited in Miethe, 1982) and for creating criminal policy (Levi & Jones, 1985).

Very little has been done in the way of scaling crime seriousness in New Zealand; only two such studies exist to date. Spier, Luketina and Kettles (1991*a*) developed a crime seriousness scale based on court sentencing data. Although a useful research tool for measuring trends over time in offending, and in the seriousness of offences for which sentences are imposed, it is not a crime seriousness scale based upon public perception.

This led Davis (1992) to scale public perceptions of crime seriousness using magnitude estimation. Results indicated that public perceptions of crime severity could be obtained from a New Zealand sample using this scaling method. It was also found that a high degree of congruence existed between the public and the judiciary, with less congruence between the public and legislature, and public and the police. Societal consensus was also found between different subgroups when his sample was examined by age, sex, occupation and whether or not subjects had been a victim of crime within the last 12 months. However, certain disadvantages exist in scaling crime seriousness using the technique of magnitude estimation one being that subjects find the idea of conceptualising the seriousness of a given crime to that of a standard crime difficult (Walker, 1978).

The present study used the method of paired comparisons to scale the seriousness of various crimes. The paired comparison technique may have some advantages over the magnitude estimation technique. Moreover, the paired comparison technique has never been applied to scale crime seriousness in New Zealand.

THE CONCEPTUAL BASIS FOR CRIME SERIOUSNESS RESEARCH

Criminal law is a list of specific forms of human conduct, or crimes, which have been prohibited by political authority (Haskell & Yablonsky, 1974). In order for the prohibition to be effective, the criminal must be punished. There must also be a high level of consensus within society regarding the perception of the relationship between the crime committed and the sanction allocated. The conceptual basis for arriving at a

hierarchy of criminal offences according to public evaluations of crime seriousness is linked to the just allocation of punishment and to the investigation of societal consensus regarding such punishment.

The Just Allocation of Punishment

Sanctions can take on various forms depending upon their purpose. Sanctions can rehabilitate or act as a punishment that deters crime and/or seeks retribution from the offender (McFatter, 1982). For the purposes of punishment, the severity of a crime is central to the fair allocation of punishment to individuals who transgress the law.

As a matter of justice the severity of the punishment should be proportional to the severity of the crime (Hart, 1958). In order for this proportional fit to be achieved, two ranking processes need to be underway simultaneously, one for the offence the other for punishment (Fox & Freiberg, 1990). This concept of proportionality has been of fundamental importance to adherents of both deterrence and retributive (or Just Deserts) sentencing theory.

Although deterrence and retribution are not the only bases for sentencing theory, they are the only models to be guided by the principle of proportionality, where the severity of the punishment allocated should be commensurate with the seriousness of the crime. For example, punishment can be allocated for rehabilitative purposes. Rehabilitative sentencing theory is a treatment-orientated sentencing approach that seeks to "cure" the offender of their criminal behaviour. It is usually applied to certain categories of offenders, such as the young and repeat offenders (Von Hirsch, 1983*a*). As such, rehabilitative sentencing theory is inconsistent with the principle of proportionality. It is based upon the offender (Hamilton & Rytina, 1980) and the perceived consequences of punishment (Warr, Mier & Erikson, 1983) rather than upon the perceived seriousness of the crime.

Deterrent sentencing theory is based upon ideas of crime prevention. The proportional fit between the severity of a crime and the punishment allocated for that crime should be just great enough to deter potential criminals from being attracted by the rewards of any given crime (Beccaria, 1794/1963). However, the deterrence rationale is not based solely upon the seriousness of the crime as the frequency of the crime is also an important component (Warr et al., 1983). As such, the idea of proportionate punishment can be given firmer grounding in retributive sentencing theory than in deterrence, because crime seriousness is the central criterion for retributive sentencing theory (Von Hirsch, 1983*a*).

The seriousness of a crime can be broken down into two components: the extent of harm done and the extent of the criminal's blameworthiness, or culpability. Retributive sentencing theory is based upon these two components. Punishment should be distributed among convicted offenders in a manner consistent with the amount of implicit blame and harm done (Hart, 1958; Von Hirsch, 1983*a,b*). Retributive sentencing theory is an important concept as it forms the underlying ideological basis for the crime seriousness paradigm (Cullen, Link, Travis & Wozniak, 1985).

The crime seriousness paradigm is considered retributive as it focuses solely on the seriousness of offences in terms of harm done and the blameworthiness of the offender. This retributive philosophy is deeply entrenched and even influences research on criminal behaviour. As Cullen et al. (1985) state, research involving the ranking of crimes according to their severity is based on a sentencing system guided by retributive principles. (The more serious the offence the more serious the punishment.)

An important question arises as to the accuracy of this paradigm. Do the public think the punishment should fit the crime, and, if so, do members of the public prefer a retributive sentencing philosophy that fits the punishment to the crime ?

Studies in the area of crime seriousness have supported the concept of proportionality. Public evaluations of crime and punishment have indicated that the punishment should fit the crime (e.g., Blumstein & Cohen, 1980; Geshcheider, Catlin & Fontana, 1982; Hamilton & Rytina, 1980; Warr, Gibbs & Erikson, 1982). Further research (Warr et al., 1983) suggests that it is also safe for crime seriousness research to prescribe to a retributive paradigm, as the paradigm is reflected in public theories of sentencing. The concordance that is emerging from the research literature is that the general public do in fact favour retributive sentencing. It seems that the average lay person is less concerned with utilitarian aims of sentencing (i.e., deterrence and rehabilitation) than with the principle of desert.

For example, Warr et al. (1983) found that respondents used the perceived seriousness of the offence as the central criterion for fixing punishment, as suggested by the retributive theory of criminal punishment. The perceived frequency of an offence, as suggested by utilitarian theories, was not used as a criterion. In addition to this finding, Parton, Hansel and Stratton (1991) showed that the perceived seriousness of a crime is evaluated in retributive terms, that is, by the injury and loss sustained by the victim

Adherence to retributive sentencing theory has been shown to hold even when members of the public are presented with other sentencing theories. For example, McFatter (1982) examined the perceived usefulness of four different punishments (incapacitation, deterrence, rehabilitation and retribution) for different types of crime (car theft, assault, rape and murder). Both punishments and crimes were given to university students and six practising court judges. Importance weights derived from the data indicated that retribution was the most important factor for all crimes among the judges and for all crimes except murder among the students. It seems that the perceived appropriateness of penalties depends primarily on their usefulness in giving offenders what they deserve, rather than on their incapacitative, rehabilitative or deterrence properties.

In summary, public opinion supports the concept of proportionality: the punishment should fit the crime. Furthermore, it has been demonstrated that this matching of crime and punishment is undertaken using the common principle of retribution. This is an important finding as data on public evaluations of crime seriousness are only meaningful if people use the principle that underlies the crime seriousness paradigm.

Although it appears that most people agree that the punishment should fit the crime, it does not necessarily follow that it does. Ideally, public perceptions of crime severity should be in consensus with the legal perceptions of crime severity as expressed in statutory and/or actual punishment. If the legal gradation of crimes is not too far out of line with public opinion, the sanctioning of these crimes will be effective. Legal prescriptions of crime severity receiving a high degree of public support are more likely to result in formal action against violators than are laws lacking in widespread support (Haskell & Yablonsky, 1974). The law in this sense is effective; legal authority is strengthened and violators of such crimes are punished. However, the outcome of a legal prescription of crime severity receiving low public support may result in sanctions for crime becoming ineffective. For instance, the reduction in criminal conduct regarding the crime in question will not take place. A legal prescription that is out of line with public perception may also be counter-productive, meaning that as well as a lack of reduction in criminal offending, society's belief in, and obedience to, legal authority will diminish. A substantial body of crime seriousness research has been directed towards ascertaining the degree of consensus between the legal gradation of crime seriousness and public perceptions of crime seriousness for these reasons.

Several studies have found discrepancies between public perceptions of crime seriousness and legal sanctions imposed upon criminal offenders. Rose and Prell (1955) found that students' seriousness ratings of 13 minor felonies were not correlated with the sentencing policies of the courts in California. Gibbons (1969) found support for these findings when he had respondents assign punishments to 20 different crimes, the punishments of

which ranged from execution to no penalty. Gibbons established that there were discrepancies between public sentiment and legal practices for victimless crimes such as possession of cannabis. The sanctions for these crimes were perceived as being disproportionably high in relation to the seriousness of these crimes. Vehicular manslaughter and child molestation offences were also found to be out of line with public sentiment; the public gave offenders harsher sentences than those meted out by the court. Gibbons also confirmed that when public sentiments and legal penalties did match they were with visible crimes, crimes against the person or those that involved coercive attacks upon property. Thus, crimes that citizens would have heavily punished were usually in line with actual court sentences.

Geshcheider et al. (1982), using magnitude estimation and crossmodality matching, found that although crime severity (as judged by respondents) and punishment (as prescribed by the New York penal code) were related, the relationship was not perfect. But, in general, they found that as the seriousness of the crime increased the prescribed sentence also increased.

When comparing the two studies by Geshcheider et al. (1982) and Gibbons (1969) contradictory findings present themselves. Gesheider et al.'s results obtained 13 years after Gibbons' suggested that the discrepancies between legal sanctions and public perceptions of crime severity involved crimes against the person, where punishment was seen as not severe enough by the public. Differences between the public and legal sanctions were also seen for crimes against property for which the public viewed the punishment as too harsh in relation to the severity of the crime, the reverse of which was found in Gibbons' study. Although caution must taken in the comparison of the two studies (as different scales were used to judge crime seriousness), these results suggest temporal changes. For example, public opinion may have moved towards recommending harsher penalties for property and person orientated crimes while statutory opinion remained unchanged. Thus, the consensus between legal and public perceptions in the

gradation of crime seriousness that existed in Gibbons' study no longer existed in Gescheider et al.'s study. As Geshcheider et al. noted, the punishment may not fit the crime because of the relatively slow response of the judicial system to rapid changes in public opinion on the seriousness of criminal offences.

Rose and Prell (1955) termed this slow response "cultural lag", or a clash between law and normative structure, where the law represents an older cultural evaluation and respondents of crime seriousness research represent a contemporary evaluation. The concept of "cultural lag" has been advanced as one reason why perfect public and statutory consensus may not be reached in crime seriousness research.

Another reason that may account for varying degrees of consensus between the two perceptions may lie in the limitations of crime seriousness research itself. Research seeks to fit the punishment to the crime, rather than to the criminal as the criminal justice system does (Gescheider et al., 1982). Courts deal with individuals and individual circumstances, aspects which are missing from crime seriousness research. The severity of a sentence as evaluated by the courts is a function of two factors: the crime and, to a lesser extent, the criminal history of the offender. It is therefore possible that empirical research may never find a perfect relationship between the public's perceptions of crime seriousness and punishment prescribed by the courts and legislators.

Strathern (cited in Wuillemin, Richardson & Moore, 1986) stated that the viability of a nation's criminal code is determined not only by the degree of consensus between the legal gradation of crime seriousness and public perceptions of crime seriousness, but also by the degree of consensus members of society reach about what are regarded as serious offences. Consequently, another important area within crime seriousness research concerns the degree of consensus that exists between members of the public.

The evaluation of societal consensus

Consensus among members of society is important for two reasons. First, as previously stated, there should be shared standards amongst members of society in order for a nation's criminal code to function effectively. For this reason Rossi, Waite, Bose and Berk (1974) argued that consensus in crime seriousness *"should be reflected in the criminal code, the behaviour of judges and the actions of law enforcement agencies"* (p.224). Secondly, societal consensus is also necessary within the crime seriousness paradigm, for as Rossi et al. (1974, p.224) state: *" To be of any theoretical or practical use, a measure of crime seriousness requires that a society show consensus about the order of seriousness of specific criminal acts".*

Contradictory findings exist in studies that have evaluated consensus between differing societal groups in their perceptions of crime seriousness. Support for consensus concerning the gravity of criminal acts has been drawn from research that consistently shows that the ranking of criminal acts according to their severity is comparable between such diverse groups such as prosecutors (Roth, 1978), offenders and non-offenders (Figlio, 1975), police and the general community (Levi & Jones, 1985) and police and students (Sellin & Wolfgang, 1964). Consensus in the perception of crime severity has also been found to exist across cultures (e.g., Ackman, Normandeau & Turner, 1967).

However, research supporting consensus in the perception of crime seriousness has not gone unchallenged. When studies examining consensus amongst people according to their sex, occupation, socioeconomic status, race, age, whether they had previously been victims of a crime and education were compared, contradictory findings were obtained. While some studies have found little or no differences amongst people in their perceptions of crime severity based upon these factors, other studies have found widespread disagreement (e.g., compare Rossi et al., 1974, with Miethe, 1984).

Sex. While Rossi et al.'s (1974) study showed that no significant sex differences exist, the predominant finding has been that sex differences do exist in the perception of crime severity. The most consistent finding has been that females judge crimes against the person as relatively more serious and property offences as relatively less serious than their male counterparts. For example, Makela (1966) measured sanctions preferred by judges and other groups for different subsets of crimes. Although differences in severity demands for punishment between groups in the general population were small, a sex difference was found. Women were in favour of more severe penalties for offences against persons and more lenient in their suggestions of penalties for offences against property. Gebotys, Roberts and DasGupta (1988) also observed significant sex differences in the ratings they obtained for certain offences. These researchers found that females generated higher seriousness ratings for offences against the person than did males. They also found that males judged offences against property to be relatively more serious than females.

Lamp (1982) also investigated differences between males and females in their perception and evaluation of crime and criminals. Three major sex differences were observed in his sample of student and non-student subjects. First, sex differences were found in the ranking of the most serious crimes. Respondents of both sexes agreed that murder and rape were the two most serious crimes. However, males rated murder as most serious while females rated rape as most serious. Both males and females included arson, robbery and assault in what they perceived to be the most serious crimes, but disagreed as to the relative seriousness and rank order of these crimes. Rape, robbery, assault and child abuse were perceived as relatively more serious by females than males. On the other hand, murder, arson and theft were seen as relatively more serious by male subjects than female.

Secondly, sex differences were also found for crimes perceived as most threatening. Females felt the most threatening crime to be rape, whereas males indicated the most

threatening crime as murder. In general, females were relatively more concerned about robbery, rape and kidnapping than males. However, males were relatively more concerned than females about assault, arson, and theft.

The third finding related to differences found concerning the relative responsibility for the crime committed. Both males and females ranked the individual criminal as most to blame for the crime committed. However, males assigned the second greatest responsibility to the family, while females saw the family as less responsible for the crime committed than society as a whole.

Occupation. Research investigating possible occupational differences in the perception of crime severity and sentencing philosophies is sparce. Occupation was one variable that Samuel and Moulds (1986) investigated, along with the other demographic variables of age, sex, race, marital status, education and income, in their study of recommended punishments. Respondents in their study were a random sample of the general public who were asked in a telephone interview to recommended a punishment (either fine, released on probation or prison term in years) after listening to one of six crime descriptions. The crime descriptions ranged from theft ($20) to murder. It was found that age, race, income, victimisation and occupation all effected recommended punishment. Moreover, people in higher status occupations generally recommended longer sentences for serious injury offence than those in lower status occupations.

However, other researchers such as Lamp (1982) have found no evidence to suggest that occupational differences effect perceptions of crime severity. In fact, agreement has been found between occupations as diverse as students and police (Levi & Jones, 1985; Sellin & Wolfgang, 1964), and prosecutors and students (Roth, 1978).

Socioeconomic status. Very few studies have explored socioeconomic status as a determinant in societal consensus. The research that has been carried out has produced

conflicting results. For example, Walker (1978) examined social class variations in perceptions of crime seriousness, finding that members of lower social classes generally rated crimes (excluding violent crimes) as more serious than those in the higher socioeconomic classes. Levi and Jones (1985) further found that respondents low on the socioeconomic scale specifically regarded property crimes as more serious than respondents in higher socioeconomic classes. However, other studies have found agreement in seriousness ratings between disparate socioeconomic classes (e.g., Davis, 1992; Thomas, Cage & Foster, 1976).

Race. Race is another variable that has received little attention in the area of societal consensus. Hawkins (1980) examined the differences in punishment philosophies across races. Results showed that, in general, significant racial differences existed, the main finding being that blacks tended to assign more punishment for all crimes than did whites. (See also Berk & Rossi, 1977; Miethe, 1984, for similar results.) However, contradictory results have been obtained by Thomas et al., (1976) and and Rossi et al., (1974). Both these studies found that the respondents' race had no effect upon the degree of consensus found amongst subjects in their ratings of crime seriousness.

Age. While few studies have explored age as a factor that may influence consensus, the main finding has been that age does effect both punishment philosophies and consensus amongst respondents. For example, Levi and Jones (1985) found that the relative order of offence seriousness did not change with age for most crimes. However, those considered 'morals offences' and homicide were perceived to increase in seriousness with age. (See Miethe, 1984, Rossi et al., 1974, and Thomas et al., 1976, for similar findings.) However, a few studies, such as Davis (1992) and Sellin and Wolfgang (1964), have found no significant differences in ratings of crime seriousness amongst respondents of various age groups.

Prior victimisation. A number of researchers have examined the differences in both punishment philosophies and seriousness ratings for prior victimisation of respondents. While Davis (1992) and Levi and Jones (1985) found no significant differences between victimised and non-victimised respondents, other studies have. For instance, Gebotys et al. (1988) found a paradoxical relationship between crime seriousness ratings and victimisation. Those respondents who had been recently victimised had lower seriousness ratings than non-victimised respondents. Sparks, Glenn and Dodd (1977, cited in Levi & Jones, 1985) speculate that having been a victim of a crime may lead people to take a less serious view of crime in general. Respondents in their study who had been victimised in the previous year rated every offence lower than those who had been victims at an earlier stage or those who were non-victims. However, it is important to note that ratings of victims of violent and non-violent crimes were not compared in these studies.

Education. Educational attainment has also been explored as a possible factor in influencing the degree of societal consensus and the sentencing philosophies held by respondents. Although little research exists for this variable, differences in educational attainment amongst respondents appears to be the only factor for which studies have had consistent findings. One such study, (Kutschinsky, 1970, cited in Newman & Trilling, 1975), compared several Scandinavian studies regarding attitudes towards punishments, on the basis of education. Two major findings were advanced, the first being that a higher level of education is accompanied by a tendency to want longer sentence lengths and less tolerance of young offenders. It was also found that educational attainment and leniency towards criminals are inversely related, with those with highest educational attainments being the least lenient.

In summary, conflicting findings exist as to whether certain individual background differences influence the degree of consensus found amongst members of society in ratings of crime severity and punishment philosophies. Three possible causes for the conflicting results found in the area of societal consensus, and its determinants, can be

found in the literature. These causes are: the measure of consensus employed, the type of crime assessed and the appropriateness of the statistical analysis used to evaluate the degree of consensus.

Measures of consensus: Agreement versus consistency. The degree of consensus found between different subgroups in society with regard to crime seriousness ratings varies according to whether consensus is measured by agreement (absolute consensus), or consistency (relative agreement) (e.g., Rossi and Henry, 1980; Miethe, 1984). Agreement (or absolute consensus) is attained when respondents give similar quantitative scores to each offence. Consistency (or relative consensus) is achieved when offences are ranked in the same order. It is important to note that the existence of absolute consensus also reflects relative consensus (Rossi et al., 1974). However, one can achieve perfect relative consensus with little or no absolute consensus.

Miethe (1984) and Rossi and Henry (1980) defined two more distinct types of consensus according to the degree of item inclusion, global or local, in the crime seriousness measure. Global consensus refers to absolute or relative agreement across a wide range of differing offences, for example, from murder to minor offences such as parking tickets. Local consensus refers to absolute or relative consensus on items that are close together on the scale, such as crimes against the person.

Miethe (1984) used and reanalysed data from Rossi et al.'s (1974) study, which employed 125 white and 75 black residents, to investigate differences between each type of consensus. Results indicated a high level of global relative consensus between both races across all acts ($r = .897$). Both black and white residents' mean ratings for violent, white collar and victimless crimes were also highly correlated, $r = .930$,with a moderate correlation found for public order crimes, $r = .837$. However, for property offences little local relative consensus was found, $r = .667$. Regression analysis revealed both global

absolute consensus and local absolute consensus for violent, white-collar, and public order offences, but not for victimless or property crimes.

Which is the most appropriate measure of consensus in crime seriousness research and how should this be evaluated? Cullen et al. (1985) suggest using Shrout and Fleiss' (1979) idea that it depends upon whether the raters are conceived as random or fixed. Agreement (or absolute agreement) is appropriate for the random effects model where one randomly selects judges from the population to which one wishes to generalise the results. As such, agreement is appropriate for investigators that wish to generalise about the degree of consensus that exists in a nation. Consistency (or relative consensus), on the other hand, is an appropriate measure if raters are viewed as fixed effects; that is, if the raters' opinions are all that are of interest. Since most investigators wish to generalise crime seriousness to the larger population, Shrout and Fleiss' logic would imply that absolute agreement should be the measure of choice when assessing consensus (Cullen et al., 1985). This is not to say that relative consensus is unimportant. However, its usefulness is greatest when investigating the opinions of specific community groups, such as law makers and law enforcers.

The research objective is also an important determinant in the choice of which type of consensus to use. Miethe (1984) states that while a finding of absolute consensus appears to be the most relevant for policy issue and theoretical development, relative consensus may be sufficient for allocating resources for crime prevention.

The type of crime rated. Different levels of subgroup consensus in seriousness ratings of crime have also been found to exist in relation to the type of crimes rated. Virtually all the studies reviewed have found that both absolute and relative consensus exist for the severity and punishment of traditional crimes, especially violent offences against the person. However, when one looks at property crimes, 'quasi criminal' crimes (such as adultery) and victimless crimes (such as possession and use of drugs), considerably more

variation in both types of consensus is found. For instance, Sparks et al. (1977, cited in Levi and Jones, 1985) found that racial and social class differences existed for property offences which were rated as more serious by blacks than by whites and by working class than by middle class groups. Violent offences, however, were rated similarly by all racial groups and social classes.

Similar findings were obtained by Levi and Jones (1985) who concluded that there was a high level of agreement (both among the public and the police) on offences such as violence but disagreement amongst respondents for property crimes, such as fraud and burglary, and for victimless crime. Moreover, high levels of agreement were obtained for violent offences amongst subjects of differing ages, ethnic groups, socioeconomic class and geographical regions. Hawkins (1980) also found fewer differences between racial groups in punishment assigned for crimes against the person than for crimes against property and white collar crimes.

Kadish (1967, cited in Chilton & DeAmicis, 1975) suggested that victimless crimes are acts about which there is little or no consensus. Chilton and DeAmicis (1975) tested this assumption by using the method of magnitude estimation to obtain seriousness scores from students on 24 separate crimes. Results indicated that crimes which could be described as 'victimising' offences, those threatening physical harm or permanent loss of property, were consistently assessed as most serious. However, victimless crimes, for which harm done was less clear or threatening, were not only assigned lower seriousness scores but also produced considerable disagreement among respondents as to the degree of seriousness for these crimes. In particular there was a considerable lack of agreement (or absolute consensus) concerning the seriousness of the use of all illicit drugs, supporting previous findings by Rossi et al. (1974) and Sellin and Wolfgang (1964). Results such as these have lead commentators in the field of crime seriousness research to argue that consensus may be artificially produced when there is an over-representation of offences such as those against the person for which there is a high level of both relative

and absolute agreement. Because of the small variances found between subgroups for such offences, their inclusion in the scale may inflate ratings of global consensus while masking possible disagreement for other types of crime (Chilton & DeAmicis, 1975; Miethe, 1982; Rossi & Henry, 1980).

The over-inclusion of high and low seriousness crimes, with an under-representation of crimes in the middle range, has also been noted as a possible contributor to artificially producing consensus. Subgroups have been found to agree on the crimes that are most serious and least serious while disagreeing on the seriousness of middle of the range crimes (Rossi & Henry, 1980). Along similar lines Miethe (1984) concluded that the degree of consensus found was guided by the seriousness of the crime rated, as the highest level of individual agreement was found for the most serious acts (violent crime) and the least serious acts (public order offences). Violent crimes were also found to contribute most to the overall level of absolute and relative consensus, whereas property crimes contributed the least to both types of consensus.

In short, research indicates that there is a high degree of consensus (both absolute and relative) amongst people that murder is a heinous crime and that parking tickets are trivial. There is also a high degree of consensus (both absolute and relative) on the seriousness of violent crime against property and the person. However, people differ in their perceptions of seriousness for some crimes, for example, theft and fraud. A low level of consensus is also expressed for non-traditional and victimless crimes.

Statistical analysis. Miethe (1982) emphasised the inconsistency of using measures that ignored individual variability as a yard-stick for consensus when measures of dispersion are the single best indicator of agreement. Unfortunately, such measures of dispersion have been neglected in past studies of crime seriousness research, with a few exceptions (e.g., Hawkins, 1980; Rossi et al., 1974). Past researchers have instead invariably computed correlation coefficients on the basis of aggregate data (e.g., group means and

geometric means). As Rossi and Henry (1980) note, it is possible to attain high correlations between subgroups mean ratings (indicating consensus) but with high individual variability.

Cullen et al. (1985) investigated Miethe's (1982) criticisms by examining the extent to which the nature of data analysis techniques influenced consensus in seriousness evaluations. Results confirmed Miethe's comments on the statistical sources of consensus in crime seriousness. They concluded that traditional measurement techniques that ignore individual variance provide only a limited assessment of the concept of consensus and can inflate levels of consensus found.

Conclusions in the area of consensus

There are no easy conclusions regarding the existence of consensus in public evaluations of crime seriousness; varying degrees of consensus have been found. The degree of consensus among members of society has been shown to differ according to whether consensus is measured by agreement (absolute consensus) or consistency (relative consensus) and the rating task given to respondents (i.e., the type of crime evaluated). Furthermore, inappropriate use of statistical procedures can inflate levels of consensus.

However, it can be stated with some certainty that for traditional high and low seriousness crimes, considerable consensus exists. One may speculate that for these crimes, measures of crime seriousness are tapping into societal consensus. On the other hand, when it comes to specific criminal acts, not traditionally criminal, with no visible victim or of moderate seriousness, individual differences come into play.

Spier et al. (1991*a*) excluded public opinion in their seriousness scale of imprisonable offences, although they asserted that, ideally, a measure of seriousness should include public views. The reason for omitting societal views was because of the doubt they cast

over subgroup consensus. Thus, it was stated: *"..it is probably not valid to apply to the New Zealand situation crime seriousness scales based on community surveys developed outside this country"* (p.16). They further asserted that because of the doubt surrounding the issue of consensus, different cultural groups may not have adequate representation.

Although Spier et al.'s (1991*a*) scale is useful for the purpose of research, it comes no closer to evaluating public opinion of the seriousness of crime in New Zealand. In fact, these researchers reject the possibility of conducting a crime seriousness survey within New Zealand based upon the public's perception of crime seriousness. Their research also does not contribute towards the evaluation of consensus, or lack of consensus, in different cultural and minority groups within New Zealand, which may effect crime seriousness ratings. As both Chilton and DeAmicis (1975) and Miethe (1984) conclude, more systematic examination of different types of consensus, and factors that contribute to consensus is required. Despite Spier et al.'s pessimistic view that ratings of crime seriousness obtained from members of the general public can never produce a valid scale of crime seriousness, several such scales have been produced overseas. These scales include category scaling, magnitude estimation scaling and paired comparison scaling.

THE SCALING OF CRIME SERIOUSNESS

The measurement of crime seriousness is not as straight-forward as it may seem. The measurement problem inherent in crime seriousness research is that perceived severity of criminal offences is a subjective property with no conventional metric. Methods for scaling crime seriousness must enable investigators to quantify a qualitative phenomenon. To overcome this difficulty, scaling techniques have been adopted from the field of psychophysics, which measure subjective levels of sensations, and from the early days of attitude scaling research (Thurstone, 1927*a*,*b*).

Unidimensional scaling of crime seriousness over the years has consisted of three psychometric techniques, these being category scaling, magnitude estimation scaling and paired comparison scaling.

Category scaling

In category scaling the judgements of subjects are placed into categories, or intervals. The method involves constructing a rank-ordered continuum of stimuli. In a typical survey that measures crime seriousness using category scaling, subjects are presented with a range of crimes in either summary label or vignette form. For each crime, subjects are required to locate the seriousness of the crime along a scale of successive categories. For example, a numerical scale may be given to the subject ranging from 1 (low seriousness) to 9 (high seriousness). The subject would then rate the seriousness of the crime by circling the number that represented their perception of the seriousness of the crime. Scale estimates for crime seriousness using category scales are determined from the sample distributions of crime seriousness ratings. Estimates of the magnitude of crime seriousness and individual variability can also be obtained (Bridges & Lisagor, 1975).

Rossi et al. (1974) found that when they employed category scaling in their study, most respondents performed the task easily. They also found that the distribution of ratings tended to be more dense on the high seriousness end of the 9-point scale, with 9 frequently being assigned to high seriousness crimes. Rossi et al.'s results indicated that crimes clustered into property, person and misdemeanour crimes. Crimes against the person, especially murder, received high seriousness ratings. Crimes against property received significantly lower ratings and offences classified as 'misdemeanours' (such as disturbing the peace) received the lowest ratings available.

The category scale has the advantage of being easy to visualise and understand (Figlio, 1975). However, it is limited in that it is numerically constraining; subjects can only assign numerical values to crime seriousness provided to them by the category scale (Stevens, 1975). It is perhaps for this reason that it has rarely been used in the scaling of crime seriousness.

Magnitude estimation scaling

Magnitude estimation scaling has been the most widely used method for scaling crime seriousness. This scaling procedure was developed by Stevens in the 1950s and popularised in crime seriousness research by Sellin and Wolfgang (1964). Magnitude estimation is a direct scaling procedure based on an unconstrained matching of one continuum to another. It involves judging strength and order of stimuli on an expansive and continuous scale.

In a typical crime seriousness study employing magnitude estimation, subjects are asked to assess crimes of varying seriousness. These are usually presented as descriptions, as seen in the instruction below. In order to assess an offence for its seriousness, the offence must be judged against an offence whose seriousness is a 'standard'. For example, in Sellin and Wolfgang's survey of crime seriousness (1985, cited in Collins, 1988) respondents were presented with the following instruction: *"A person steals a bicycle parked in the street. This has been given a score of 10 to show its seriousness. Use this first situation to judge all the others . For example, if you think a situation is 20 TIMES MORE serious than the bicycle theft, the number you should tell me should be around 200, or if you think it is HALF AS SERIOUS, the number you should tell me is around 5"* (Collins, 1988; pp.61-62).

Subjects are also told that there is no constraint on what number to use, as long as they show how serious they think the crime is. Crimes are scaled according to their geometric

means. Individual scores are used to produce a scale value based on a geometric mean rather than an arithmetic mean value. The reason for this is that the geometric mean is less susceptible to extreme scores. As with category scaling, the degree of individual variability across subjects on the seriousness of crimes can be computed.

Using magnitude estimation, Sellin and Wolfgang (1964) constructed a seriousness index of delinquent events. Sellin and Wolfgang later attempted to employ this scaling technique in a National Survey of Crime Severity the limitations of which were later demonstrated by Parton et al. (1991). The National Survey of Crime Severity was carried out to provide a measure of crime seriousness according to public opinion.

Subsequent studies using Sellin and Wolfgang's (1964) basic research design and scaling technique have focused on three main areas: the relative seriousness of various crimes (e.g., Cullen et al., 1985), the degree of subgroup consensus (Figlio, 1975; Hawkins, 1980) and cross cultural investigations (Ackman and Normandeau, 1967; Stevens, 1975; Newman, 1976; Kavalseth, 1980; Davis, 1992).

Results of these replications have been consistent, with two major findings being established. Firstly, ratings obtained for various criminal acts show that certain crimes are consistently rated as more serious than others, with crimes clustering into violent, property and victimless categories. Secondly, consensus has been found to exist for traditional and violent crimes. However, various racial, educational, sex, socioeconomic class and age differences have been found to influence the degree of consensus found for other categories of crime, for example, victimless crime.

The presumed basis for the continued use of magnitude estimation scaling as a method for acquiring public perceptions of crime severity over other methods, is that magnitude estimation has been presented as having ratio scale properties (Gescheider et al., 1982; Stevens, 1975). However, as argued and demonstrated by Marks (1982, cited in Parton

et al., 1991) and Duncan (1984), an absolute zero point cannot be achieved with magnitude estimation, without which a true ratio scale is not obtained. Thus, magnitude of estimation is, in effect, an interval scale with an arbitrary zero point - as is paired comparisons.

Apart from the questionable zero point, other limitations are contained within magnitude estimation scaling. One such limitation is that subjects are required to display greater abstraction in their thought processes (Figlio, 1975). For example, people have to evaluate the seriousness of a crime in terms of a fraction of the standard crime given. This evaluation may be difficult for subjects (Walker, 1978).

For this reason, it is necessary to assess if a respondent can perform the magnitude estimation procedure through use of a training task (which may be complex) before any data are collected (Parton et al., 1991). A common example of such a training task is to have respondents assign numerical values to estimations of line lengths so that the numerical values and line lengths correspond (Zwislocki, 1983, cited in Parton et al., 1991). Some data collected using magnitude estimation may be flawed as some researchers have not used a training task (e.g., Sellin & Wolfgang, 1985, cited in Parton et al., 1991), or they have carried out the training task in an uncontrolled environment (e.g., Davis, 1992).

Analytical problems with the method have also been observed. Collins (1988) noted two such problems. Firstly, because respondents of magnitude estimation methods are not numerically constrained, and can thus assign a seriousness score of anything from zero up, artificial upper limits must be set. The problem that arises is that although a standardised upper limit has been established (Collins, 1988), crime seriousness research using magnitude estimation has not yet utilised these standards. This limits the ability to compare results obtained from different surveys. The second problem that arises also involves the issue of standardisation. Geometric means are commonly employed to

obtain scale vales. However, difficulty occurs when data obtained produce zero cases for X_j, the seriousness score recorded for the ith respondent (i =1, 2,...n). This becomes problematic as the geometric mean is undefined when X_j is equal to zero, and as such, cannot be used. Collins noted that previous research has not stated how data with values of zero for any X_j have been modified or eliminated. Thus, variations between studies in their results may merely be due to the different practices researchers have used to address these zero scores. Again, although Collins established a standardisation procedure for this problem, it has not been used by researchers to date. However, it bears mentioning that the probable reason for researchers not employing the two standardising procedures developed by Collins is that they themselves are somewhat arbitrary.

Paired comparisons scaling

Less popular in crime seriousness research has been Thurstone's (1927*a,b*) scale of comparative judgement. Scaling using the method of paired comparison is an indirect scaling procedure and involves subjects comparing pairs of stimuli. For each pair of stimuli presented, a subject must choose the one which dominates the other on some specified attribute, or attributes, such as seriousness.

In the measurement of crime seriousness, subjects typically are given pairs of crimes (in random order) and are asked to circle or underline the most serious of the two. If subjects find this impossible for any pair, they are usually asked choose randomly which they find most serious. That is, the method normally uses a forced-choice technique, with a rating of equal seriousness not being allowed.

The basic *law of comparative judgement* is bounded by two assumptions. The first assumption is that subjective stimuli, that differ to some extent in magnitude, give rise to a *discriminal process* in an observer which can be translated to a value on a psychological continuum. A *discriminal process* can be defined as a theoretical construct that

represents the evaluative response that takes place in an individual when having to distinguish between stimuli on some given attribute. The most common discriminal process associated with any given stimulus Thurstone (1927*a*) called the *modal discriminal process.* It is this value of the modal discriminal process which provides the scale value of the stimulus on the psychological continuum.

The second assumption underlying Thurstone's law of comparative judgement is that repeated presentation of a stimulus will not necessarily give rise to the same magnitude of subjective experiences in an individual, or individuals, but rather, as Thurstone (1927*a*) postulated, there is a normal distribution of such discriminal processes. The standard deviation of this distribution of discriminal processes associated with a stimulus Thurstone (1927a) termed the *discriminal dispersion* . The discriminal dispersion, just as for the scale values, may change for different stimuli.

The scale difference between the discriminal processes of two stimuli which are involved in the same judgement for any single presentation of a pair of stimuli, is referred to as the *discriminal difference* (S_i - S_j), where S_i is the proportion of judgements $_i$ greater than $_j$, and S_j, the proportion of judgements $_j$ greater than $_i$. The discriminal difference is a function of the proportion of judgements $_i$ greater than $_j$.

Figure 1 illustrates a psychological continuum with the distributions associated for four stimuli, the associated scale values being S_1, S_2, S_3 and S_4 respectively. Each stimulus carries with it a discriminal dispersion, which is, as seen, not always equal for differing stimuli. Also shown in Figure 1 are the modal discriminal process and the discriminal difference.

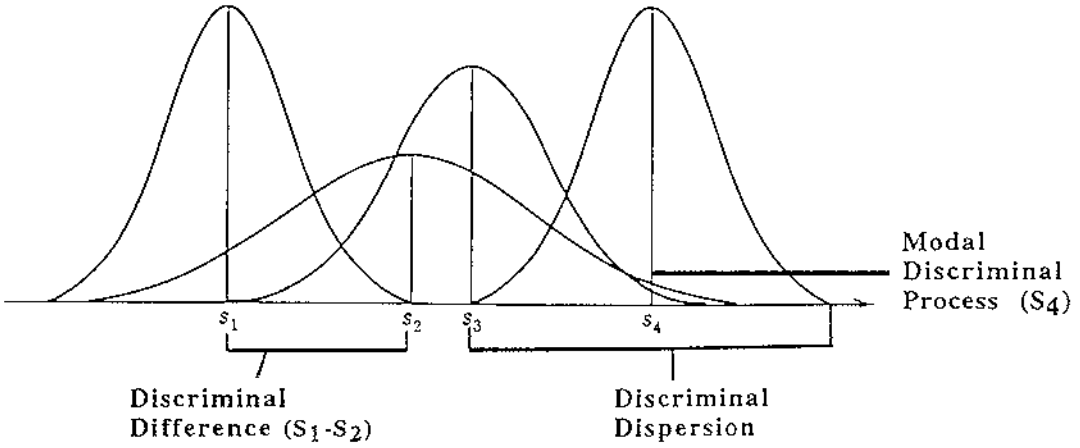

Figure 1. Distributions of the psychological continuum of discriminal processes associated with four stimuli. Also shown are examples of the discriminal difference, the discriminal dispersion and the modal discriminal process. (After Torgerson, 1958; p.157.)

The actual law of comparative judgement is based upon the scale separation of the modal discriminal processes (S_i and S_j) on the psychological continuum. The scale separation is established by converting comparative judgements into empirical frequencies that correspond to the number of times i is judged greater than j, which is expressed as a proportion p_{ij}. The values of p_{ij} are then transformed into unit normal deviates, or *z* scores, z_{ij}. The relationships between p_{ij} and z_{ij} are described graphically in Figure 2.

It may be observed that when p_{ij} is equal to 0.50, or p_{ij} is equal to p_{ji}, as in (a) in Figure 2, then z_{ij} will correspond to the zero point on the abscissa. When the proportion of comparative judgements, p_{ij}, is greater than 0.50, or $p_{ij} > p_{ji}$, as in (b) in Figure 2, then z_{ij} will fall to the right of the zero point of the abscissa and be positive in sign. Conversely, when p_{ij} is less than 0.50, or $p_{ij} < p_{ji}$, as in (c) in Figure 2, then z_{ij} will fall to the left of the zero point on the abscissa and be negative in sign.

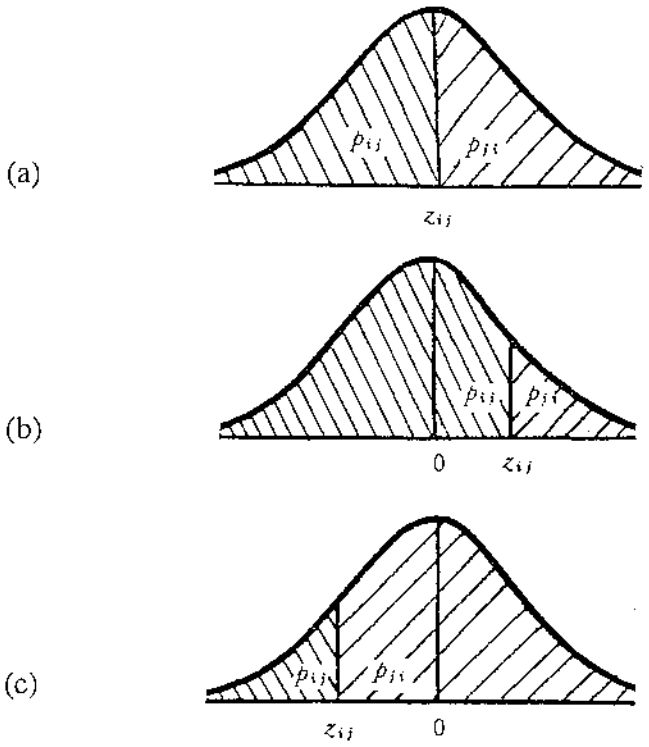

Figure 2. The unit normal deviate for transformation for p_{ij}. (a) $p_{ij} = p_{ji}$, (b) $p_{ij} > p_{ji}$ and (c) $p_{ij} < p_{ji}$. Measurements on the abscissa are in units of standard deviation. (After Edwards, 1957, p 26.)

Thurstone (1927*a*) expressed the scale separation between the two modal discriminal processes, S_i and S_j, in terms of the equation:

$$Z_{ij} = \frac{S_i - S_j}{\sqrt{\sigma_i^2 + \sigma_j^2 - 2r_{ij}\,\sigma_i\,\sigma_j}} \qquad (1.0)$$

In which:

$\mathbf{S_i - S_j}$ represents the mean psychological distance between the stimuli $_i$ and $_j$ respectively;

$\mathbf{Z_{ij}}$ is the standard deviate corresponding to the number of times stimulus $_i$ is judged greater than stimulus $_j$;

σ_i and σ_j are the discriminal dispersions (standard deviations) of the distributions S_{ij} and S_{ji} respectively;

r is the coefficient of correlation between S_{ij} and S_{ji}.

There is one such equation for every pair of stimuli, of which there are:

$$ {}_nC^2 = \frac{n\,(n-1)}{2} \qquad (1.1) $$

Where ${}_nC^2$ is the number of combinations of n stimuli taken 2 at a time. Division by 2 is required because the order of presentation of any pair of stimuli is ignored.

Practically, equation 1.0 can only be applied to the judgements of a single observer who compares a series of stimuli by the method of paired comparison, when no equal judgements are allowed. The size of each $(S_i\text{-}S_j)$ separation can only be determined if the values are known on the right-hand side of equation 1.0. Z_{ij} is found from the knowledge of the obtained proportion p_{ij}; however, the remaining parameters are unknown. For this reason, Thurstone (1927*a*) distinguished "five cases" with respect to the application of the law of comparative judgement. (See Coombs, Dawes & Tversky, 1970; Guilford, 1954; Thurstone, 1927*a* and Torgerson, 1958 for details.) As Case V is the simplest and most widely used in psychological scaling (Coombs et al.), it is the only case described here.

Case V rests on the assumption that equal dispersions in judgement represent equal distances along the subjective continuum and that the correlational term is the same for all pairs. The first assumption of Case V is that the standard deviations are all equal. If $\sigma_i = \sigma_j = \sigma$, then equation 1.0 becomes:

$$ Z_{ij} = \frac{(S_i - S_j)}{\sqrt{\sigma_i^2 + \sigma_j^2 - 2r_{ij}\,\sigma_i\,\sigma_j}} $$

$$ = \frac{(S_i - S_j)}{\sqrt{2\sigma^2 - 2r_{ij}\,\sigma^2}} $$

$$ = \frac{(S_i - S_j)}{\sqrt{2\sigma^2\,(1-r_{ij})}} \qquad (1.2) $$

Case V further assumes that the inter-correlations are all equal to one another so that equation 1.2 can be rewritten as:

$$Z_{ij} = \frac{(S_i - S_j)}{\sqrt{2\sigma^2 (1-r)}} \qquad (1.3)$$

Under the second assumption made, $\sqrt{2\sigma^2 (1-r)}$ is held constant and is the common unit of measurement of the scale separations of the various pairs of stimuli. This common unit of measurement can be given the value of 1.00, without loss of generality, so that:

$$Z_{ij} = S_i - S_j \qquad (1.4)$$

Thus, under these two assumptions, $Z_{ij} = \frac{(S_i - S_j)}{\sqrt{\sigma_i^2 + \sigma_j^2 - 2r_{ij}\,\sigma_i\,\sigma_j}}$ equation 1.0 simply becomes $Z_{ij} = S_i - S_j$ equation 1.4. As Mosteller (1951, cited in Edwards, 1957) has shown, Thurstone's procedure for finding the scale values of stimuli is a least squares solution.

Thurstone (1927*b*) was the first to apply the law of comparative judgement to the scaling of crimes and offences as perceived by the public. Thurstone presented 266 university students with 117 (19 possible pairs taken two at a time and in random order of presentation) pairs of crimes to assess the perceived severity of the crime in relation to all other crimes presented. The crimes ranged from murder to vagrancy. Resulting data were turned into a proportion matrix and scaled using the Case V version of the scaling model. The scalability of these offences using the method of paired comparison was demonstrated, his model fitting the data exceedingly well as the data matrix could be reconstructed from the scale values with only a small margin of error.

Since Thurstone's (1927*b*) study, various replications have used the method of paired comparisons to assess the perceptions of crime severity (Borg, 1988; Carroll, Pine, Cline & Kleinhans, 1974; Coombs, 1967; Krus, Sherman & Krus, 1977; Rose & Prell, 1955). The general conclusion that can be made from research using Thurstone's (1927*a,b*) scaling procedure is that crimes can be plotted on an interval scale, according to their seriousness, by employing the theory presented by Thurstone. Although the scale

constructed using Thurstone's law of comparative judgement has no ratio properties, (i.e., one crime cannot be described as X times more serious than another), the distance between crimes according to their seriousness can be examined.

Sex differences have also been found to exist in the seriousness ratings of crimes as measured by paired comparisons. The principle difference found is that females judge both property related and person related crimes as relatively more serious than males (Borg, 1988; Coombs, 1967). However, most research using paired comparisons scaling has investigated temporal changes in the relative perception of crime seriousness. For example, Coombs (1967) found that although there was a high level of concurrence in the ranking of crimes between his study and Thurstone's (1927*b*) study, several temporal differences existed. Coombs' equivalent sample of American students perceived crimes against property and crimes against the person as relatively more serious than sex crimes, with the exception of rape. Rape, the highest ranking offence in 1927, dropped to second place in 1966 and has continued to stay in this position (Borg, 1988; Carroll et al., 1974; Krus et al., 1977).

Krus et al. (1977) followed up both Thurstone's (1927) and Coombs' (1967) studies, finding that the trend originally observed by Coombs continued. Sexual crimes continued to descend the seriousness scale while crimes against property and the person were perceived as relatively more serious, with crimes against the person as being the most serious. Social issues of the time and changing perceptions of these crimes were identified as possible explanations for the shift in the perception of crime seriousness.

Borg (1988), using Thurstone's (1927*b*) method, tested a sample of 80 West German subjects, again demonstrating temporal changes. In general, crimes against the person were seen as relatively more serious than crimes against property. However, compared to Thurstone's sample and Coombs' (1967) sample, crimes against property rose in relative seriousness and crimes against the person fell in relative seriousness.

The law of comparative judgement is not without its limitations. One weakness that has come to the attention of researchers with the method of paired comparisons is the difficulty of indeterminate cases of 1.00 and 0.00 proportions (Edwards, 1957). These values of p_{ij} equal to 1.00 or 0.00 are indeterminate as the distributions of the discriminal processes for the two stimuli in question have failed to overlap. When distributions of discriminal process fail to overlap it is usually because comparative judgements have been obtained for stimuli that lie at opposite ends of the psychological continuum.

Researchers working with the method of paired comparisons are advised to treat judgements for which p_{ij} are 1.00 or 0.00 as missing data in practical scaling applications (Torgerson, 1958). Edwards (1957) further suggested that judgements of p_{ij} equal to or greater than 0.98 or equal to or less than 0.02 are also to be treated as missing data. The main reason for treating proportions of 0.98 (and above) and 0.02 (and below) as missing data is that these proportions produce z values that are situated at the extreme ends of the normal distribution. These z values are problematic as they have larger normal deviate unit distances between them than those found at the centre of the normal distribution. Edwards (1957) provides the example of moving from p_{ij} 0.98 to p_{ij} 0.99 with the corresponding shift in z_{ij} from 2.054 to 2.326. However, moving from p_{ij} 0.50 to pij 0.51 only produces a corresponding shift of z_{ij} from 0.00 to 0.025. The distance between z values in the first instance was 0.272 normal deviate units, while in the second only 0.025 normal deviate units. Thus, the more towards 1.00 or 0.00 the proportions obtained are, the more unstable the z values become. Others have suggested alternative p_{ij} boundary standards (e.g., see Guilford, 1954).

A second limitation of Thurstone's method of paired comparisons is that the method employs forced-choice (Borg, 1988). In the usual data collection procedure, a subject is forced into deciding whether stimulus $_i$ dominates stimulus $_j$. As such, the subject cannot state that they perceive both stimulus' as equal in magnitude along a psychological

continuum. When perceived inequity does occur, subjects are instructed to chose one stimuli in the pair at random.

Borg (1988) argued that this method has its advantages in psychophysics but when applied to attitudinal scaling subjects often protest against this format because of the difficulty they have in randomly chosing one stimuli if both are perceived as equal on an attitudinal dimension. However, this random choice is no more difficult than the decisions that subjects are required to make in a magnitude estimation study. Moreover, this perceived limitation is not a good reason for not employing the method of paired comparisons for scaling crime seriousness. The validity of a scaling technique is determined by theory and data, not the opinion of respondents using the technique.

As well as being grounded in theory, the method of paired comparisons has several practical advantages justifying its use. For instance, one advantage is that as stimuli are presented only two at a time subjects need only focus on one pair of crimes at any one time. Another advantage of the method is that the instructions are easy to comprehend, and little or no training is required. An example of the simplicity of the paired comparisons technique is given by Walker (1978) when she compared category scaling, magnitude estimation scaling and paired comparisons scaling. Results showed that subjects had little difficulty understanding the paired comparisons questions; subjects could easily decide which of the two offences was the most serious. However, the use of the techniques of category and magnitude of estimation scaling were found to be more difficult. One reason that can be offered for this difficulty is that both category and magnitude estimation scaling procedures require subjects to assign numerical values to evaluations of the seriousness of offences. This is a process conceptually more demanding than choosing which crime out of a pair is more serious.

However, even though paired comparisons scaling has no such numerical constraints, it remains problematic as to which scaling technique should be used on theoretical grounds.

Indeed, one could argue that all three techniques have some validity and comparing the results of their applications may be more profitable than trying to decide which is the "best" technique.

Comparison of crime seriousness scales

Although little comparative research has been carried out, the general conclusions from those studies that have been done is that similar rankings of crime seriousness are obtained from all three scaling techniques. As such, it can be stated that crime seriousness appears to be a robust phenomenon that can survive alternative measurement approaches.

Bridges and Lisagor (1975) compared category scaling and magnitude of estimation scaling. Their results confirmed that magnitude and category scaling techniques produced similar distributions and estimates of seriousness magnitude. These results supported prior research by Stevens (1972, cited in Bridges and Lisagor, 1975). Another study carried out by Walker (1978) found a considerable degree of consistency in the ratings of crime seriousness when the same subjects assessed the relative seriousness of a set of offences using category scaling, magnitude estimation scaling and paired comparisons scaling. Because of this consistency, Walker concluded that numerically assessing the seriousness of a set of offences is a meaningful operation for the general public, irrespective of the scaling technique employed. Sheley (1980) reviewed both magnitude estimation and paired comparisons survey procedures and found little evidence to suggest that one procedure distorted crime seriousness ratings more than the other.

In summary, it seems that all three scaling techniques used for crime seriousness may have their advantages and disadvantages. Moreover, similar results are obtained when any one of the techniques is used to scale crime seriousness. Therefore, the choice of technique appears somewhat arbitrary. For the present investigation it was decided to

scale crime seriousness using the method of paired comparisons for two important reasons. First, the method allows subjects to focus on just two crimes at a time and requires a simple judgement. Thus, the technique is simple and straight-forward and does not require training subjects. Second, crime seriousness has been scaled in a New Zealand study using only magnitude estimation (Davis, 1992). The method of paired comparisons was used to find out if similar conclusions about crime seriousness in a New Zealand setting can be made using a different scaling technique.

USING PUBLIC PERCEPTION OF CRIME IN DETERMINING POLICY

By far the most discussed use of crime severity scales in the research literature is the employment of these scales for criminal justice policy. Many commentators and researchers in criminology are of the opinion that crime and justice policy should be influenced by what the public wants. Our society accepts the notion that the role of the state is to serve its citizens. The state is therefore under increasing pressure from society to provide greater accountability. Waller (1982) remarks that the concept of state accountability is comprised of two factors. First, there is the view that one must show how public funds have been used. Secondly, there is the opinion that public funds should be used to serve the needs of the public, reflecting the view that tax monies should be used for what the public wants.

Crime seriousness research in public perceptions of crime seriousness can provide policy information to assess the wants of society as well of as to help meet objectives of crime and justice policy. Criminal justice policy objectives are concerned with maintaining public order, maintaining public confidence in the rule of law and in reducing the incidence of crime (Home Office, 1977, cited in Waller, 1982). As such, policy implications for crime seriousness research include issues related to sentencing practices and allocation of scarce crime resources. Both promise and problems are inherent in framing criminal law around public evaluations of crime seriousness.

Promise

Sentencing practices, governed by legislative statutes and allocated by judges, play an important part in maintaining public order. Public perceptions of legitimacy and fairness of these sentencing practices are obtained when public perceptions of crime severity and legal prescriptions of crime severity are in consensus. Only then is public confidence in the rule of law maintained.

Considerable emphasis has been placed on utilising crime severity scales to provide a measure of the appropriateness of sentencing practices according to public perceptions (e.g., Hoffman & Hardyman, 1986). Some authors have gone so far as to suggest that an empirically derived offence severity scale should be the basis for determining statutory penalties using a retributive sentencing model (e.g., Nevares-Muniz, 1984).

As it stands, sentencing decisions are influenced only by officially approved considerations embodied in statute, practice direction or case-law. However, sentencers admit to having some regard for what they believe to be public opinion (Walker, 1985). Unfortunately, as Walker asserts, sentencers' beliefs of public opinion are based upon occasional protests in the news media rather than on scientifically planned surveys of relevant public opinion. As Douglas (1986, cited in Fox and Freiberg, 1990) commented: *"Reliance on the media as a guide to public opinion squares poorly with its assessment of the problematic nature of media reports as the basis for understanding an underlying reality "* (p.332).

Concordance between legal and public perceptions of crime seriousness can also be achieved if public opinion is incorporated into revamping or creating statute maxima for criminal offences (e.g., Fox and Freiberg, 1990). As Davis (1992) asserts, the New Zealand criminal justice system is acting under the influence of many outdated maximum penalties developed over a century ago *" with no rational basis or relevance to modern*

views on the seriousness of crimes " (p.19). Davis further notes that any revision of the 1961 Crimes Act is done without the knowledge of public opinion, leading to unchanged and unchallenged concurrence between the attitudes of society and legislature. Thus, there may no longer be a high concordance between the perceived severity of a crime by both the public and dispensers of justice and statutory punishment.

Crime seriousness scales also have the potential for measuring disparity between law makers and law enforcers. Previous studies of the sentencing practices of high court judges in Victoria (Australia) have shown that there is a wide disparity between the scales of gravity applied in actual practice and those defined by legislation. For example, Freiberg and Fox (1986, cited in Fox and Freiberg, 1990) found that the ordering of crimes according to the severity of statutory maximum penalties bears little relationship to their ranking according to the sentence actually imposed.

Because police are concerned with maintaining public support, so that crime can be reduced, there should also be a high degree of consensus between the service it visibly provides and the service the public demands. Leibrich, Galaway and Underhill (1984) created a relative seriousness scale based on the urgency of police clearance in New Zealand. High ratings were given to many offences which in most seriousness scales, based upon public opinion, have been rated as much lower relative to the serious offences. For example, wounding with intent and aggravated assault were rated as being less serious than selling or supplying cannabis.

Spier et al. (1991*a*) asserted that applications of Leibrich et al.'s (1984) scale appeared limited. However, such a measure of crime seriousness based upon police clearance, when compared to public evaluations may show up inconsistencies between public and police opinion and police and legislative opinion. For example, it may be that although the selling of cannabis has a higher priority for police clearance than aggravated assault,

both the public and the police may perceive selling cannabis as less serious than aggravated assault.

Policy implications for crime seriousness scales also extend to their use as a tool for the allocation of crime resources needed for the reduction of crime (Roth, 1979, cited in Miethe, 1984). For example, the allocation of scarce police resources and time may be guided by public evaluations of crime seriousness. Increasing the number of officers on routine preventative control without changing their tasks may have no effect on either reducing crime or public feelings of security (Kelling, 1974, cited in Bowie & Elliston,1982). Tasks and priorities of police time and clearance may be changed with the incorporation of public opinion of crime severity. This would also increase public support and feelings of security, as the visible services provided by the police would be in concordance with what the public wants.

The lack of consensus found across subjects in crime seriousness research concerning victimless crimes also bears upon decriminalisation policies. Lack of consensus for victimless crimes may equate to the suggestion that they are offences in which the harm is unclear or at least questionable (Chilton & DeAmicis, 1975). As such, these crimes should not have severe penalties allocated by legislators and judges, or have high police priority. If harsh sentences are given, crime resources are seen to be misappropriated, and the police are seen as spending a disproportionate amount of time and manpower on victimless crime, public sentiment is likely to be out of line with the law. Repercussions for this lack of consensus between law and society may lead to the loss of policy objectives. Public order may be threatened as the law and its enforcement agencies lose their legitimacy and the actual rate of crime is not reduced as the public are uncertain as to whether a victimless crime actually constitutes a crime. For these reasons, legislators should take into account whether public perceptions of crime severity for victimless crimes are out of step with those of the law, and as such, be decriminalised.

The decriminalisation of victimless crimes is unlikely to provide any detrimental effects on society but has the potential to increase public legitimacy. For example, Suggs (1981) found that when Nebraska introduced a law which decriminalised the possession of small amounts of cannabis it was a success in terms of increasing legitimacy of the law. The legislation, as evaluated by the public and police, was regarded as less traumatic than the old law even though harsher fines were imposed for possession (these harsher fines replaced prison terms and court appearances). Suggs also found that decriminalisation of cannabis use did not lead to increased cannabis usage.

However, although it can be argued that public opinion should play a part in maintaining policy objectives, policy should not be a direct reflection of public opinion. There may be legitimate reasons why officials responsible for policy choose not to follow public opinion; afterall, both political and practical considerations also influence the policy process (Waller, 1982). Another reason policy cannot mirror public opinion is that there are problems inherent in using public perceptions of crime seriousness.

Problems

Although researchers and commentators agree that, ideally, public opinion of crime severity should be used as input for policy decisions, many argue that this may not be achievable (e.g., Fox and Freiberg, 1990; Spier et al., 1991*a*). If research on public perceptions of crime severity are to be used for policy making, current problems associated with using public perceptions of crime seriousness must be acknowledged.

There appear to be two major problems, the first of which is the concern that the general public may be overly punitive. The second problem is the probable lack of consensus in perceived crime seriousness among the various groups in society.

Grim retributivism. As noted previously, empirical evidence exists that shows that the public favour a sentencing system based upon just desserts. As McAnany (1981, cited in Fogel & Hudson, 1981) states: *" The fear of many reformers is that given the opportunity, the local community would exercise grim retributivism against the unwelcome stranger, the deviant, the reject in society "* (pp.45-46).

These fears are by no means unfounded; surveys of public opinion have shown that people believe that the courts do not deal harshly enough with criminals, that stiffer sentences would decrease recidivism and that leniency by the courts is an important cause of increasing crime (Doob and Roberts, 1988, cited in Fox and Freiberg, 1990). At the time of writing 290, 000 New Zealanders are petitioning for harsher sentences to be meted out to violent criminals. Cullen et al. (1985) and Rossi and Henry (1980) express concern about basing the criminal justice system too much on community standards of retribution. Such action may subordinate criminal sanctions based upon rehabilitation of offenders.

The issue of consensus. Definitive conclusions regarding the existence of consensus within groups of society are lacking in crime seriousness research. A variety of variables have been found to effect the degree of consensus (e.g., sex and educational attainment of the respondent). The lack of agreement across various groups, while in itself not posing insurmountable problems, may effect the degree to which public evaluations of crime seriousness influence policy decisions. The more disagreement perceived by the policy makers, the more likely it is that they will want to downplay public input. The question of how much consensus is required before policy makers would treat the public's perception of crime seriousness with respect is moot (Miethe, 1984). Miethe makes the further point that a similar question can be asked of the quality of research on perceived crime seriousness.

Revised implications for using public perception in determining policy

Clearly, there are some serious problems facing the utility of the public's perception of crime seriousness to policy making. However, public evaluations of crime seriousness may still provide useful feedback to criminal justice policy makers, enforcers and dispensers. Attempts to translate seriousness ratings into legal sanctions would be best accomplished under conditions of absolute consensus. However, the finding of relative consensus is considered sufficient for the allocation of crime resources (Miethe, 1984). Further, both relative and absolute consensus exist for traditional and high and low serious crimes. For these crimes there may be room for using it as corrective feedback, such as Hoffman and Hardyman (1986) suggest, and could possibly be given weight in framing sentencing revisions and allocation of crime resources.

In summary, promise and problems of crime seriousness scales constitute a fine balancing act in their use for creating criminal justice policy and maintaining criminal justice policy objectives. The problems inherent in crime seriousness research must be brought into perspective and weighed against the problems inherent in the present system of policy creation for sentence and crime resource allocation. At present, the allocation of offences to a higher or lower class of seriousness, or the reallocation of crime resources, is often justified by some reference to ***intuited*** public opinion (Fox & Freiberg, 1990). To make the most informed policy decisions a valid measure of crime seriousness must be employed to gauge ***actual*** public opinion. To this end, the present investigation set out to obtain some data on perceived crime seriousness, using a scaling technique that has not yet been tried in a New Zealand setting.

THE PRESENT STUDY

The present study had three research objectives. The first of these was to add to the meagre amount of data on crime seriousness obtained using the method of paired comparisons. The second objective was to explore two key factors which complicate the scaling of crime seriousness: sex differences and group differences. Group differences were examined by investigating perceived crime seriousness in two disparate occupational groups. The third objective was to examine the degree of consensus between the present study's sample evaluations of crime seriousness and New Zealand legislature and judiciary evaluations.

Objective 1 Little has been done in the way of scaling public perceptions of crime seriousness using the method of paired comparisons; this scaling technique has never before been applied to crime seriousness in New Zealand. In fact, only one such study of the public's perception of crime seriousness has been conducted in a New Zealand setting, irrespective of the scaling technique employed. Davis (1992) scaled public perceptions of crime seriousness using magnitude estimation. The first objective of the present study was to collect data on perception of crime seriousness using the method of paired comparisons.

Objective 2 There is sufficient evidence in the research literature to warn against the assumption that consensus is a social fact. Rather than dismiss the use of crime seriousness scales in New Zealand based on the perceptions of the community at large, as Spier et al. (1991*a*) asserted, the present study argues for the need for research that assesses the degree of consensus, or lack of consensus, in New Zealand community groups. Attempts should be made to explain consensus in terms of New Zealand respondent characteristics and broad classes of crime type. As such, the second objective was to investigate the effect of two respondent variables, sex and occupation.

It may be recalled that, in general, both sex and occupation have been found to influence the degree of consensus found between different groups of people in their estimations of crime seriousness and in the punishment philosophies they hold. The predominant finding obtained from previous research regarding sex as a factor in the degree of consensus found has yielded results suggesting that sex is an important variable. Female respondents have been found to perceive crimes against the person as more serious, and property crimes as less serious than male respondents. Females have also tended to prescribe harsher penalties and higher seriousness ratings for crimes against the person than males. Males, on the otherhand, have been found to be harsher in both their suggestions for penalties and in their ratings of seriousness for property orientated crimes when compared to females (Makela, 1966; Gebotys et al., 1988). Furthermore, rank-order differences in the seriousness of specific crimes, such as rape and murder, between males and females have been found (Lamp, 1982). Occupational differences between respondents on the effect of perceptions of crime severity have received very little attention. Samuel and Moulds (1986) found evidence to suggest that occupation effected the consensus found in recommendations of punishment for specific crimes. On the basis of these findings, it was expected that both the subject's sex and occupation would produce a difference in the ranking of crimes according to their seriousness. It was also expected that these differences would be crime specific.

Objective 3 In general, previous overseas research has found a discordance between legislative evaluations of crime seriousness and public perceptions of crime seriousness (e.g., Gibbons, 1969; Gescheider et al., 1982; Rose & Prell, 1955). In addition to this finding Davis (1992) showed that while there was less relative consensus to be found between New Zealand public perceptions of crime severity (as measured by rank order of crimes) and New Zealand legislative prescriptions (as measured by legislative maxima). than between the ranking of crime seriousness according to public opinion and the ranking of crime severity according to the New Zealand judiciary (as measured by court sentencing data). As such, another objective within the issue of consensus was to

investigate whether discordance existed between New Zealand legislative ranking of crime seriousness and New Zealand community evaluations of crime seriousness. The expectation, based upon previous research, was that discordance in the ranking of crime seriousness would exist between the New Zealand legislation and the subjects' perception of crime severity, but similar rankings would be obtained between subjects and the New Zealand judiciary.

EXPERIMENT 1
METHOD

SUBJECTS

Data for the first experiment were collected from two disparate occupational groups. One group was made up of New Zealand Army personnel (Linton Military Camp) and the other university students (Massey University). These groups were chosen as their occupations were thought to be completely different in content and style. Army personnel are restricted to a hierarchical and controlled environment. Students, on the other hand, work in an egalitarian and liberal environment. Thus, in general, while army personnel are taught not to question their superiors students are encouraged to question and are rewarded for doing so.

Massey University subjects

Forty Massey University psychology and business studies students volunteered to complete the study. Of these, 20 subjects were male and 20 female, the age range being 18 - 43 years (mean = 23.60, s.d = 7.37). Students were recruited by obtaining permission to seek subjects during their normal class times.

New Zealand Army subjects

Forty volunteer subjects from Linton Military Camp were also recruited. Again 20 of these subjects were male and 20 female, with an age range of 18 - 39 years (mean = 24.11, s.d. = 5.64). Three army male responses were later deleted, two for incorrect questionnaire responses and one for inconsistent responding. Army personnel were recruited by obtaining permission to contact officers in command of the medical, infantry

and transport units who spoke to and obtained names and phone numbers of those who wished to participate in the study. The army personnel recruited consisted of a random sample of both commissioned and non-commissioned ranks

THE QUESTIONNAIRE

Selection of crime types

Using pre-tests, Borg (1988) established that many students in his study had only vague notions of crimes such as libel, forgery and counterfeiting which composed Thurstone's original (1927*b*) crime stimuli. Therefore, for the present experiment only those crimes that Borg (1988) had found to be readily understood were used, with the exception of three crimes: seduction of minors, abortion and adultery. These crimes were not used as they were considered as 'sexual morals offences' (Levi & Jones, 1985), that is, crimes concerned with morality rather than criminality. For example, Borg found that his subjects did not perceive adultery as a crime. Seduction of minors was not used as it can be thought of as a 'swing' crime; it can be a crime or it can be a 'sexual morals offence' depending upon subjects' perceptions. For example, it no doubt would be perceived as a crime if it involved the sexual violation of a five year old. However, if it involved sex between two consenting 15 year olds, it is unlikely that it would be considered as anything more than a sexual morals offence.

Thus, the three crimes of seduction of minors, abortion and adultery were replaced with possession and use of cannabis, manslaughter and kidnapping. Manslaughter and kidnapping were thought to be two serious crimes and were used to balance the number of serious to non-serious crimes. The stimuli used in the present study (in alphabetical order) were arson, aggravated assault, burglary, kidnapping, manslaughter, murder, possession and use of cannabis, rape, receiving property dishonestly obtained and theft.

A majority of crimes in the list of offences that Borg (1988) employed were changed to conform with New Zealand legal statutes (that is, with those from the Crimes Act, 1961; and the Misuse of Drugs Act ,1975). For example, assault and battery in Borg's crime list was changed to aggravated assault (Crimes Act, 1961). Single word descriptors were used where possible.

Davis (1992) in his New Zealand study employed 25 crimes taken from the New Zealand 1961 Crimes Act, the 1962 Transport Act and the 1975 Misuse of Drugs Act. A pilot study carried out by Davis indicated the crimes he used were readily understood. In the present study, where every possible pair of crimes was presented to each subject, it was not considered necessary, nor desirable, to overtax the subjects using 25 crimes. The 10 crimes chosen spanned the range of crime seriousness (murder to possession and use of cannabis) and compared well in this respect with Davis' seriousness range (murder to illegal bookmaking). The present study selected 10 crimes, of which eight were also selected by Davis (manslaughter and arson being the two crimes excluded in the Davis study).

The 10 crimes in the present study were presented as 45 pairs of offences taken two at a time in all possible pairs (ignoring order of presentation within each pair). The sequence of pairs was randomly arranged with ordering within each pair also determined randomly, with the restriction that no one crime could follow in more than three consecutive pairs, (see Appendix A for a copy of the questionnaire).

Instructions

The instructions given to subjects were essentially the same as those of Thurstone (1927*b*) and Coombs (1967). The forced-choice format was described and subjects were asked to circle which of the two crimes they judged the more serious in a pair. Subjects

were also asked to be consistent in their decision making and an example of consistent answering was given (see Appendix A).

PROCEDURE

Arrangements were made to run subjects in small groups of between three and ten. Each subject was seated at a desk and handed a copy of the questionnaire and an informed consent form (see Appendix A). Subjects were told that the purpose of the study was to ascertain public opinion as to the seriousness of a set of 10 crimes. The instructions were then read out and any questions answered. Subjects were encouraged to respond reasonably quickly to each item pair; it was pointed out that a long drawn out thought process was unnecessary. Subjects were also informed that they could check over their answers and change any they thought incorrect or inconsistent with other answers. Each subject was given up to 20 minutes to answer the questionnaire. Subjects were then debriefed and thanked for their participation in the study.

MEASURES

In order to develop a scale of crime seriousness for the 10 crimes contained in the present study, Thurstone's law of comparative judgement (Case V) was employed. The method of paired comparisons has not yielded any sophisticated statistical measures. The comparison of data is predominantly visual as the method arranges each crime along a scale. Because the origin of the scale produced is arbitrarily set to zero, the resulting crime seriousness scale has the properties of an interval scale. Crimes can be rank-ordered and for any unit increase in the scale there is a unit increase in the seriousness of a crime. Distances between crimes can be meaningfully discussed, in terms of some distances, or separations, between crimes on the scale being greater/smaller than others. However, because the scales produced have no true zero point, it cannot be interpreted from the scales that subjects perceive crime x as twice as serious as crime y. Instead it

can only be noted, for example, that crimes against the person are perceived as relatively more serious than crimes against property. As such, absolute consensus, or the agreement in the seriousness scores obtained for each crime, could not be assessed. However, a measure of relative consensus could be obtained by comparing the rank-order of crimes as perceived by different subject groups. A measure specific to paired comparisons data termed *discriminability* could also be obtained to assess the overall agreement within each subject group.

Coombs (1967) defined discriminability as: *" The homogeneity of the subjects in their attitudes . The more uniformity or agreement between subjects, the greater the range of scale relative to the unit of measurement"* (p.70). Thus, discriminability is simply obtained by calculating the range of the scale values relative to the standard error of discriminal differences, which for Case V data is unity. Two extreme scale ranges exist, one that indicates maximum homogeneity and the other which indicates maximum heterogeneity between subjects in a group. For subjects to be maximally homogeneous each subject would agree with every other subjects as to which crime in every crime pair was most serious. Thus, the proportion matrix obtained would hold only proportions of 1.00 or 0.00. This would result in the maximum scale range possible between each crime pair relative to all others. Subsequently, the scale range between the lowest and highest scale value would be at its greatest, indicative of maximum homogeneity. Subjects would be maximally heterogeneous if they were divided 50/50 on the seriousness for each crime pair. This would result in the obtained proportion matrix containing only proportions of 0.50, the range between each crime and every other crime would therefore be zero. The resulting scale range would therefore also be zero indicating maximum heterogeneity. Thus, the more agreement between subjects the greater the range of the scale value relative to the unit of measurement.

In the initial Case V solution, the unit of measurement was the standard error of the discriminal differences and was assumed to be constant for all pairs of stimuli. As such,

Coombs (1967) stated that the range of the raw-scale values for each sample can be used as a rough index of the relative discriminability of the stimuli and can be used to compare between subject groups.

EXPERIMENT 1
RESULTS AND DISCUSSION

This results and discussion section is divided into four parts. The first part examines the internal consistency for each subject in their judgements of crime seriousness. The number of circular triads (Kendall, 1962) was used as a reliability measure to investigate how consistent each subject was when answering the questionnaire. The crime seriousness scales, constructed from the paired comparisons data, are next presented and discussed. Part three presents the data on the two independent variables, occupation and sex. The fourth section examines the degree of consensus between the crime seriousness rankings obtained in the present study and the rankings according to New Zealand legislation (measured in statute maxima) and judiciary (measured in custodial sentence length).

CIRCULAR TRIADS

Internal consistency is important in the method of paired comparisons as it indicates whether the basic assumptions for the law of comparative judgement (Case V) have been met (Guilford, 1954). A high level of internal consistency suggests that standard deviations are equal and discriminal dispersions are normal. A high level of internal consistency also suggests that subjects are comparing stimuli on the single dimension of seriousness (Walker, 1978).

In order for there to be a high degree of internal consistency, subjects are required to be transitive in their decision making for most of the crime pairs presented to them. If a subject decides that the crime of murder is more serious than the crime of theft, and the crime of theft is more serious than the crime of aggravated assault, the subject should

report that murder is more serious than aggravated assault. Each time a subject disobeys this transitivity rule for any three crimes, a circular triad is said to exist.

The number of circular triads for each subject in the present study was found using a special purpose computer programme that incorporated Kendall's (1962, pp. 146-148) equations for computing both the actual and expected number of circular triads.The number of circular triads per person were then transformed into coefficients of consistence using Kendall's (1962) method. These equations are, in effect, ratios of the observed to total circular triads possible. Thus, if a subject obtained a coefficient of consistence of 0.98, he/she would be 98% consistent as to which crime was seen as more serious than the other in all the pairs presented. The coefficients of consistency derived from the present data were categorised at 0.05 intervals.

Coefficients of consistency per person were employed, instead of the absolute number of triads per person, as this procedure standardised data and allowed for comparison of the present sample's consistency with previous research. Unfortunately, direct comparisons still could only be made with Coombs' (1967) study as other studies did not compute internal consistency and did not provide sufficient information to do so (e.g., Carroll et al., 1974; Borg, 1988; Krus et al., 1977). Approximations to Coombs' absolute frequencies of circular triads per subject were transformed into coefficients of consistence and also categorised at 0.05 intervals. Coefficients of consistence for both the present study and Coomb's study are presented in Table 1.

Table 1 indicates that the present study's coefficients of consistence ranged from 100% consistency to 75% consistency, with nearly all subjects obtaining 95 to 100% consistency. Coombs' (1967) study had coefficients of consistency that ranged from 100% to 10% consistency. However, only half of Coombs' subjects obtained 95% to 100% consistency, compared to 94.9% of the subjects in the present study. A possible explanation for Coombs' sample being less consistent than the present sample is that the

current study's instructions specifically directed subjects to be consistent in their answering, and gave an example of the transitivity required (see Appendix A) whereas Coombs' instruction did not. It seems likely that if subjects are alerted to the possibility of inconsistent answering they are able to avoid it. This could be an important finding for future research that utilises paired comparisons. On the other hand, the high consistency obtained for the present study may be no more than the result of the small number of crimes utilised in the present study, 10, as opposed to the 20 crimes employed by Coombs. It is reasonable to expect consistency to decrease as the number of crimes presented to subjects increases.

Table 1.
Coefficients of consistence, categorised at 0.05 intervals. Equivalent data, derived from Coombs (1967), are included for comparison. The data obtained from Coombs do not sum to 100 as they were taken from a graph, and as such, only approximate frequencies could be obtained.

COEFFICIENTS OF CONSISTENCE	% OF SUBJECTS	
	Present study	Coomb's (1967) study
0.00 - 0.05	0.00	0.00
0.06 - 0.10	0.00	0.00
0.11 - 0.15	0.00	0.06
0.16 - 0.20	0.00	0.00
0.21 - 0.25	0.00	0.00
0.26 - 0.30	0.00	0.06
0.31 - 0.35	0.00	0.00
0.36 - 0.40	0.00	0.00
0.41 - 0.45	0.00	0.06
0.46 - 0.50	0.00	0.06
0.51 - 0.55	0.00	0.00
0.56 - 0.60	0.00	0.00
0.61 - 0.65	0.00	0.00
0.66 - 0.70	0.00	0.06
0.71 - 0.75	1.30	1.20
0.76 - 0.80	0.00	2.70
0.81 - 0.85	0.00	5.60
0.86 - 0.90	1.30	17.10
0.91 - 0.95	2.50	21.40
0.96 - 1.00	94.90	49.00

Defining a cut-off point at which a subject's internal consistency is too low can be established by comparing the obtained frequency distribution of circular triads with the distribution of circular triads expected by chance (Coombs, 1967). Any subjects who are found to have a number of circular triads that falls, say, within three standard deviations

of the number expected by chance, could be deleted from the analysis. The actual cut off point is somewhat arbitrary. For the present study, the expected number of circular triads for a subject responding at random is $(^1/_4)\binom{n}{3} = 30$ (where n is the number of crimes) with a variance of $(^3/_{16})\binom{n}{3} = 22.5$ (Kendall, 1962).

Coombs (1967) found that only one of his subjects fell within three standard deviations of the expected number of triads in his study. Further analysis of Coombs' data (which he had not undertaken) shows that this subject was only 10% consistent in responding. There were also six further subjects who were less than 50% consistent in their answering. These margins of error were considered too great for the present study, especially if the smaller number of crimes used is taken into consideration. Hence, instead of using information from a graph to define an inconsistency cut-off point , such as Coombs did, an arbitrary minimum of 80% consistence rate was required. This resulted in one subject being deleted from the data analysis.

It is interesting to note that when this (arbitrary) cut-off point is applied to both Coomb's (1967) study and the present study, 93.1% of his sample were consistent more than 80% of the time and 98.7% of the present study's sample were consistent more than 80% of the time. Thus, it can be stated with some confidence that nearly everyone who used the paired comparisons method as a means of rating crime seriousness in both studies had little difficulty using and understanding the method. This replicates the previous finding of Walker (1978) who found that very few subjects had difficulty in deciding which of two crimes was perceived as more serious in paired comparisons ratings.

SCALE CONSTRUCTION

Proportion matrices. Responses to each pair of crimes were totalled across all subjects to produce a frequency matrix (the number of times each crime was seen to be the more serious of a pair). These relative frequencies were then converted to proportions and then

to z scores using normal curve tables. The data, in proportion form, are shown in Table 2, and in z-score form in Table 3.

Table 2.
The proportion matrix for all subjects collapsed across sex and occupation. The row offence is more serious than the column offence. It was assumed that for a crime presented with itself a proportion of 0.50 would be obtained (Guilford, 1954). These proportions have been omitted from the tabulated data.

	1	2	3	4	5	6	7	8	9	10
1.Cannabis		0.792	0.922	0.961	0.961	1.000	1.000	1.000	1.000	1.000
2.Receiving	0.208		0.896	0.987	0.974	0.961	1.000	0.961	1.000	1.000
3.Theft	0.078	0.104		0.935	0.961	0.961	0.987	0.948	1.000	1.000
4.Burglary	0.039	0.013	0.065		0.896	0.844	0.987	0.987	1.000	1.000
5.Arson	0.039	0.026	0.039	0.104		0.663	0.663	0.883	0.974	1.000
6.Ag. Assault	0.000	0.039	0.039	0.156	0.337		0.519	0.870	0.948	1.000
7.Kidnapping	0.000	0.000	0.013	0.013	0.337	0.481		0.779	0.844	1.000
8.Manslaughter	0.000	0.039	0.052	0.013	0.117	0.130	0.221		0.519	1.000
9.Rape	0.000	0.000	0.000	0.000	0.026	0.052	0.156	0.481		0.909
10.Murder	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.091	

In all, nine such proportion matrices were constructed. These matrices, for males and females within the total sample, as well as for students and army personnel and males and females within these two occupational groups, can be found in Appendix B.

It is important to note that numerous boundary proportions were found in all the matrices created. For example, in Table 2, the crime of murder was chosen by every subject as being more serious than aggravated assault. Even though these results demonstrated complete agreement among subjects, such boundary proportions cannot be scaled on a finite scale because the equivalent z values for these scores are indeterminate. It is remarkable that previous research that has employed the method of paired comparisons, such as Carroll et al. (1974), Coombs (1967), Krus et al. (1977) or Thurstone (1927*b*), appear to observe no such boundary problems. From Coombs' proportion matrix, for example, one can conclude that 2% of his subjects must have evaluated murder as less serious than receiving stolen goods. As Borg (1988) asserts, this can only be interpreted as error of judgement. Thus, the true dominance proportion of murder paired with receiving stolen goods should have been 1.00 instead of 0.98. One has to wonder about the validity of scaling data where the scale value may be dependent upon subject error.

Errors of judgement may have been present in the above mentioned studies and absent in the present study because of the relatively larger sample sizes that previous studies used. As sample size increases so to does the chance of obtaining error in the judgement of seriousness for crime pairs presented to subjects, or, for that matter, the odd subject who does genuinely hold an unconventional view of how relatively serious different crime are. Another possible cause of error in previous research may have been the inclusion of subjects in the data analysis that had been found to have low internal consistency. For example, Coombs (1967) included subjects in his data analysis that had a coefficient of consistence of only 0.28. As the present study deleted subjects with less than a 0.80 coefficient of consistence, there was little room for error to be brought into the data analysis through such low internal consistency.

z matrices. To construct scales of crime seriousness, the law of paired comparisons (Case V) requires that the proportion matrices be standardised by transforming them into *z* matrices. Such a matrix can be observed in Table 3 which presents a matrix of *z* scores for each crime pair. Each cell in Table 3 represents the *z* score for the earlier found proportion of people who rated one crime as more serious than the other (see Table 2).

It can be seen from Table 3 that when the p_{ij} is greater than 0.50, z_{ij} receives a positive algebraic sign, while for a proportion less than 0.50, z_{ij} receives a negative sign. For example, the proportion of 0.844 of people who rated rape as more serious than kidnapping in Table 2 became a standardised *z* score of 1.01. Table 3 also shows that there were missing data for a number of crime pairs. Blank cells in the *z* matrices indicate that for these crime pairs boundary proportions were found. *z* matrices were completed for all nine proportion matrices (see Appendix B).

Table 3.
z Matrix for all subjects, collapsed across sex and occupation.

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.00	0.81	1.42	1.76	1.76					
2.Receiving	-0.81	0.00	1.26	2.23	1.94	1.76		1.76		
3.Theft	-1.42	-1.26	0.00	1.51	1.76	1.76	2.23	1.63		
4.Burglary	-1.76	-2.23	-1.51	0.00	1.26	1.01	2.23	2.23		
5.Arson	-1.76	-1.94	-1.76	-1.26	0.00	0.42	0.42	1.19	1.94	
6.Ag. Assault		-1.76	-1.76	-1.01	-0.42	0.00	0.05	1.13	1.63	
7.Kidnapping			-2.23	-2.23	-0.42	-0.05	0.00	0.77	1.01	
8.Manslaughter		-1.76	-1.63	-2.23	-1.19	-1.13	-0.77	0.00	0.05	
9.Rape					-1.94	-1.63	-1.01	-0.05	0.00	1.34
10.Murder									-1.34	0.00

Note. To facilitate table spacing, z scores are two decimal places only.

Deriving the crime seriousness scale. Thurstone's (1927*a*) law of comparative judgement (Case V) was used to construct the scale of crime seriousness. However, because obtained data contained numerous boundary proportions, which resulted in missing data in the z matrices, Thurstone's original scaling procedure could not be used. A scaling procedure suggested by Edwards (1957) for Case V missing data was therefore employed. Edwards' scaling procedure involves calculating a matrix of successive differences of the column (C) entries of the z -score matrix. For example, the mean of the sum of z_{ij} (C2) - z_{ij} (C1) was used to create a scale value instead of using the mean of the sum of z_{ij} for each crime, as in Thurstone's original method. A matrix of successive differences for the total sample is presented in Table 4. The first column, C2 - C1, represents the z scores found for receiving stolen property minus the z scores found for possession and use of cannabis for every crime. Thus, for example, in Table 4, the entry 0.160 for theft in column C2 - C1 is obtained by subtracting the z score found for possession and use of cannabis paired with theft (C1 = -1.419) from the z score found for receiving stolen property paired with theft (C2 = -1.259).

A matrix was thus obtained that contained the difference of the z values of adjacent columns of the z matrix. In all, nine such matrices were created (see Appendix B). To derive the scale value for each crime, the mean of each column in this new matrix was calculated. This was achieved by summing the cells in each column and dividing by the number of times a successive difference was found (see Table 4). For instance, it can be

seen in Table 4 that the mean for C2 - C1 is 0.228. To scale the data, S_1 (the least serious crime) is arbitrarily set to zero. As can be seen from Table 4, this is the crime of possession and use of cannabis.

Table 4.
Matrix of successive differences for the total sample collapsed across occupation and sex.

STATEMENTS	COLUMN DIFFERENCES								
	C2-C1	C3-C2	C4-C3	C5-C4	C6-C5	C7-C6	C8-C7	C9-C8	C10-C9
1. Cannabis	0.813	0.606	0.344						
2. Receiving	0.813	1.259	0.967	-0.283	-0.181				
3. Theft	0.160	1.259	1.514	0.248		0.464	-0.601		
4. Burglary	-0.464	0.712	1.514	1.259	-0.248	1.216			
5. Arson	-0.181	0.181	0.505	1.259	0.420	-0.003	0.773	0.753	
6. Ag. Assault			0.752	0.591	0.420	0.048	1.079	0.500	
7. Kidnapping				1.809	0.370	0.048	0.769	0.242	
8. Manslaughter		0.137	-0.601	1.036	0.064	0.358	0.769	0.048	
9. Rape					0.317	0.615	0.963	0.048	1.335
10.Murder									1.335
SUM	1.141	4.154	4.995	5.920	1.162	2.746	3.752	1.591	2.670
N	5	6	7	7	7	7	6	5	2
MEAN	0.228	0.692	0.714	0.846	0.166	0.392	0.625	0.382	1.335
SCALE VALUES									
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
0.000	0.228	0.920	1.634	2.480	2.646	3.038	3.663	4.045	5.380

To solve for the other scale values is simply a matter of cumulating the mean values for each of the successive difference columns. So, for example, the solutions for the next three scale values in Table 4 are:

S_1 Cannabis = 0.000
S_2 Receiving Stolen Property = 0.000 + 0.228 = 0.228
S_3 Theft = 0.228 + 0.692 = 0.920
S_4 Burglary = 0.920 + 0.714 = 1.634.

The complete set of scale values is shown in the last row of Table 4.

The difference between Thurstone's (1927*a*) original scaling method and Edwards' (1957) Case V incomplete data method is that the latter method gives scale separations of adjacent stimuli whereas the former method gives scale values expressed in terms of their deviations from the mean of all the scale values. However, it can be noted that the two methods yield identical relative scale values for stimuli.

As well as the scale for all the subjects (collapsed across sex and occupation) shown in Table 4, separate scales were obtained for each occupational group and for each sex using Edwards' (1957) procedure. All scale values are shown in Table 5 and the actual scales in Figure 3.

Table 5.
Crime seriousness scale values for all subjects, and for the data broken down by gender and occupation. Scale values for the crimes of arson, aggravated assault, kidnapping, manslaughter, rape and murder could not be obtained for female students. This was because the number of column differences the scale value for burglary was to be based upon were non-existent. A scale value for the crime of murder for male army personnel could also not be obtained for similar reasons.

OFFENCE	TOTAL SAMPLE			STUDENT SAMPLE			ARMY SAMPLE		
	Total	Female	Male	Total	Female	Male	Total	Female	Male
1.Murder	5.380	6.135	6.134	6.358		6.192	5.718	5.621	
2.Rape	4.045	5.099	4.207	5.076		4.547	4.319	4.585	4.230
3.Manslaughter	3.663	4.708	4.077	4.798		4.273	3.878	3.998	4.175
4.Kidnapping	3.038	3.445	3.357	3.831		3.536	2.997	3.282	3.594
5.Ag. Assault	2.646	3.781	2.509	3.355		2.694	2.816	3.573	2.811
6.Arson	2.480	3.090	2.732	3.113		3.039	2.474	2.893	2.707
7.Burglary	1.634	2.198	1.929	2.145		1.919	1.718	2.386	1.779
8.Theft	0.920	1.292	1.289	1.490	1.414	1.433	1.075	1.544	1.247
9.Receiving	0.228	0.637	0.183	0.600	0.707	0.682	0.261	0.570	0.931
10.Cannabis	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

It can be seen from Table 5 that only seven complete scales could be produced. Complete scales for female student subjects and male army personnel could not be obtained. As seen in the successive difference matrix for female students in Appendix B, no column differences for the crime of burglary (C4 - C3) existed within the *z* matrix, resulting in the N and mean scale difference value for this crime equating to zero. This lead to an incomplete scale (see the *z* matrix for the female student sample in Appendix B). For similar reasons a scale for male army personnel could not be constructed.

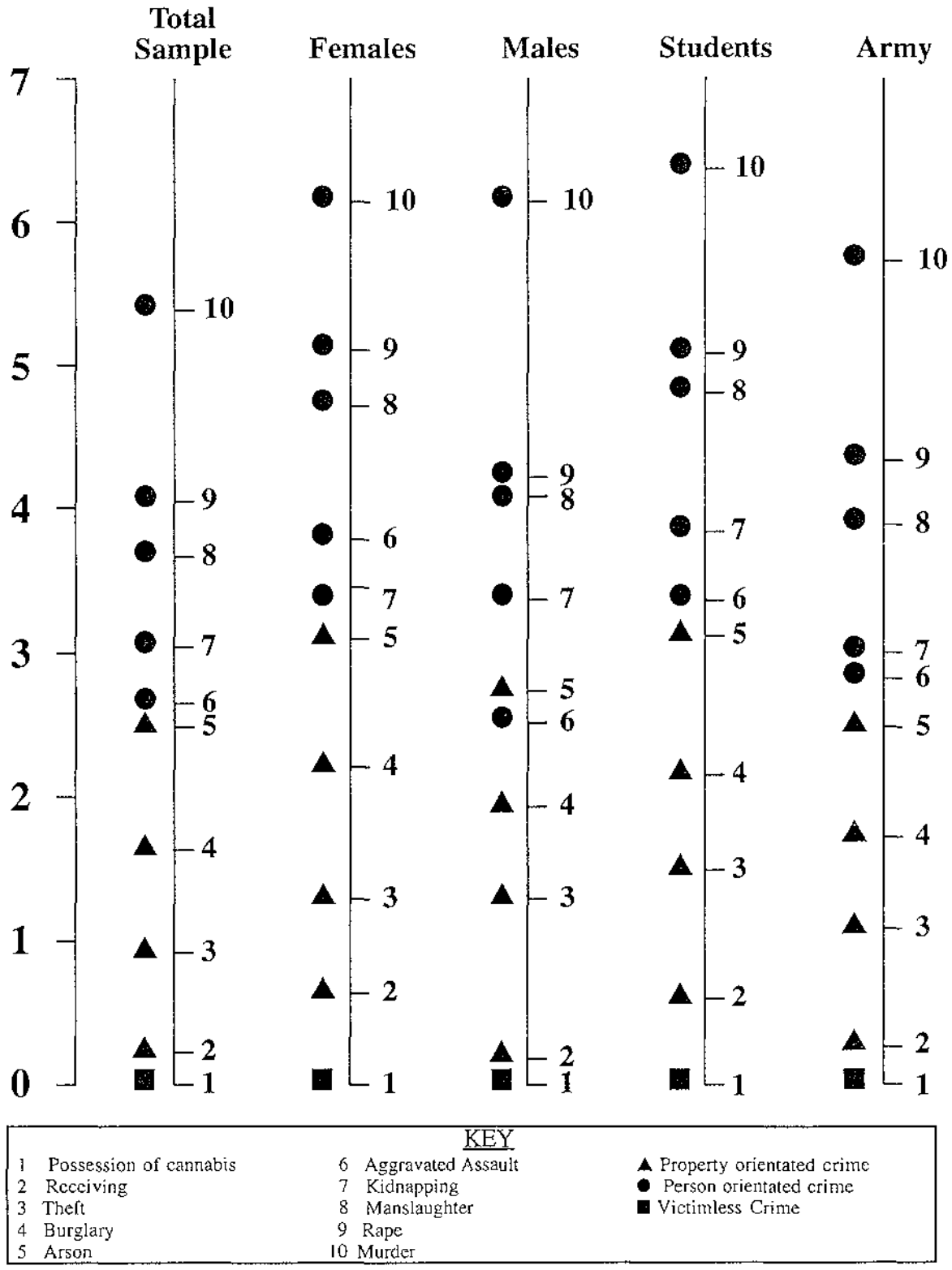

Figure 3 . Interval scales for crime seriousness for the total sample and for occupation and gender. In almost all cases, every crime against the person was perceived as relatively more serious than property crime. Also rank orderings differed between females and males for the crimes of arson, aggravated assault and kidnapping. No rank order differences were found between the occupational groups.

INTERPRETING THE SCALED DATA

Classification of offences

Figure 3 demonstrates that Thurstone's (1927*b*) original classification of crimes into those against the person and those against property held. For all groups, crimes against the person, especially murder and rape, were placed highest on the scales of crime seriousness. Crimes against property, that did not involve acts against the person, were placed relatively lower down the scales. This finding supports numerous other studies which have also found that person orientated crimes are perceived as more serious than property oriented crimes (e.g., Coombs, 1967; Krus et al., 1977; Rossi et al., 1974). From Figure 3 it can be clearly seen that all groups viewed the victimless crime of possession and use of cannabis as the least serious crime, a result similar to that of Chilton and DeAmicis (1975) and Rossi et al. (1974).

Consensus

The usefulness of the ratings of crime seriousness obtained in the present study is, to some extent, contingent upon the degree to which different groups were in consensus. If major differences were to be found between the sexes and across occupational groups, then the agreement that the public perception of seriousness should be taken into account in say, setting sentence length, is weakened. For this reason, female perceptions of crime seriousness were compared with those of males and student perceptions with those of army personnel.

Effects of sex and occupation As can be seen in Figure 3, all groups agreed with the total sample for the ranked seriousness of specific crimes. Murder was always perceived as the most serious crime, followed by rape and manslaughter. Groups also agreed that possession and use of cannabis was the least serious crime and that receiving stolen

goods, theft and burglary followed in ascending order of seriousness. But differences between groups emerged for the crimes of kidnapping, aggravated assault and arson. Thus, it seems that relative consensus exists for the most serious crimes and the least serious but not for those crimes that are perceived as moderately serious. This result is very similar to that of Cullen et al. (1985) and Miethe (1984). These researchers also found that there was considerable agreement between various groups of people as to which crimes were perceived as the most serious and least serious, but little agreement between people for crimes that were considered as moderately serious.

Occupational differences As Figure 3 demonstrates, complete relative consensus for all crimes presented was found between the two occupational groups. All 10 crimes were ranked identically when student and army personnel perceptions of crime seriousness were compared. This result supports previous studies that have shown that occupational differences in relative consensus of crime seriousness are minimal (Lamp, 1982). Indeed, it has been demonstrated that such diverse groups as university students and police produce similar rank ordering for crime seriousness (Sellin and Wolfgang, 1964).

However, a *direct* comparison of the crime pairs, rather than the *indirect* comparison obtained by comparing the scaling of the ten offences, did reveal several occupational differences in the perception of crime seriousness. Noticeable proportional differences in the percentage of students versus army personnel who viewed one crime as more serious than the other existed for specific crime pairs. For instance, more students (82.5%) perceived manslaughter as more serious than kidnapping compared to army personnel (73%). Several other such differences existed, predominantly between crime pairs that involved manslaughter, aggravated assault and arson. Just why these variations exist, and why they exist for some crime pairs and not others is not known. While they may be due to the unreliability of the data, it also could be that students and army personnel varied in their discrimination of crime seriousness between certain pairs because of their occupational differences.

Because the paired comparisons scales have an arbitrary zero point, absolute scale scores could not be compared. However, when relative seriousness for each crime was compared on the basis of occupation, by comparing relative positions on these crime seriousness scales, differences between the two occupations were found. In general, students perceived all crimes as relatively more serious than army personnel. This was seen to be especially true for crimes against the person.

As can be seen in Figure 3, the scales obtained for the two occupations have noticeably different lengths, with the student scale relatively longer than the army scale. When discriminability over all 10 crimes was measured for army personnel and students the range was 5.718 and 6.358 respectively, relative to the standard error of discriminal differences, which was set to one. The difference in range of 0.640 between these two scales suggests that students were more homogeneous in their perceptions of crime seriousness than army personnel.

It is not clear exactly why students were found to be more homogeneous than army personnel. As both army personnel and students had equivalent mean ages and age distributions, age was not thought to be a factor that contributed towards the lower homogeneity found within the army sample. However, previous research suggests that educational attainment may be a possible explanation. Kutschinsky (1970, cited in Newman & Trilling, 1975) found that leniency towards criminal behaviour decreased as educational attainment increased. While students were of the same educational attainment, army personnel may have had marked discrepancies in educational attainment which possibly ranged from a tertiary qualification to school certificate. This discrepancy in educational attainment amongst army personnel may have led to there being more variance in perceptions of crime seriousness for this sample than for the student sample.

Sex differences Relative consensus was not as strong between the sexes as it was between the two occupational groups. However, these rank order differences constituted

only minor variations in the perception of crime severity between males and females as both groups produced rank orders of crime severity that were highly correlated (r=.964). The differences that did occur between these two groups were found to exist for the three crimes of kidnapping, aggravated assault and arson. As seen in Figure 3, females perceived aggravated assault as the most serious crime out of the three, followed by kidnapping and then arson. Males, on the other hand, perceived the crime of kidnapping as the most serious out of the three crimes followed by arson and aggravated assault. These results are similar to those found by Lamp (1982) who also observed rank order differences in crime seriousness between his male and female sample. Lamp explained these rank order differences in terms of personal threat. In general, Lamp found that those crimes perceived as more personally threatening to females than to males were also perceived as more serious by females than by males. Thus, differences in rankings of kidnapping, aggravated assault and arson according to their seriousness between males and females in the present study may reflect the difference in the degree of personal threat that each of these crimes represents for females and males. Of course, this "explanation" begs the question of why females/males should view kidnapping as more or less personally threatening.

Differences in relative seriousness were also found when the crime positions on the male and female crime seriousness scales were compared. As seen in Figure 3, with the exception of murder, kidnapping and theft, females perceived all crimes as relatively more serious than males.

Despite the differences found between males and females regarding the rank-order and the relative seriousness of certain crimes, no difference was found between the sexes regarding homogeneity. As can be seen in figure 3, females had a scale range of 6.135 and males had a similar scale range of 6.134. These results differ from those obtained in previous research. Both Thurstone(1927*b*)and Coombs (1967) found that males were slightly more homogenous than females in their judgements of crime seriousness. For

instance, Coombs found that the males in his sample were more homogeneous than females by a scale range difference of 0.260 units. A possible explanation for the variation in results found between the present study, and the results found by Thurstone and Coombs, may be the elimination of the 'sexual morals offences' of seduction, adultery and abortion in the present study. These offences may have increased the level of heterogeneity within the female sample in Thurstone's and Coombs' study.

Apart from the differences in perceived seriousness between males and females in the ranking of crimes according to their seriousness, variations were also found between the sexes when crime pairs were directly compared with each other. Of interest were the noticeable differences in proportions found between males and females for crime pairs that included the crime of rape. In general, rape was seen as more serious by females than males for all other crimes, excluding the least serious crimes of possession of cannabis, receiving stolen property, theft and burglary. For instance, all of the female subjects, 100%, regarded rape as more serious than arson, while 94.6% of males thought this to be the case. A larger percentage of females (97.5%) than males (91.9%) also perceived rape as more serious than aggravated assault. Also, 55% of females viewed rape as more serious than manslaughter compared to 51.4% of males. More females than males also perceived rape as more serious than murder. While 15% of females perceived this, only 2.7% of males had the same opinion. These results suggest that a larger percentage of females perceived rape to be more serious than arson, aggravated assault, kidnapping, manslaughter, and even murder, than males.

These findings are similar to those of Lamp (1982) who also found found that females perceived rape as more serious than males in relation to other crimes. A possible explanation for these findings is again the perceived threat of the crime itself. Lamp found that females perceived rape as the most personally threatening crime. It seems that personal threat translates into an increase in perceived seriousness of a crime.

In summary, a high degree of relative consensus was found to exist between students and army personnel in their perceptions of crime severity. Between males and females this consensus existed for the most serious and least serious crimes but not for those crimes middling in seriousness on the crime severity scales, namely: arson, aggravated assault and kidnapping. Discriminal differences between the scales obtained for students and army personnel indicated that students were the more homogeneous of the two occupations in their perceptions of crime severity. When discriminability was measured for males and females, males and females were found to be equally homogeneous in their perception of crime seriousness.

COMPARISONS WITH THE NEW ZEALAND JUDICIARY AND LEGISLATURE

It is of some interest to investigate the degree of relative consensus among the rank ordering of crimes found in the present study, the New Zealand Judiciary, and the New Zealand Legislature. There are a number of ways this could be done. However, to make the present results as comparable as possible with other similar studies, it was decided to define crime seriousness with respect to maximum statutory sentence length for the New Zealand legislature and with respect to average sentence length imposed (for 1990) for the judiciary. Female and male rankings for the present study were separated from the total sample as different rank orders were obtained for these samples compared to the total sample. However, as both occupational samples obtained the same rankings as the total sample the rankings for these groups were not separated. The rankings of crime seriousness are shown in Table 6.

Table 6.
Ranking of crime seriousness according to rankings obtained for the present study, the New Zealand Judiciary and the New Zealand Legislature.

CRIME	PRESENT STUDY (1)			JUDICIARY (2)	LEGISLATION (3)
	Total	Females	Males		
1 Murder	1	1	1	1	1
2 Rape	2	2	2	2	2
3 Manslaughter	3	3	3	3	2
4 Kidnapping	4	5	4	4	2
5 Aggravated Assault	5	4	6	6	9
6 Arson	6	6	5	5	2
7 Burglary	7	7	7	7	6
8 Theft	8	8	8	9	7
9 Receiving Stolen Property	9	9	9	8	7
10Possession and use of Cannabis	10	10	10	10	10

(1) Rankings of crime serious obtained for the present study.

(2) Ranking of each crime according to the average sentence length imposed for each crime by the New Zealand courts in 1990 (Spier, Southly & Norris, 1991*b*). Davis (1992) found that both the seriousness scale created by Spier et al. (1991*a*) and data from this report produced identical ranking of crime seriousness. As such, the average custodial sentences for year ending 1990 were selected for the 10 crimes in the present study.

(3) Ranking of each crime according to maximum statutory sentence length as found in the Crimes Act (1961) and the Misuse of Drugs Act (1975).

The findings obtained from these rankings indicated that relative consensus between the present study the judiciary and the legislature existed for the crimes ranked as most serious, these being murder, rape and manslaughter. Relative consensus also existed for the crime of possession and use of cannabis, ranked least serious. However, minor differences existed between the rank ordering of crime seriousness in the present study when compared with both the judiciary (r =.976) and legislature (r = .740). However, it can be noted that the differences between the rank order for the current results and those of the judiciary were no more different than the rank orderings between the judiciary and the legislature (r = .825).

Table 6 shows that relative consensus between the present study and the judiciary was greater than for the present study and legislature. This finding agrees with that of Davis (1992). There was a high degree of relative consensus between the present study's data and the judiciary (r = .976). Relative consensus was again shown to be crime specific

with the least relative consensus found for crimes perceived as moderately serious. The female sample showed different rank ordering for the crimes of aggravated assault, arson and kidnapping compared to judiciary rankings. Females perceived aggravated assault as the most serious crime out of the three, followed by kidnapping and arson; the judiciary perceived kidnapping as most serious followed by arson and aggravated assault. No differences were found between male rankings of these crimes and that of the judiciary. All groups in the present study judged theft as more serious than receiving of property dishonestly obtained. This perception in seriousness for these two crimes was reversed in the judiciary's rank order of these crimes.

Differences in rank ordering of crimes according to their seriousness between the New Zealand judiciary and legislature were evident from the present analysis, a finding which was noted by Davis (1992) and Fox and Freiberg (1990). While the legislature ranked manslaughter, kidnapping and arson as equal in seriousness, the judiciary ranked these crimes as third, fourth and fifth respectively. Aggravated assault was seen as relatively more serious than burglary by the judiciary, the rank order of which was reversed in legislation.

In summary, results indicated that the ranking of crime seriousness according to the legislature and the judiciary were similar to the rank ordering of crimes obtained for the present study. This was especially so for crimes perceived as high or low in seriousness. For instance, all three groups perceived murder as the most serious crime and possession and use of cannabis as least serious crime. However, deviations existed between these three groups in the rank ordering of crimes categorised as moderately serious. This was especially true when the present study's ranking of crime seriousness was compared to that of the legislature. In general, the results of the comparison between the present data, the judiciary and legislature show few signs of the cultural lag that Geshcheider et al. (1982) and Rose and Prell (1955) found.

THE PRESENT DATA COMPARED TO DAVIS (1992)

Very little New Zealand research has been conducted in the area of public perceptions of crime severity. Only one other study has been carried out in New Zealand, the work of Davis (1992). As such, his study constitutes an important piece of research. Davis employed the magnitude estimation technique to scale crime severity according to public opinion. These data were then compared with those of the New Zealand police, judiciary and legislature. In his study Davis had 115 respondents, randomly selected from the general public in Christchurch, complete a survey. The survey involved respondents estimating the seriousness of 25 crimes that ranged from murder to book making. The crimes were presented to subjects in both single-word and one to two sentence descriptions. The questionnaire contained a training exercise as well as questions relating to the demographic variables of age, sex, occupation and whether or not respondents had been a victim of a crime in the last 12 months. Also contained in the questionnaire was a Criminal Attitude Scale (CAS) designed to measure the relationship between respondent attitudes towards specific crimes and their attitudes towards offenders.

The main findings that Davis (1992) obtained are as follows. First, his results indicated that subjects gave the same crime seriousness estimates to crimes, no matter whether they were presented as a single word description or a one sentence vignette. However, as Davis himself states, this finding may have been the result of both forms of the crime being given to the same subject. The subjects may have realised that they had previously estimated the seriousness for a crime given as a single word description earlier and therefore given the crime vignette the same crime seriousness estimate. The second major finding was that high levels of both relative and absolute consensus existed amongst the respondents in his study. The demographic variables of sex, age, and socioeconomic status or whether subjects had been victims of a crime within the last 12 months appeared to have no effect on crime seriousness estimates, except for a few specific crimes. For example, brothel keeping and defacing a coin were seen as more serious by females than

males. However, in general, widespread agreement was found in the perceptions of crime seriousness by the subjects in his study. The third finding related to the comparison of ratings of crime seriousness by his subjects with those of the New Zealand police (measured by clearance priority), judiciary (as measured by average sentence length) and legislature (as measured by statutory maxima). It was found that subjects agreed most with the New Zealand judiciary, followed by moderate agreement with the police and least agreement with the legislature. Finally, it was shown that there was no relationship between his subjects' CAS scores and their seriousness ratings of the 25 crimes.

There are several similarities between the results of Davis (1992) and the present study. First, both studies found high degrees of relative consensus between different groups of subjects. Specifically, the present study found high levels of relative consensus between the sexes, a finding which supports that of Davis. Similar findings were also obtained when perceptions of crime seriousness obtained in the two studies were compared with that of the New Zealand legislation and and judiciary. Both studies found that crime seriousness rankings correlated highly with that of the New Zealand judiciary, but that this correlation was lower when compared to that of the New Zealand legislation.

Furthermore, a comparison of the rank order for the crimes common to Davis' study and the current study showed them to be nearly identical (see Table 7). This finding is consistent with that of Walker (1978), who also found that the different scaling techniques of magnitude estimation and paired comparisons produced similar rankings of crime seriousness.

Table 7.
Ranking of crime seriousness obtained for the crimes in the present study compared to those of Davis (1992) for the same crimes. Only eight crimes were compared as Davis' selection of crimes did not include manslaughter and arson, as did the present study.

CRIME	PRESENT STUDY			DAVIS (1992)
	Total	Female	Male	Total
1 Murder	1	1	1	1
2 Rape	2	2	2	2
3 Kidnapping	3	3	3	3
4 Aggravated Assault	4	5	4	4
5 Burglary	5	4	5	5
6 Theft	6	6	6	7
7 Receiving Stolen Property	7	7	7	6
8 Possession and Use of Cannabis	8	8	8	8

Table 7 shows that nearly identical rankings were obtained for both the total sample and the male sample, the rank order correlations for these groups both being r = .989. However, rank order differences did exist when the females in the present study were compared with Davis' total sample. The one clear difference, shown in Table 7, was that all groups in the present study perceived theft as more serious than receiving property dishonestly obtained. However, Davis' (1992) total sample perceived receiving property dishonestly obtained as more serious than theft. One could speculate that this difference is due to the different subject samples used in the two studies. Davis used a sample drawn from the population of Christchurch, whereas the present sample consisted of miliary personnel and students whose age ranges were different to that of Davis' subjects. While the present study contained only subjects whose ages ranged from 18 - 43 years, Davis' study also contained a number of mature adults aged 56 years and over.

In summary, after comparing the results obtained in the present study with those of Davis' (1992) study similar conclusions regarding relative consensus between differing groups of people were found. Both studies found that a high level of consensus existed not only between the sexes but also between subjects surveyed and the different levels of the justice system. A high level of relative consensus also existed between the two studies when the rank ordering of the crimes were compared. These results suggest that

the different scaling techniques used produce similar results concerning the issue of crime seriousness.

EXPERIMENT 1: CONCLUSIONS

The results of the present study, using a New Zealand sample, compare well with previous studies using the method of paired comparisons. Moreover, the results are also in agreement with those obtained using the technique of magnitude estimation for scaling crime seriousness. These results are not surprising since previous research has shown that consistent results are obtainable (e.g., Walker, 1978). Thus, it does not seem to matter, at least for the purpose of rank ordering crimes in terms of their seriousness, whether magnitude estimation or paired comparisons is used.

A principle part of the present study was to investigate whether the scaling of crime seriousness varied across sex and across different subject groups. If this were the case, then then the argument put forward by Spier et al. (1991*a*) that public perception of crime seriousness cannot be meaningfully used to help determine sentence lengths would have some credence. The finding that there were few major differences between students (viewed as a well educated group that works in a liberal thinking environment) and mostly lower rank army personnel (viewed as working in a rigid hierarchical structure and having less education) mitigates against this view. Also in regard to sex differences, some were found but these were not especially large differences. The differences in relative consensus between males and females concerned the crimes of aggravated assault, kidnapping and arson which were viewed as moderately serious. Females ranked aggravated assault as the most serious of the three crimes, followed by kidnapping and arson. Males, on the other hand, ranked kidnapping as the most serious crime of the three, followed by arson and then aggravated assault.

Crime seriousness was found to be more in line with the way the judiciary perceived the seriousness of the crimes presented in the present study than the way the legislature perceived these crimes. It may be concluded that New Zealand court judges are either of the same opinion regarding crime seriousness as the small sample of the New Zealand public studied, and/or keep abreast of public opinion.

While the results of the present study are generally in agreement with previous findings regarding high levels of relative consensus found, it seems quite astonishing that subjects, across studies, across cultures, and using different scaling techniques can produce results that are reasonably similar on the basis of a single-word description of a crime. In interpreting seriousness, it is not unreasonable to assume that subjects base their judgements on their personal knowledge (direct or indirect) of particular crimes. Yet the range of seriousness within a crime can be enormous. For instance, an armed robbery may range from a simple affair involving the robber and the victim where a gun is brandished just once as a threat, to the terrorising of a large group of people for several minutes, even hours, and where one or more victims may be injured. Why does this variation appear not to be reflected in judgements of seriousness? Is it that people do not or cannot take such variations within a crime into account?

On the bases of the above reasoning, it could be predicted that if crime seriousness was varied within a crime by making the crime more or less severe, these variations would have no effect on the scaling of the perceptions of crime seriousness. A second study was undertaken to investigate whether the positions of the 10 crimes on the crime seriousness scale could be manipulated by altering the degree of severity of the crime.

EXPERIMENT 2
INTRODUCTION

It was clear from the first experiment that although subjects were not given a definition of crime seriousness they had no trouble understanding the term with respect to their own perceptions. For instance, it was demonstrated that subjects readily and consistently discriminated between crimes according to their seriousness. It could also be concluded from Experiment 1 that the meaning imparted to the term, "seriousness", was shared to large extent, as high levels of consensus were found between groups that varied in terms of sex and occupation. Results from the first experiment, then, suggest that subjects did not fluctuate markedly in their perception of the seriousness of each crime held.

A second study was carried out to find out if subjects would take into account the level of seriousness within a crime (thereby altering the position of the crime on the crime seriousness scale), and whether this would alter the level of relative consensus found between subjects and groups.

There are several methods available to the researcher for altering the level of a crime's seriousness. Seriousness is a conceptual dimension of criminality that in terms of the law has two major components, culpability and harm. The culpability, or the blameworthiness, of the offender involves the perception of the extent of the offender's awareness, motivation and intention in relation to the crime. The culpability of an offender in criminal law is usually enhanced if an offender holds a criminal record.

The concept of harm refers to the injury inflicted by a prohibited act, or the omission of an act (Von Hirsch, 1983*a*). It has been demonstrated that there are at least two different types of harm that can be inflicted by the offender, these being physical and economic (Parton et al., 1991). The results of the first experiment in the present study support the

above observations in that when ad hoc classifications of the 10 stimulus crimes were made, subjects classified crimes into those that produced physical harm (crimes against the person), and those that produced economic harm (crimes against property).

Perceptions of crime seriousness can be increased or decreased by changing the level of either culpability of the offender or the harm inflicted by the offender. For instance, Hawkins (1980) and Sebba (1980) both concluded that the culpability of the offender and offender's characteristics, such as status, age, sex and income, affected respondents' perception of the seriousness of a crime. Generally, as the offenders culpability increased, so too did the perceived seriousness of the crime.

However, although vignettes containing offender characteristics and identities might add realism to a study, they also produce possible interactions between respondent characteristics and item characteristics (Parton et al., 1991). It then becomes difficult, if not impossible, to determine what factor, or factors, are influencing the experimental outcome. Because of these problems, offender intent and offender characteristics were not used in the present study to alter a crime's seriousness. Instead, the seriousness level of each crime was varied by quantity of harm done, the dollar value stolen or the amount of physical harm meted out.

Previous research has already shown that quantity of harm done can alter a subject's perception of seriousness for a particular crime. For instance, Sellin and Wolfgang (1964) established that the relationship between the judged seriousness of theft was a power function of the dollar value stolen. As the dollar value increased so did the seriousness of the crime. Replication by Davis (1992) yielded a similar power function relationship.

Figlio (1975) found similar results in his study. He concluded that the relationship between the seriousness score given to theft was strongly related to the dollar value

stolen. Figlio also found that the amount of injury sustained by a victim was related to the seriousness score given, with the seriousness of a crime increasing with the amount of injury sustained.

Multidimensional studies have shown that seriousness is composed of two or more dimensions, the primary dimensions of which have been found to be victim harm, constituting violence against the person and property damage (Forgas, 1980; Hansel, 1987; Howe, 1988). These studies have also shown that people do in fact hold stereotypic views of a crime. The underlying perception of crimes and the basis for making evaluations of a crime's seriousness have been found to include, commonness, intentionally, concern (Forgas, 1980), sex relatedness (Hansel, 1987) moral wrongfulness and deserved punishment (Howe, 1988). The common finding amongst studies is that harm done to the victim is always found to be the central criterion for evaluating a crime in terms of its seriousness. These studies provide further evidence that altering the level of harm done would also alter the level of a crime's seriousness.

The aim of the second experiment was to find out whether the relative positions of crimes could be altered on a crime seriousness scale by varying each crime's seriousness. The results of the first experiment show good agreement with similar studies in that the rank order and grouping of the crimes (single-word descriptors) are very similar. This result strongly implies that people hold a crime scenario of a particular crime that is largely unaffected by the range of potential seriousness of the crime. However, such a view is inconsistent with the results found using magnitude estimation which show that both the dollar value stolen (Davis, 1992; Sellin & Wolfgang, 1964) and the amount of physical injury sustained by a victim (Figlio, 1977) change the perceived seriousness of a crime.

These results are also inconsistent with studies that clearly show that people *do* have stereotypes of crimes that include the harm done to the victim which is used for assessing the seriousness of a crime (Hansel, 1987). It seems inconceivable that crime stereotypes

are constant across differing groups of people when experience and knowledge of crimes is inevitably different for different people. Indeed, Hansel's research suggests that there is little consensus between people in their stereotypic crimes. Hansel, using a sample of students, found that crime stereotypes were organised around basic perceptions of the level of violence, amount of property harm, sex relatedness or "sickness" of the crime. These stereotypic dimensions were found to be employed by subjects to evaluate crime seriousness. Hansel found significant disagreement regarding crime stereotypes. Different people adopted different crime stereotypes. However, this did not effect the the high level of consensus found amongst groups of people regarding their evaluations of crime seriousness. For example, rape was perceived by some subjects as a function of harm done and sex relatedness, for others others a function of violence. Even so, these differences in stereotypes did not alter the consensus found between people in the seriousness rating given for rape. Rape was perceived by everybody to be a serious crime.

Thus, while on the one hand, the high level of relative consensus found in the first experiment in the present study could suggest that variation in a crime scenario may not alter the relative seriousness of that crime, previous research suggests that changing the value of the dollars involved, or the physical harm done, will alter relative seriousness.

The principle objective of the second study was to find out if manipulating the perception of crime seriousness could alter the position of crimes on the crime seriousness scale. Variation in crime seriousness in Experiment 2 was achieved by having two 1-2 sentence descriptions of the same crime that differed only in the dollar value of goods stolen, or the amount of punishment inflicted. If crime positions changed on the crime seriousness scale, previous research that strongly implies that people do take into account the potential range in seriousness of each crime would be confirmed. Such a finding would also imply that harm done to the victim is employed by subjects to evaluate the seriousness of a

crime, lending support to research that suggests that harm done to the victim is a factor in the perception of a crime's seriousness.

The second study was also carried out to find out if relative consensus could be increased if a crime description, as opposed to a single-word descriptor, would increase relative consensus among subjects and between males and females. Both stereotypic research and research that suggests that crime seriousness can be altered through the amount of harm meted out to the victim indicates that more agreement between subjects will be reached if a crime description is given to subjects.

EXPERIMENT 2
METHOD

SUBJECTS

Data for the second experiment were collected from 51 Massey University psychology students who volunteered to take part in the study. Twenty four of these subjects were male and 27 were female. Ages ranged between 18 and 42 (mean = 23.00, s.d = 5.88) and 18 and 49 (mean = 25.15, s.d = 7.72) for females and males respectively.

QUESTIONNAIRE

Questionnaire format

Using the method of paired comparisons meant that numerous crime pairs had to be presented to the respondents (190 in all). Thus, the most efficient way to present the pairs was by computer. Pairs of crimes were simultaneously presented, one pair at a time, with the first crime in each pair being positioned above the second crime. When they were ready, subjects initiated proceedings by entering a code number which produced on-screen instructions for the questionnaire.

Following the instructions, five practice trials were presented for each subject after which the 190 crime pairs followed. The data were filed automatically and the resulting files were able to be transported into appropriate statistical packages for analysis.

Selection of crime types

The second study employed the same 10 crimes as those used in Experiment 1. However, unlike Experiment 1, each crime was dichotomised into a high and low serious version. The seriousness of each crime was controlled for by varying the quantity of harm inflicted upon the victim, either physical or economical, in a one sentence crime description. The crime descriptions were similar to those used by Sellin and Wolfgang (1964). For instance, theft was described as either "*The offender takes a cheque book and credit cards and obtains $1, 500 worth of stolen cash and goods*" (high serious manipulation), or "*The offender takes a cheque book from a bag and cashes a cheque to the value of $50.00 dollars*" (low serious manipulation). The quantity of money taken operationalises the level of seriousness.

The length for each of the crime seriousness descriptions was controlled for, the same number of words being used for the high and low serious version of each crime. The word limit had a minimum of 12 words and a maximum of 25 words. As far as was possible, only the numbers involved in manipulating seriousness were varied, the remainder of the description being identical for the two crime versions. The crime seriousness vignettes can be seen in Appendix C.

These high and low seriousness crime stimuli were presented as 190 pairs of offences taken two at a time in all possible pairs. As with Experiment 1, the ordering within a crime pair and the sequence for these pairs were arranged randomly.

Instructions

The instructions given to the subjects were essentially the same as those in Experiment 1, again outlining the forced-choice format, with added instructions on how to use the computer programme. These instructions were as follows:

The purpose of this study is to ascertain the opinion of several groups of people about crime seriousness. You will shortly be given a list of short descriptions of different crime types which have been arranged in pairs, 190 in all. Please decide which crime in the crime pair is more serious, the top crime description or the bottom crime description. For example, for the pair:

The offender kills the victim by repeated stabbings (20 times) before fleeing

The offender enters a back yard, while walking home one night. A bicycle is stolen which the offender uses to ride home

You would probably decide that the top crime is more serious and depress the letter T on the keyboard. If you decide that the bottom crime is more serious then you would depress the letter B on the keyboard. If you find a pair of crimes that seem equally serious (or equally inoffensive) be sure to pick one anyway, even if you have to make a random choice. Although you need think about your answer please do not use any long drawn out thought process. It should only take a couple of seconds to decided which crime in the pair is the more serious.

PROCEDURE

Arrangements were made with subjects so that data could be collected in small groups of three to ten. Subjects were seated at personal computers (one per subject) and given an informed consent form, a second form with a subject code number on it and questions relating to their age and sex (see Appendix C for an example).

The computer-based questionnaire did not allow for errors of judgement. If a subject depressed either T or B by mistake they could not return to that trial to correct it. Subjects were instructed to write down on the second form provided the trial number the mistake occurred on and the prefered answer. These data were entered in the subject's computer file at a later date by the experimenter.

The subjects were asked to read and sign the informed consent form and then enter their subject code number into the computer. The instructions on the computer screen were then explained and any questions the subjects had were answered. They were then given five practice trials before going on to answer the questionnaire. No time limit was imposed with the average time taken to complete the questionnaire being approximately 30 minutes.

EXPERIMENT 2
RESULTS AND DISCUSSION

The presentation of the results and discussion for Experiment 2 are divided into two main sections. The first section, as with Experiment 1, assesses the degree of internal consistency for each subject through a circular triad count. The interpretation and discussion of the scales obtained for each crime are presented and discussed in section two.

CIRCULAR TRIADS

The number of circular triads found for each individual was transformed into a coefficient of consistence. These coefficients of consistence were then categorised at 0.04 intervals and are shown in Table 8.

Table 8.
Coefficients of consistence categorised at 0.04 intervals. The coefficients ranged from 80% consistency to 100%, with 90% of subjects attaining a consistency coefficient of 88% or above.

Coefficients of Consistence	Number of subjects
0.97 - 1.00	25
0.93 - 0.96	16
0.89 - 0.92	6
0.85 - 0.88	2
0.80 - 0.84	2

It can be seen from Table 8 that all subjects in Experiment 2 displayed a high level of internal consistency, well above the cut off point in Experiment 1 (80%) for removing subjects from the analysis. The very high levels of internal consistency found replicated the previous findings of Experiment 1 showing that subjects have little difficulty in using the method of paired comparisons for the type of crimes used and the number of crimes to be scaled. In Experiment 1 it was found that the level of internal consistency was higher

than that found in previous research by Coombs (1967). It was proposed that the high level of internal consistency may have been due to differences in the number of crimes used in each study. Coombs, by using 19 crimes, may have produced a lower internal consistency than Experiment 1 which employed 10 crimes. The present study (Experiment 2) further suggests that it is not the number of crimes that may effect the level of internal consistency, but the number of different crimes. Coombs utilised 19 different crimes, while the present study utilised 20 crimes, but only 10 different crimes.

CRIME SERIOUSNESS SCALES

The paired comparisons data for the total sample are presented in Table 9. As with Experiment 1, each cell in the proportion matrix contains the proportion of subjects who perceived the column offence as more serious than the row offence. The columns and rows are presented in ascending order of seriousness. Data for females and males were also analysed separately and these proportions are presented in Tables 10 and 11 respectively. These proportions were then transformed into corresponding z matrices. (See Appendix D.)

As for Experiment 1, a large number of boundary probabilities existed. These boundary probabilities were again treated as missing data. Crime pairs that consisted of high and low levels of the same crime were also treated as missing data because of their unreliability. These crime pairs were considered unreliable as subjects were probably forced into the choice of judging the high seriousness version of the crime as the most serious out of the pair. As the two crimes were the same, varying only in terms of quantity of harm done, it would have been illogical for subjects to select the crime with the lesser quantity or harm done as the *more* serious. Leaving these crime pairs in the analysis may have had the effect of overly influencing the scales which may have resulted in the alteration of the true rank-order of the crimes presented.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Low Cannabis	0.500	0.822	0.843	0.902	0.980	1.000	0.961	0.941	0.980	1.000	0.941	1.000	1.000	0.941	1.000	1.000	1.000	1.000	1.000	1.000
2. Low Receiving		0.500	0.804	0.824	0.980	0.569	0.765	0.971	1.000	0.980	0.902	0.980	0.971	0.941	1.000	1.000	1.000	1.000	1.000	1.000
3. Low Burglary			0.500	0.549	0.765	0.529	0.843	0.922	1.000	0.980	0.922	1.000	1.000	0.961	1.000	0.961	1.000	1.000	1.000	1.000
4. Low Theft				0.500	0.627	0.667	0.814	0.863	0.980	0.961	0.922	1.000	1.000	0.843	0.980	0.980	1.000	1.000	0.980	1.000
5. High Receiving					0.500	0.627	0.804	0.784	0.941	0.980	0.804	0.980	1.000	0.882	1.000	1.000	1.000	1.000	1.000	1.000
6. High Cannabis						0.500	0.608	0.706	0.765	0.843	0.765	0.882	0.922	0.922	0.902	0.980	0.980	0.971	0.980	1.000
7. Low Ag. Assault							0.500	0.608	0.569	0.765	0.608	0.804	0.843	0.863	1.000	0.961	1.000	0.980	1.000	1.000
8. Low Arson								0.500	0.510	0.892	0.765	0.922	1.000	0.863	0.980	1.000	1.000	1.000	1.000	1.000
9. High Theft									0.500	0.980	0.647	1.000	0.980	0.824	0.980	0.980	1.000	1.000	1.000	1.000
10.High Burglary										0.500	0.647	0.922	0.902	0.706	0.882	0.980	0.922	0.980	0.941	1.000
11.Low Manslaughter											0.500	0.510	0.510	0.980	0.765	0.686	0.824	0.882	0.941	1.000
12.Low Kidnapping												0.500	0.529	0.608	0.866	0.941	0.961	1.000	1.000	1.000
13.High Arson													0.500	0.608	0.667	0.804	0.863	0.922	0.922	0.980
14.High Manslaughter														0.500	0.549	0.569	0.647	0.745	0.843	0.843
15.High Ag. Assault															0.500	0.569	0.804	0.980	0.941	0.980
16.High Kidnapping																0.500	0.843	0.922	0.961	0.941
17.Low Rape																	0.500	1.000	0.804	0.902
18.HighRape																		0.500	0.608	0.902
19.Low Murder																			0.500	0.980
20.High Murder																				0.500

Table 9.
Proportion matrix obtained for the total sample. The column offence is judged as more serious than the row offense. It is assumed that the obtained proportion for a crime paired with itself has a proportion of 0.50. Only the top right half of the table is presented as the lower half is its complement.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Low Cannabis	0.500	0.889	0.889	0.889	1.000	1.000	0.963	0.963	1.000	0.926	1.000	1.000	1.000	0.926	1.000	1.000	1.000	1.000	1.000	1.000
2. Low Receiving		0.500	0.852	1.000	0.815	0.667	0.926	1.000	0.778	0.852	1.000	1.000	0.963	0.963	1.000	1.000	1.000	1.000	1.000	1.000
3. Low Burglary			0.500	0.667	0.593	0.593	0.889	1.000	0.815	0.889	0.963	1.000	1.000	0.815	0.963	1.000	1.000	1.000	1.000	1.000
4. High Receiving				0.500	0.593	0.593	0.852	1.000	0.852	0.852	1.000	1.000	1.000	0.889	1.000	1.000	1.000	1.000	1.000	1.000
5. Low Theft					0.500	0.741	0.852	0.963	0.852	0.852	0.963	0.963	1.000	0.815	0.963	1.000	0.926	1.000	1.000	1.000
6. High Cannabis						0.500	0.741	0.778	0.630	0.741	0.852	0.852	0.926	0.889	0.923	0.963	1.000	1.000	1.000	1.000
7. Low Arson							0.500	0.519	0.704	0.667	0.963	0.889	1.000	0.815	1.000	0.963	1.000	1.000	1.000	1.000
8. High Theft								0.500	0.593	0.667	0.963	1.000	1.000	0.815	1.000	1.000	1.000	1.000	1.000	1.000
9. Low Ag. Assault									0.500	0.519	0.667	0.778	0.741	0.852	1.000	1.000	1.000	0.963	1.000	1.000
10.Low Manslaughter										0.500	0.630	0.593	0.593	1.000	0.667	0.778	0.815	0.852	0.963	1.000
11.High Burgalry											0.500	0.963	0.926	0.667	0.667	0.963	0.852	0.963	0.963	1.000
12.Low Kidnapping												0.500	0.556	0.593	0.963	0.889	0.926	1.000	1.000	1.000
13.High Arson													0.500	0.593	0.852	0.852	0.889	0.926	0.926	1.000
14.High Manslaughter														0.500	0.519	0.593	0.704	0.778	0.889	0.815
15.High Kidnapping															0.500	0.519	0.889	0.852	0.963	0.889
16.High Ag. Assault																0.500	0.815	1.000	1.000	0.963
17.Low Rape																	0.500	1.000	0.704	0.926
18.High Rape																		0.500	0.593	0.889
19.Low Murder																			0.500	1.000
20.High Murder																				0.500

Table 10.

Proportion matrix obtained for the female sample. The column offence is judged as more serious than the row offence.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Low Cannabis	0.500	0.875	0.792	0.958	1.000	0.917	0.917	0.917	1.000	1.000	0.958	1.000	1.000	1.000	1.000	0.958	1.000	1.000	1.000	1.000
2. Low Receiving		0.500	0.750	0.833	0.542	0.958	1.000	0.750	1.000	0.958	0.958	0.958	0.958	1.000	1.000	0.958	1.000	1.000	1.000	1.000
3. Low Burgary			0.500	0.500	0.542	0.875	0.958	0.875	1.000	1.000	0.958	1.000	1.000	1.000	0.958	0.958	1.000	1.000	1.000	1.000
4. Low Theft				0.500	0.583	0.667	0.833	0.833	1.000	0.958	1.000	1.000	1.000	0.958	1.000	0.875	1.000	1.000	0.958	1.000
5. High Cannabis					0.500	0.667	0.667	0.583	0.750	0.833	0.792	0.917	0.917	0.875	1.000	0.958	0.958	0.917	0.958	1.000
6. High Receiving						0.500	0.708	0.750	0.875	0.958	0.750	0.958	1.000	1.000	1.000	0.917	1.000	1.000	1.000	0.958
7. Low Arson							0.500	0.500	0.500	0.875	0.875	0.958	1.000	1.000	1.000	0.917	1.000	1.000	1.000	1.000
8. Low Ag. Assault								0.500	0.558	0.875	0.750	0.833	0.917	1.000	0.917	0.917	0.975	1.000	1.000	1.000
9. High Theft									0.500	1.000	0.625	1.000	0.958	0.958	0.958	0.833	1.000	1.000	1.000	1.000
10.High Burglary										0.500	0.667	0.875	0.875	0.792	1.000	0.750	1.000	1.000	0.917	1.000
11.Low Manslaughter											0.500	0.583	0.625	0.750	0.708	0.958	0.833	0.917	0.917	1.000
12.Low Kidnapping												0.500	0.625	0.833	0.917	0.625	1.000	1.000	1.000	1.000
13.High Arson													0.500	0.542	0.750	0.625	0.833	0.917	0.917	0.958
14.High Ag. Assault														0.500	0.667	0.708	0.792	0.958	0.875	1.000
15.High Kidnapping															0.500	0.625	0.792	1.000	0.958	1.000
16.High Manslaughter																0.500	0.583	0.708	0.792	0.875
17.Low Rape																	0.500	1.000	0.917	0.875
18.High Rape																		0.500	0.625	0.917
19.Low Murder																			0.500	0.958
20.High Murder																				0.500

Table 11.
Proportion matrix obtained for the male sample. The column offence is judged as more serious than the row offense.

Because of the missing data, the Edwards Case V incomplete data method was employed to scale the 20 crimes according to their perceived seriousness (Edwards, 1957). Successive difference matrices for the total sample, and for female and male subjects within this sample, are also shown in Appendix D. The scale values obtained for the total sample and for both sexes are shown in Table 12. The three corresponding seriousness scales are presented in Figure 4. The rank order of the 20 crimes for both females and males and the total sample are also shown in Table 12.

Table 12.
The scale values obtained for high and low seriousness crimes. Also included are the seriousness rank orders based upon the scale value found for each crime. Values for males and females separately are also shown.

CRIME	TOTAL (N = 52)		FEMALES (N = 27)		MALES (N = 24)	
	SCALE VALUES	RANK ORDER	SCALE VALUES	RANK ORDER	SCALE VALUES	RANK ORDER
High Murder	5.772	1	5.262	1	5.617	1
Low Murder	5.295	2	5.010	2	5.239	2
High Rape	5.196	3	4.528	3	5.106	3
Low Rape	4.755	4	4.286	4	4.738	4
High kidnapping	4.297	5	3.647	6	4.312	6
High Assault	4.025	6	3.909	5	4.153	7
High Manslaughter	3.917	7	3.536	7	4.477	5
High Arson	3.721	8	3.252	8	4.089	8
Low Kidnapping	3.522	9	3.163	9	3.684	9
Low Manslaughter	3.201	10	2.599	11	3.266	10
High Burglary	3.096	11	2.617	10	3.130	11
High Theft	2.133	12	2.034	13	2.543	12
Low Arson	2.012	13	1.905	14	2.038	14
Low Assault	1.932	14	2.336	12	2.378	13
High Cannabis	1.556	15	1.519	15	1.751	16
High Receiving	1.206	16	1.096	17	1.791	15
Low Theft	0.964	17	1.133	16	1.560	17
Low Burglary	0.809	18	0.973	18	1.082	18
Low Receiving	0.552	19	0.503	19	0.782	19
Low Cannabis	0.000	20	0.000	20	0.000	20

It is interesting to note that a clear distinction in seriousness between property and person orientated crimes persisted in Experiment 2 even after crime seriousness was manipulated. In general, it can be seen from the total sample scale that person orientated crimes (with the exception of low seriousness aggravated assault) were perceived as more serious than property oriented crimes (the exception being high seriousness arson). It seems that in general, harm done to the person persists as more serious than that done to property even if the harm done to the person is relatively minor

Manipulation of a crime's seriousness

Scaling of the crimes using Edwards (1957) modification of Thurstone's Case V method indicates that the perceived seriousness of a crime was altered by manipulating the quantity of harm inflicted upon the victim or by the dollar value taken from the victim. There are three major findings that support this conclusion. First, subjects were found to perceive the low seriousness crime descriptions of each crime as less serious than the corresponding high seriousness crime descriptions without exception. This was true whether quantity of harm done was measured in dollar value or the amount of bodily injury sustained. If one examines the 10 crimes individually, it can be seen (Table 13) that almost all subjects were influenced by the manipulation irrespective of the crime. The vast majority of subjects, irrespective of sex, perceived the high seriousness manipulation as the more serious crime.

It is impossible to tell whether the very small proportion of subjects who appeared to perceive a low seriousness description of a crime as more serious than a high seriousness description actually did so, or if this was simply due to subject error. For example, the results in Table 13 show that 2% of subjects perceived the theft of $50 as more serious than the theft of $1, 500. Given the overall pattern of results, it seems likely that such findings are due to subject error.

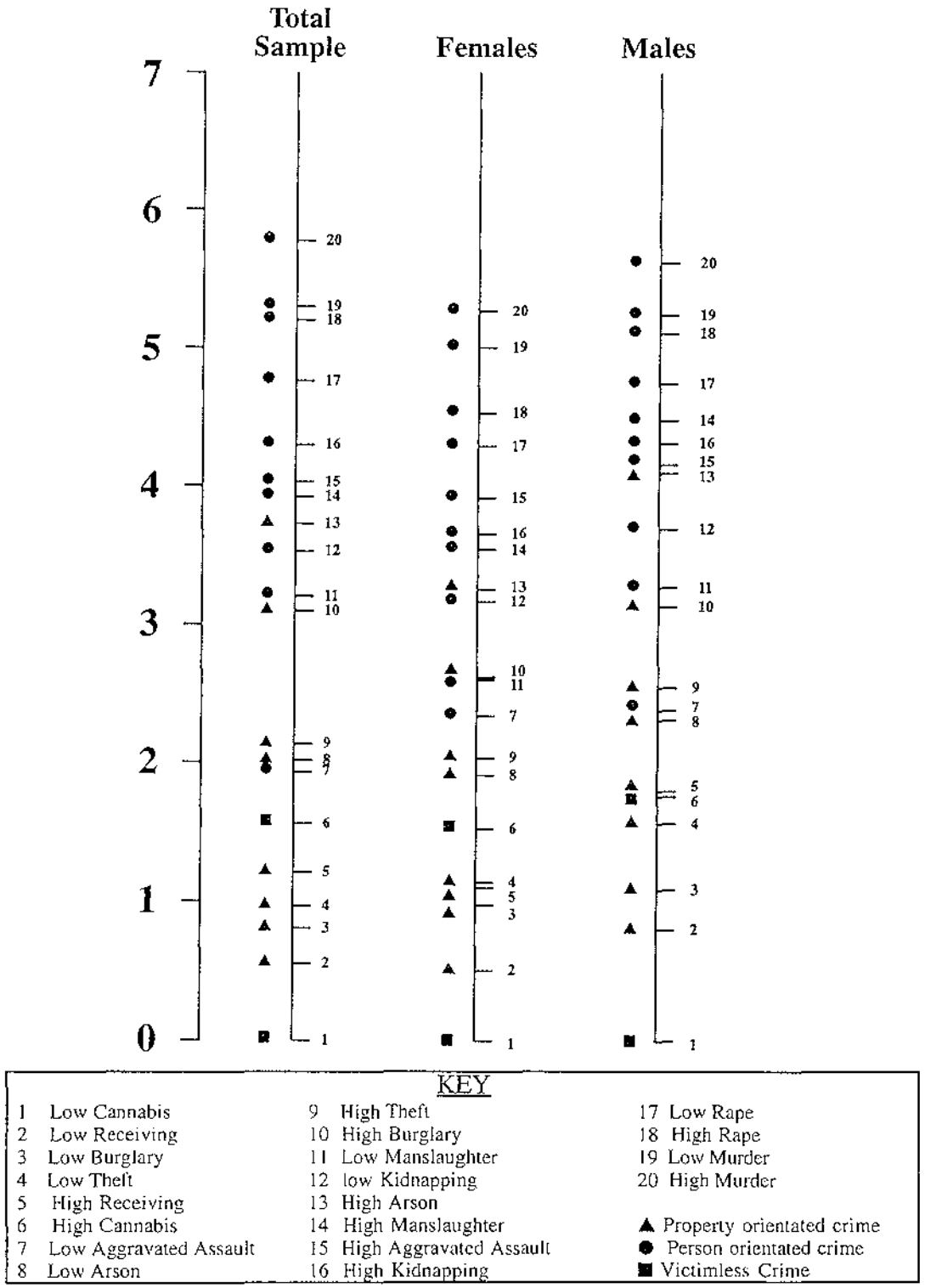

Figure 4. Interval scales for crime seriousness for the high and low seriousness descriptions of each offence that was presented in Experiment 1. In general, all property offences, both high and low seriousness descriptions, were perceived as less serious than person orientated offences, with two exceptions; high arson and low aggravated assault. Shifts along the scale from the low seriousness description of a crime to the high seriousness description of the same crime were larger for some offences than for others. In general, these shifts in seriousness were larger for crimes against property than for crimes against the person. Also shown are the scales for the male and female subjects.

Table 13.
Proportion of subjects who perceived the high serious manipulation of each crime as more serious than the low seriousness manipulation.

CRIME PAIR	TOTAL SAMPLE (N = 51)	FEMALES (N = 27)	MALES (N = 24)
High - Low Murder	0.980	1.000	0.958
High - Low Rape	1.000	1.000	1.000
High - Low Kidnapping	0.941	0.963	0.917
High - Low Manslaughter	0.980	1.000	0.958
High - Low Assault	1.000	1.000	1.000
High - Low Arson	1.000	1.000	1.000
High - Low Theft	0.980	0.963	1.000
High - Low Burglary	0.980	0.963	1.000
High - Low Receiving	0.980	1.000	0.958
High - Low Cannabis	1.000	1.000	1.000

Secondly, position itself was altered on the crime seriousness scale through high and low seriousness descriptions. For instance, as can be seen in Figure 4, the position of the crime of arson relative to other crimes for the total sample was different depending upon the seriousness level. Figure 4 shows the crime positions, in order of seriousness, differed according to whether the crime in question was a low or high seriousness description. A high seriousness description of arson was seen as less serious than a high serious description of aggravated assault. However, a low seriousness description of arson was seen as relatively more serious than a low serious description of aggravated assault. Theft was seen as more serious relative to burglary when a low serious crime description was given. The reverse was true when a high seriousness description was given. Similar results were found for the crimes of aggravated assault and manslaughter. When a low serious description of manslaughter was given to subjects, they perceived this crime as more serious than aggravated assault. However, aggravated assault was seen as more serious than manslaughter for the crimes with high serious descriptions. It is evident from these findings that both the high and low seriousness descriptions employed in the present study could alter the position of any given crime on the crime seriousness scale.

Thirdly, it can be seen from Figure 4 that not only could the position of each crime (in relation to all other crimes) be altered, but, that there were also notable shifts from the low

seriousness version of each crime to the high seriousness version. Some of these shifts along the scale are larger than others. Large differences in scale positions for low and high seriousness versions of the same crime were found for the crimes of burglary (which had a difference of 2.286 scale units), arson (1.709 scale units), theft (1.168 scale units), and aggravated assault (2.093 scale units). These scale unit differences are clearly larger than those obtained for other crimes such as rape, for which there was only a shift of 0.441 scale units from the low seriousness version to the high.

It is interesting to note that property orientated crimes shifted more along the scale than person orientated crimes, with the exception of aggravated assault for which there was a very significant shift in scale position. This finding suggests that the difference in the perception of seriousness between high and low seriousness descriptions for property orientated crimes was greater than for person orientated crimes. A possible explanation for these findings is that the range in seriousness within a person orientated crime may play less of a role in the evaluation of its seriousness than for a property orientated crime. (E.g., murder will always remain the most serious crime in comparison to all other crimes no matter what the range in seriousness within this crime is.)

In summary, the seriousness of a crime can be manipulated successfully through the quantity of harm done to the victim. In property oriented crimes, seriousness can be manipulated through quantity measured in dollar value. For example, as the monetary value of theft increased so too did the seriousness of the crime. This finding is similar to that of Davis (1992) and Sellin and Wolfgang (1964) who also found that the seriousness of a crime was directly related to the dollar value taken by an offender in a crime. For person orientated crimes, the quantity of harm done, measured by the amount of physical injury meted out to the victim, also manipulated the position of the crime on the seriousness scale. As the amount of physical violence to the victim increased, so too did perception of the crime's seriousness. These findings support those of Figlio (1975) who also found that the perceived seriousness of a crime increased as injuries meted out

to the victim increased. Both these findings suggest that subjects employed the amount of harm done to evaluate the seriousness of a crime.

Clearly manipulating quantity of physical harm done, or dollar amount, can alter the scale position of all the crimes used in the present study relative to all other crimes. The shifts in scale position for crimes in the present study ranged from large scale differences for property oriented crimes to small scale differences for person oriented crimes. These scale differences according to Thurstone's (1927*a*) theory can be translated into perceptual differences in crime seriousness. The present study indicates that there can be very large perceptual differences in the seriousness of a crime depending upon the range of seriousness within the crime itself. This implies that the crime scenario subjects may be using to judge the seriousness of a single-word description may also vary across subjects by as much. This suggests that reported variations in the levels of consensus across studies, and within studies, may be due to subjects having different crime scenarios which can be, it is assumed, made more uniform through a crime description.

If consensus between studies and within studies is shown to change according to whether or not a crime description (as opposed to a single-word descriptor) is given, contradictory findings between studies regarding consensus between different groups of people may be explained. Such results would also suggest that a crime description (however brief) should be given to control for this possible variation in crime scenarios and thus increase consensus. Therefore, further analysis was carried out to investigate whether or not consensus between subjects and between groups of subjects increased when a verbal description of a crime was used in place of a single-word descriptor.

Consensus among subjects and between groups

Among subjects Individual crime pairs were examined to investigate whether variability among subjects in their perceptions of crime seriousness decreased using crime

descriptions, as opposed to single-word descriptors. This was achieved by comparing direct proportions of specific crime pairs from the total sample in Experiment 1 (which employed single-word descriptors) with that of the high and low versions of specific crime pairs in Experiment 2 (which employed 1-2 sentence descriptions). The two crimes at the extremes of the crime seriousness scale (i.e., murder, rape, receiving stolen property and possession and use of cannabis) were left out of the crime pair analysis because boundary proportions were obtained for these crimes when paired with all other crimes.

Variability between subjects within the different samples could be measured using Coombs' (1967) concepts of homogeneity and heterogeneity. It may be recalled that proportions for specific crime pairs that are near 0.50 indicate heterogeneity. The closer to a proportion of 1.00 the crime pair proportion becomes, the more homogeneous the sample judging the crime pair is. The proportions obtained for the 8 crimes when paired with each other are presented in Table 14.

As can be seen from Table 14, homogeneity was slightly increased when a high serious crime description was given for the crime pairs containing burglary, arson and aggravated assault as more serious than the other crime in the pair. Homogeneity was also increased for the crime pairs of kidnapping-arson and kidnapping-aggravated assault. However, for all other crime pairs in which high seriousness descriptions were given homogeneity decreased. When low serious descriptions were compared to the single-word descriptions similar results were found. Although there was an increase in homogeneity across subjects for the crime pairs of aggravated assault-theft, kidnapping-theft, kidnapping-burglary, kidnapping-arson and kidnapping-aggravated assault, there was a decrease for all other crimes.

When high and low serious descriptions were averaged, no large increases in homogeneity for individual crime pairs were found. In summary, although problematic

ceiling effects are present, homogeneity across subjects generally did not increase with the presence of a crime description as opposed to a single-word crime descriptor. If anything, variability across subjects within a group slightly increased.

Table 14.
Proportions obtained for crime pairs in Experiment 1, using single-word descriptions, and Experiment 2, using high and low one-two sentence descriptions. The first crime in the pair is more serious than the second. Thus, for the first crime pair burglary is perceived as more serious than theft by 93.5% of all subjects in Experiment 1. It can be noted that no substantial increase in consensus is achieved by using a sentence description as opposed to a single-word description.

CRIME PAIR	EXPERIMENT 1: TOTAL SAMPLE Single-word descriptions	EXPERIMENT 2: TOTAL SAMPLE High-seriousness descriptions	EXPERIMENT 2: TOTAL SAMPLE Low-seriousness descriptions
Burglary-Theft	0.935	0.980	0.451
Arson-Theft	0.961	0.980	0.863
Arson-Burglary	0.896	0.902	0.922
Ag. Assault-Theft	0.961	0.980	0.814
Ag. Assault-Burglary	0.844	0.882	0.843
Ag. Assault-Arson	0.663	0.667	0.392
Kidnapping-Theft	0.987	0.980	1.000
Kidnapping-Burglary	0.987	0.980	1.000
Kidnapping-Arson	0.663	0.804	0.922
Kidnapping-Ag. Assault	0.519	0.569	0.804
Manslaughter-Theft	0.948	0.824	0.922
Manslaughter-Burglary	0.987	0.706	0.922
Manslaughter-Arson	0.883	0.608	0.765
Manslaughter-Ag. Assault	0.870	0.451	0.608
Manslaugter-Kidnapping	0.779	0.431	0.490

Sex A similar analysis was carried out comparing the two subject groups, males and females. As can be seen in Table 12 and Figure 4, high relative consensus (based on the rank ordering of crimes) existed between the sexes for crimes perceived as highly serious: both high and low seriousness murder and high and low seriousness rape as well as high seriousness kidnapping. High relative consensus also existed between the sexes for crimes perceived as least serious: low seriousness burglary, low seriousness receiving of stolen property and low seriousness possession and use of cannabis. With the exception of high seriousness arson, low seriousness kidnapping and low seriousness arson, all other crimes appeared in comparatively different rank-orders.

However, these rank-order differences between males and females in their perception of crime seriousness were minor as both rank orders of crimes according to their seriousness were found to be highly correlated ($r = .986$). This finding is similar to that of Experiment 1 in which female and male rank ordering of crimes according to their seriousness were also found to be highly correlated ($r = .964$). Adding a one-two sentence crime description did little to increase or decrease relative consensus between the sexes.

It therefore appears that differences in levels of relative consensus among subjects and between groups are not due to the presence (or absence) of a crime description, or to the fact that crime scenarios are present and may be different for different people. These results are similar to those of Davis (1992) who found no significant differences in perceived crime severity for single-word descriptors of crimes versus one-two sentence descriptions of crimes. The implication from Davis' finding is that the level of consensus across groups differing in age, socioeconomic status or sex also did not change.

The above findings also lend support to the results obtained by Hansel (1987). Hansel found that even though crime stereotypes were remarkably varied for different subjects, these differences in stereotypes did not alter the high levels of consensus found between different groups of subjects. The present findings suggest something of a paradox. It seems that relative consensus may slightly *decrease* when subjects are given a specific crime to consider, compared to a single-word descriptor. Thus, to maintain an acceptable level of relative consensus, the present results suggest that it may be best to use the single-word descriptors. On the other hand, the specific crime descriptions were clearly shown to alter markedly the relative positions of some of the crimes on the seriousness scale (see Figure 4). This result implies that relative consensus *could* be affected if different subjects held different scenarios for the same crime. However, based on the present findings, and those of Davis (1992) and Hansel (1987), one is forced to the conclusion that subjects do not envisage a specific scenario for a crime when rating its

seriousness, or, if they do, then the scenario must be remarkably similar across subjects. The question of just what lies at the basis of a subject's crime seriousness judgement is an important one and one which requires further research.

Turning now to examine specifically the male and female differences, Table 12 and Figure 4 demonstrate that the scale ranges for females (5.262) and males (5.617) differ. The difference of 0.355 suggests that the males in Experiment 2 were slightly more homogeneous than the females. This difference in scale ranges between the sexes was larger than the difference obtained in Experiment 1, the scale range difference being 0.001 there. Thus, a relatively substantial change in the level of homogeneity was found between females and males with the inclusion of a crime description as opposed to a single-word descriptor. These results suggest that women do not distinguish between crimes according to their seriousness as clearly as do men, nor do they agree with each other over the seriousness of the crimes presented to the same extent as do their male counterparts when presented with a crime description.

It is not clear why males became more homogeneous in their perception of crime severity than females when subjects were given a crime description. However, it has been well documented that perceptual differences do exist in the way men and women evaluate criminal phenomena (Lamp, 1982). For example, men have been shown to be more analytical in the perception of crime (Witkin, 1968, cited in Lamp, 1982). The differing levels of homogeneity found between the males and females in the present study when given a verbal description of a crime, as opposed to a single-word descriptor, may provide further evidence that there are certain differences in the perception of crime between males and females.

A direct comparison of the proportions obtained for males and females for individual crime pairs revealed differences between the sexes in their perception of crime severity for specific crime pairs, a finding similar to that of Experiment 1. As for Experiment 1, the

perceptual differences in the seriousness of rape between females and males was of special interest. Clear perceptual differences between women and men have been found with this crime (e.g.,Lamp, 1982).

In general, when male and female proportions obtained for the *low* seriousness manipulation of rape (when paired with every other crime) were compared, females regarded rape as more serious than males. This result is similar to that obtained by Lamp (1982) who also found that women perceive rape as more serious than men. Proportional differences in the perceived seriousness of low seriousness rape between males and females was most apparent when rape was paired with a person orientated crime as opposed to a property orientated crime.

Several large proportional differences in the perception of low seriousness rape between females and males were found and bear mentioning. While few males regarded low seriousness rape to be more serious than low seriousness murder (8.3%), a substantially larger percentage of women (29.6%) perceived this to be the case. More females (70.4%) also perceived low seriousness rape as more serious than high seriousness manslaughter than did males (58.3%). A substantial difference between the sexes was also found when low seriousness rape was paired with high seriousness kidnapping. While 88.9% of females perceived low seriousness rape as more serious only 79.2% of males thought this.

Certain exceptions did exist: Low serious rape was perceived as more serious by a higher percentage of males than females for the crimes of low seriousness theft, high seriousness burglary, low seriousness manslaughter and low seriousness kidnapping. However, in general, the low seriousness description of rape was seen as more serious by a higher percentage of females than males when compared to all other crimes where variations in percentages were found. This result is consistent with the results of Experiment 1, where it was also found that a higher frequency of females than males

rated rape as more serious than most other crimes. These results are similar to those of Lamp (1982). These findings may be due to the explanation that Lamp offered, that females perceive rape as more threatening than do males and therefore as more serious.

However, less variation between the sexes was found when the *high* seriousness rape description was paired with all other crimes. While a larger percentage of females regarded low seriousness rape as more serious than high seriousness possession of cannabis, aggravated assault , manslaughter, murder and low seriousness murder, a larger percentage of males perceived high seriousness burglary, kidnapping and low seriousness aggravated assault and manslaughter as more serious.

The element of *beating a women into submission so that she requires hospitalisation* as opposed to *raping a woman without further injuries* increased the seriousness of rape for men so that differences in the perception of rape between men and women lessened. It is interesting to consider why the introduction of the element of added violence increased rape seriousness for males so that differences between the sexes became minimal.

A possible explanation is that the crime scenario of rape may differ between men and women. Men may view rape as sexual rather than violent in nature. More women, on the other hand, may view rape as a violent rather than a sexual crime. Thus, the increase in violence or specification of violence increased the males' perception of crime seriousness for rape. It appears then, that it may not only be the threat of the crime of rape itself that may account for sex differences in the perception of rape, as Lamp (1982) surmised, but it is also possible that the crime scenario held of rape is different for females and males.

It is interesting to note that although these differences in crime scenarios may have occurred, the scale position for low and high serious rape remained the same for both men and women. Thus, relative consensus remained unchanged even though men and

women may hold different crime scenarios for rape; rape is a serious crime for both sexes.

EXPERIMENT 2: CONCLUSIONS

Crime positions on the crime seriousness scale can be manipulated through varying the quantity of harm done to the victim, indicating that subjects readily employ this information in their evaluations of crime seriousness. The implications for this finding are two-fold. First, one dimension of crime seriousness appears to be the amount of harm done to the victim, whether it be physical or financial.

Second, this manipulation of crime positions indicates that the rank order and distance on the crime seriousness scale of a crime, when compared to other crimes, is contingent upon the possible range in severity within the crime itself. A notable perceptual shift in seriousness from a low seriousness description to a high seriousness description was also found for every crime. For some crimes, mainly person orientated crimes, with the exception of aggravated assault, this change in seriousness was substantial.

However, when crime scenarios were controlled for, or made more uniform by presenting subjects with the actual amount of harm done in a crime description, little change was found in either the level of variability in subject evaluations of crime seriousness or relative consensus. The results found in Experiment 2 (using a one-two sentence crime description) reflected those found in in Experiment 1 (which employed a single-word descriptor). Variability in the evaluation of one crime as more serious than another remained the same, as did the high level of relative consensus between the sexes. It appears that the inclusion of crime description does not increase relative consensus between the sexes. A high level of consensus between subjects and groups can be reached, even though the present study also clearly shows that the range in seriousness within a crime can effect its perceptions of seriousness.

GENERAL DISCUSSION

Both Experiments 1 and 2 strongly support previous findings that the method of paired comparisons can usefully be employed to generate a crime seriousness scale (Borg, 1988; Carroll et al., 1974; Coombs, 1967; Krus et al., 1977; Rose & Prell, 1955; Thurstone, 1927*b*). Both of the present experiments also add to previous research carried out by Borg, (1988), Carroll et al., (1974), Coombs, (1967), and Krus et al., (1977) on the temporal nature of the perception of crime seriousness.

In 1927 Thurstone was the first to apply the law of comparative judgement to the perception of crime severity. Two major findings emerged from his early study. First, crimes could be categorised into person, property, or sex oriented offences. In general, these categories clustered in seriousness along the scale with crimes against the person perceived as the most serious (the exception being rape which was ranked as the most serious crime), followed by sex offences (which Thurstone classified as seduction, abortion, adultery and rape) and lastly property offences. Secondly, female and male differences were noted in the perception of crime severity with males being slightly more homogeneous than females in the ranking of crimes according to their seriousness.

Coombs (1967) followed up Thurstone's (1927*b*) research forty years later. Person orientated crimes were still seen as the most serious offences with murder replacing rape as the most serious crime on the scale. However, the distinction in seriousness between property and sex crimes was found to be less clear, with sex crimes becoming relatively less serious for Coombs' sample than for Thurstone's. When scales were produced for males and females separately, males were still more homogeneous than females.

In the 1970s, two studies were undertaken that employed the method of paired comparisons for scaling crime seriousness. Both these studies, conducted by Carroll et

al. (1974) and Krus et al. (1977), suggested that since Coomb's (1967) research, sex crimes continued to fall in seriousness relative to property orientated crimes. Crimes against the person continued to be perceived as as the most serious crimes, with murder persisting as the most serious. Borg (1988) found similar results. In general, sex crimes had progressed to the bottom of the scale in terms of seriousness, the exception being rape which still held the scale position of the second most serious crime. All crimes against the person were still judged as more serious than crimes against property .

Thus, 65 years after Thurstone's (1927*b*) original study two major factors have emerged. First, the sexual offences of abortion and adultery have lessened in seriousness over the years to the extent that in 1992 they may be considered largely moral rather than criminal offences. As such, these sex crimes were not presented to subjects in the present experiments. However, rape has retained the position of the second most serious crime (following murder) since 1967. It is important to note that the reasons for rape being perceived as a serious crime in 1992 may be different than the reasons given in 1967. It is possible that rape is perceived as a serious crime today because it is considered a crime against the person, rather than because it is seen as a sexual crime. The second factor that has emerged in since 1927 is that, in general, all crimes against the person have remained more serious than all crimes against property, even when, as in Experiment 2, a low seriousness crime against the person is paired with a high seriousness crime against property.

Experiment 1 showed that the ranking of crime seriousness is similar across groups, both occupational and gender groups, confirming both New Zealand and overseas findings that a high level of relative consensus exists between people in their perceptions of crime severity (Davis, 1992; Rossi et al., 1974). A high level of relative consensus was also found between the subjects in Experiment 1 and the New Zealand judiciary, supporting previous New Zealand research by Davis (1992). Also, similar results were obtained using the method of paired comparisons to those obtained by Davis (1992) using the

method of magnitude estimation. Thus, for a simple rank ordering of crime seriousness, it does not seem to matter which scaling technique is used.

Spier et al. (1991*a*) concluded that because of the doubt cast over the consensus found in existing overseas research, crime seriousness scales that employ public perceptions of crime severity should not be employed in New Zealand. However, the results of Davis (1992) and the present study do not support this view. If anything, these two New Zealand studies strengthen the justification for the use of public perception in evaluating crime severity, as a high level of relative consensus was found.

The results from both Experiment 1 and 2 indicate that this high level of relative consensus among subjects and between the sexes continues regardless of whether a single-word descriptor or a one-two sentence description of a crime is presented. This finding is similar to that of Davis (1992) who also found that the inclusion of a crime description did not change the level of agreement found between subjects presented with only a single-word descriptor.

However, the present findings are inconsistent with previous research which argues that people do take into account the wide range of seriousness within a crime (Davis, 1992; Figlio, 1975; Sellin & Wolfgang, 1964), and that crime stereotypes exist which are different for each individual judging the seriousness of a crime (Forgas, 1980; Hansel, 1987; Howe, 1988). Experiment 2, using a high or low seriousness description of the 10 previous crimes employed in Experiment 1, found subjects do take into account the severity of a crime when the severity is varied.

Unless subjects hold the same, or remarkably similar, scenarios it appears that they do not use specific crime scenarios from which they evaluate crime seriousness. If they did, one would expect greater variability between subjects when given a single-word descriptor as opposed to a one-two sentence description. Presently it remains unclear as

to exactly how subjects evaluate the seriousness of a crime: What does a single-word descriptor of a crime conjure up ? However, it is known, both from previous research and from the findings in Experiment 2, that part of this evaluation is based upon whether a crime is person or property orientated and the quantity of harm done, either physical or financial, to the victim.

Taking together, the results of previous overseas research, those of Davis (1992) and the present study, it appears that the pessimistic views advanced by Spier et al. (1991*a*) regarding the use of public perceptions of crime severity, based upon concern for low levels of public consensus, are unfounded. However, the present investigation indicates that the concern over research that addresses the public's perception of crime severity should shift to the problematic nature of the crime seriousness scales themselves.

LIMITATIONS OF THE PRESENT EXPERIMENTS

One obvious problem observed in both Experiments 1 and 2 relates to a weakness within the method of paired comparisons itself. This is the problem of boundary probabilities, a problem which has rarely been noted in previous research concerned with scaling crime severity using the method of paired comparisons. The lack of boundary problems in previous studies may be because both the number of crimes and subjects used in these studies were larger than in the present study. Hence, errors were more likely to occur.

The fact that previous research may have relied on subject error to define the end points of the scale gives cause for concern. The problem of boundary probabilities in the present study was especially noticeable for pairs of crimes where one crime ultimately was ranked high in seriousness and the other low in seriousness. This finding suggests that the method of paired comparisons may be better suited in investigating the perception of crimes that are perceived as relatively similar in seriousness, for example, studies investigating local relative consensus.

The large number of boundary probabilities found in both experiments lead to the use of Edwards (1957) Case V method of incomplete data scaling. This method sufficiently scaled crimes according to their seriousness for both of the present studies for relatively large samples. However, the method failed to produce scales for the smaller samples of female student subjects and male army subjects. The reason for the failure to produce scales for these smaller samples can be attributed to the large number of missing entries in the successive difference matrix for particular crimes; missing entries in the successive difference matrix are the result of no z values in the same row of the adjacent columns in the z matrix for a particular crime.

The substantial number of missing entries in the successive difference matrices for the end scale points in both experiments also raises the question of how reliable these scale points are. For the present experiments, the scale points for the crimes of rape and murder at the high seriousness end of the scale, and possession of cannabis and receiving stolen property at the low seriousness end, may be highly unreliable. Thus, it would seem that if the method of paired comparisons is to be used, a much larger number of subjects than was used in the present studies is required, especially if the sample is to be broken down into smaller groups for further analysis.

On the other hand, increasing subject numbers to decrease the high level of missing entries in the successive difference matrices may add to the unreliability of the end points of the scale, because with more subjects there is an increased likelihood that these end points are based upon subject error. The problematic nature of the method of paired comparisons may be resolved if its ordinal properties only are employed. Data can then be taken directly from the proportion tables as the sum of each crime paired with every other crime can be used to produce a reliable rank order. It is interesting to note that in all of the applications of the paired comparisons technique to date, the interval scaling property, theoretically established, is not used. (E.g., see Borg, 1988; Carroll et al., 1974; Coombs, 1966; Krus et al., 1977; Rose & Prell, 1955.) As the scale's interval

properties are very rarely, if ever, used in crime severity research it may be better to use the direct proportions obtained from the method of paired comparisons to produce a simple ordinal scale. Boundary probabilities of 1.00 or 0.00 do not create a problem of interpretation for ordinally scaled data.

Another limitation found with the method of paired comparisons was the data collected (proportions) do not lend themselves readily to assessing variability among subjects in their ratings of crime seriousness. As such, a measure of absolute consensus was unattainable. The ability to obtain a measure of absolute consensus using magnitude estimation, coupled with the problematic nature of the scales obtained through Thurstone's (1927*a*,*b*) method of paired comparisons, appears to suggest that we should use the magnitude estimation method in preference to the paired comparisons method.

However, the method of magnitude estimation is itself not without limitations. As discussed previously, several limitations exist with this method, the main one being that the method may not have true ratio properties as is claimed. Nonetheless, it is usually the case that the claimed ratio scaling properties of the magnitude estimation technique are used. For example, Davis (1992) employed the ratio scaling properties to calculate the dollar value associated with each crime. Despite the apparent advantages of the magnitude estimation scaling technique over the paired comparisons technique, by no means all researchers are willing to accept the assumptions that underlie this technique (Duncan , 1984; Marks, 1982, cited in Parton et al., 1991). Because the assumptions underlying the scale's ratio properties are questionable, the intentional use of magnitude estimation scaling for sentencing applications is also brought into question. For example, rape should receive a sentence twice as long as aggravated assault because it is twice as serious. Again, the ordinal properties of the scale, as with paired comparisons, may be less powerful but more reliable. Thus, if rank order is the primary research objective, choice of scaling technique may be a matter of personal preference rather than being based on theoretical argument or evidence.

A number of other limitations were associated with the present experiments. A short coming, seen in hindsight, with Experiment 2 was the presentation of both the high and low serious level of the *same* crime. This may have produced an instructional bias. Subjects may have realised that the intention of the researcher was to see whether they discriminated between high and low seriousness crimes. As such, it is possible that subjects may have chosen one crime as more serious than another purely because it had a greater dollar value or more physical violence associated with it. It is assumed, however, that this possible instructional bias did not affect results, because many subjects still choose a low seriousness person offence over a high seriousness property offence.

Other limitations which existed in the present experiments have been advanced by previous crime seriousness researchers. These limitations included the the small number of crimes used in the study, the limitation of using average and maximum sentences as indicators of the seriousness of the crimes by the New Zealand judiciary and legislature and the representativeness of the samples used. These limitations have previously been noted and discussed elsewhere, (e.g., Davis ,1992; Rossi & Henry, 1980). In the present study, pragmatic considerations governed sample sizes and the number of crimes used.

Finally, a possible limitation of the present research, which has been addressed by very few crime seriousness researchers, is the question of whether crime seriousness can legitimately be viewed as a unidimensional concept. As the present experiments show, the seriousness of a crime is evaluated at the least by the type of crime (person, property or victimless) and the quantity of harm done to the victims in these crimes. More research is required to evaluate the possible multidimensional nature of crime seriousness.

FUTURE RESEARCH

It seems that persuing the question of societal consensus in perceptions of crime severity may have become a redundant topic for crime seriousness researchers. A high level of relative consensus in the perception of crime seriousness has been found to exist, both in New Zealand and overseas, across differing community groups. However, before policy makers employ community perceptions in their decisions, the limitations of sample size, the number of crime types and representativeness of samples needs to be addressed. The scaling limitations found in both magnitude estimation and paired comparisons should also be addressed, although it is difficult to see how confidence in the scales can be improved beyond the ordinal level. Indeed, it may be that human beings, in whatever judgements they make, are limited to ordinal judgements.

Future research in the area of measuring public opinion of crime severity, both in New Zealand and overseas, needs to break away from traditional studies regarding public consensus and begin to focus upon the limitations of unidimensional scaling. The present study indicated that subjects do not hold specific crime scenarios. It remains to be seen exactly what subjects base their judgements on when judging the seriousness of a crime. The present study showed that both the quantity of physical harm done and the material loss experienced by a victim are probable factors in the evaluation of a crime's seriousness. Harm done may well be one factor that is used or envisioned when a single-word descriptor is given to evaluate. It may also be that any single crime conjures up, and is comprised of, other factors such as psychological damage, violence, the sexual nature of a crime and violation of human rights and privacy.

The factors involved in a person's perception of how serious a crime is are largely unknown. Future research should move away from the scaling of crime seriousness, as if it were a unidimensional concept, to examining its possible multidimensional nature. Methods such as factor analysis and multidimensional similarity scaling could be applied

to investigate these factors. These methods could also be employed to classify crimes into a hierarchy of offence seriousness according how crimes cluster in public evaluations of crime seriousness (e.g., is it a crime against property or a crime against the person). Such research might also include investigating whether or not differing groups (e.g., legal and non-legal raters) use the same factors, and give equal weight to these factors, when judging a crime's seriousness.

Once the factors (or dimensions) of crime seriousness are exposed, then further studies using the method of short crime descriptions employed in Experiment 2, could be undertaken to evaluate the degree to which each factor was able to shift the position of a crime on the crime seriousness scale.

GENERAL CONCLUSIONS

In conjunction with previous research two major conclusions can be drawn from the findings of the present study. These conclusions are as follows:

Magnitude estimation and paired comparisons scaling techniques seem to produce similar rankings in crime seriousness that are largely unaffected by group and gender differences. In the policy domain, high levels of relative consensus equate to the reliable use of public perceptions in prioritising scarce crime resources, such as police time.

Specific crime scenarios *can* change the scale position of a crime and produce substantial shifts in the perception of a crime's seriousness. Yet, the degree of consensus across studies shows that using single-word descriptors produces remarkable uniformity in the perception of crime seriousness. In conclusion, it appears that it is very unlikely that subjects use specific scenarios as the basis of their judgements in evaluating crime seriousness. If subjects are not basing these judgements on an 'image' or 'scenario' of a crime, the critical issue now is to find out more about what they do base their judgements on.

REFERENCES

Ackman, P., Normandeau, A., & Turner, S. (1967). The measurement of delinquency in Canada. *Journal of Criminal Law, Criminology and Police Science, 58,* 330-337.

Berk, R., & Rossi P. (1977). Prison reform and state elites. In D.F Hawkins, Perceptions of punishment for crime. *Deviant Behaviour, 1,* 193-215.

Beccaria, C. (1963). On crimes and punishment. (H. Paollucci, trans.). New York: MacMillan Publishing Company. (Original work published 1764.)

Blumstein, A., & Cohen, J. (1980). Sentencing convicted offenders: An analysis of the public's views. *Law and Society Review, 14,* 223-261.

Borg, I. (1988). Revisiting Thurstone's and Coombs' scales on the seriousness of crimes and offences. *European Journal of Social Psychology, 18,* 53-61.

Bowie, N., & Elliston. F. (Eds.) (1982). *Ethics, public policy and criminal justice.* Massachusetts: Gunn and Hain.

Bridges, G. S., & Lisagor, N. (1975). Scaling seriousness: An evaluation of magnitude and category scaling techniques. *The Journal of Criminal Law and Criminology, 66,* 215-221.

Carroll, R. M., Pine, S. M., Cline, C. J., & Kleinhans, B. R. (1974). Judged seriousness of watergate-related crimes. *The Journal of Psychology, 86*, 235-239.

Chilton, R., & DeAmicis, J. (1975). Overcriminalsation and the measurement of consensus. *Sociology and Social Research, 59*, 318-329.

Collins, M. F. (1988). Some cautionary notes on the use of the Sellin-Wolfgang index of crime seriousness. *Journal of Quantitative Criminology, 4*, 61-70.

Coombs, C. (1967). Thurstone's measurement of social values revisited forty years later. *Journal of Personality and Social Psychology, 6*, 85-91.

Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). *Mathematical psychology: An elementary introduction.* New Jersey: Prentice-Hall.

Crimes Act. (1961). *Reprinted statutes of New Zealand. vol 1.* Wellington: Government Print.

Cullen, F. T., Link, B. G., Travis III,L F., & Wozniak, J. F. (1985). Consensus in crime seriousness: Empirical reality or methodological artifact ? *Criminology, 23*, 99-118.

Davis, S. A. B. (1992). *Magnitude estimation of the seriousness of crime.* Unpublished masters thesis, University of Canterbury, Christchurch.

Duncan, O. D. (1984). *Notes on social measurement.*. New York: Sage.

Edwards, A. L. (1957). *Techniques of attitude scale construction.* New York: Appleton-Century-Crofts.

Figlio, R. M. (1975). The seriousness of offences: An evaluation by offenders and non-offenders. *The Journal of Criminal Law and Criminology, 66*, 189-200.

Fogel, D., & Hudson, J. (Eds.) (1981). *Justice as fairness: Perspectives on the justice model.* (p.45-46). Cincinnati: Anderson.

Forgas, J. P. (1980). Images of crime: A multidimensional analysis of individual differences in crime perception. *International Journal of Psychology, 15*, 287-29.

Fox, R., & Freiberg, A. (1990). Ranking offence seriousness in reviewing statutory maximum penalties. *Australian and New Zealand Journal of Criminology, 23*, 165-191.

Gebotys, R. J., Roberts, J. V., & DasGupta, B. (1988). News media use and public perception of crime seriousness. *The Canadian Journal of Criminology, 31*, 3-16.

Gescheider, G. A., Catlin, E. C., & Fontana, A. M. (1982). Psychophysical measurement of the judgement of crimes and severity of punishments. *Bulletin of the Pyschonomic Society, 19*, 275-278.

Gibbons, D. C. (1969). Crime and punishment: A study in social attitudes. *Social Forces, 47*, 391-397.

Guilford, J. P. (1954). *Psychometric methods.* (2nd ed.) New York : McGraw-Hill.

Hamilton, V. L., & Rytina, S. (1980). Social consensus on norms of justice: Should the punishment fit the crime ? *American Journal of Sociology, 85*, 1117-1144.

Hansel, M. (1987). Citizen crime stereotypes: Normative consensus revisited. *Criminology, 25*, 455-485.

Hart, H. M. (1958). The aims of the criminal law. *Law and Contemporary Problems, 23*, 401-441.

Haskell, M. R., & Yablonsky, L. (1974). *Criminology, crime and criminality.* Chicago: Rand McNally.

Hawkins, D. F. (1980). Perceptions of punishment for crime. *Deviant Behaviour, 1*, 193-215.

Hoffman, P. B, & Hardyman, P. L. (1986). Crime seriousness scales: Public perception and feedback to the criminal justice policymakers. *Journal of Criminal Justice, 14*, 413-431.

Howe, E. (1988). Dimensional structure of judgements of crimes. *Journal of Applied and Social Psychology, 18*, 1371-1393.

Kavalseth, T. O. (1980). Seriousness of offences: An experimental study based on a psychophysical scaling technique. *Criminology, 18*, 237-244.

Kendall, M. G. (1962). *Rank correlation methods.* (3rd ed.) London: Griffin .

Krus, D. J., Sherman, J. L., & Krus, P. H. (1977). Changing values over the last half-century: The story of Thurstone's crime scales. *Psychological Reports, 40*, 207-211.

Lamp, P. E. (1982). Sex differences in the perception of crime and criminals. *International Journal of Women's Studies, 5*, 413-422.

Leibrich, J., Galaway, B., & Underhill, Y. (1984) *Community Service Orders in New Zealand: The Research reports.* Wellington: Planning and Development Division: Department of Justice.

Levi, M., & Jones, S. (1985). Public and police perceptions of crime seriousness in England and Wales. *British Journal of Criminology, 25*, 234-250.

Makela, K. (1966). Public sense of justice and judicial practice. *Acta Sociologica, 10*, 42-67.

McFatter, R. M. (1982). Purposes of punishment: Effects of utilities of criminal sanctions on perceived appropriateness. *Journal of Applied Psychology, 67*, 255-317.

Miethe, T. D. (1982). Public consensus on crime seriousness: Normative structure or methodological artifact? *Criminology, 20,* 515-526.

Miethe, T. D. (1984). Types of consensus in public evaluations of crime: An illustration of strategies for measuring 'consensus'. *Journal of Criminal Law and Criminology, 75,* 459-473.

Misuse of Drugs Act. (1975). *Reprinted statutes of New Zealand. vol 26.* Wellington: Government Print.

Nevares-Muniz, D. (1984). The eighth amendment revisited: A model of weighted punishments. *The Journal of Criminal Law and Criminology, 75,* 272-289.

Newman, G. (1976). *Comparative deviance: Perception and law in six cultures.* New York: Elsevier.

Newman, G., & Trilling, C. (1975). Public perceptions of criminal behaviour: A review of the literature. *Criminal Justice and Behaviour, 2,* 217-36.

Parton, D. A., Hansel, M., & Stratton, J. R. (1991). Measuring crime seriousness: Lessons from the National Survey of Crime Severity. *British Journal of Criminology, 31,* 72-85.

Rose, A. M., & Prell, A. E. (1955). Does the punishment fit the crime ? A study in social valuation. *American Journal of Sociology, 61,* 247-259.

Rossi, P. H., Waite, E., Bose, C., & Berk, R. E. (1974). The seriousness of crimes: Normative structure and individual differences. *American Sociological Review, 39*, 224-237.

Rossi, P. H., & Henry, P. J. (1980). Seriousness: A measure for all purposes ? In M. Klein & J. Teilmann (Eds.) *Handbook of criminal justice evaluation*. Beverly Hills: Sage.

Roth, J. A. (1978). Prosecutor perceptions of crime seriousness. *Journal of Criminal Law and Criminology, 69*, 232-242.

Samuel, W., & Moulds, E. (1986). The effect of crime severity on perceptions of fair punishment: A Californian case study. *The Journal of Criminal Law and Criminology, 77*, 931-948.

Sellin, T., & Wolfgang, M. (1964). *The measurement of delinquency*. New York: Wiley.

Sebba, L. (1980). Is mens rea a component of perceived seriousness ? *The Journal of Criminal Law and Criminology, 71*, 124-135.

Sheley, J. F. (1980). Crime seriousness ratings: The impact of survey questionnaire form and item content. *British Journal of Criminology, 20*, 123-135.

Shrout, P. E., & Fleiss, J. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 66*, 420-428.

Suggs, D. L. (1981). A qualitative and quantitative analysis of the impact of Nebraska's decriminalisation of marijuana. *Law and Human Behaviour, 5,* 45-71.

Spier, P., Luketina, F., & Kettles, S. (1991*a*). *Changes in the seriousness of offending and in the pattern of sentencing: 1979 to 1988.* Wellington: Government Print.

Spier, P., Southly, P., & Norris, M. (1991*b*). *Conviction and sentencing of offenders in New Zealand: 1981 to 1990.* Wellington: Government Print.

Stevens, S. (1975). *Psychophysics: Introduction to its perceptual, neural and social prospects.* New York: Wiley.

Thomas, C. W., Cage, R. J., & Foster, S. C. (1976). Public opinion and legal sanctions: An examination of two conceptual models. *Journal of Law and Criminology, 67,* 110-116.

Thurstone, L. L. (1927*a*). A law of comparative judgement. *Psychological Review, 34,* 237-286.

Thurstone, L. L. (1927*b*). The method of paired comparisons for social values. *Journal of Abnormal and Social Psychology, 21,* 384 - 400.

Torgerson, W. S. (1958). *Theory and methods of scaling.* New York: Wiley.

Transport Act 1962. (1985). *Reprinted Statutes of New Zealand. vol. 16.* Wellington: Government Print.

Von Hirsch, A. (1983*a*). Commensurability and crime prevention: Evaluating formal sentencing structures and their rationale. *The Journal of Criminal Law and Criminology, 74*, 209-248.

Von Hirsch, A. (1983*b*). "Neoclassicism", proportionality and the rationale for punishment: Thoughts on the Scandinavian Debate. *Crime and Delinquency, 29*, 52-70.

Walker, M. A. (1978). Measuring the seriousness of crimes. *British Journal of Criminology, 18*, 348-364.

Walker, N. (1985). *Sentencing theory, law and practice.* London: Butterworths.

Waller, I. (1982). Public policy on crime and criminal justice: Who does and who should determine it ? In N. Bowie & F. Elliston (Eds.), *Ethics, Public Policy and Criminal Justice.* (p.353-369.) Massachusetts: Gunn and Hain.

Warr, M., Mier, R. F., & Erikson, M. L. (1983). Norms, theories of punishment, and publicly prefered penalties for crimes. *The Sociological Quarterly, 24*, 75-91.

Warr, M., Gibbs, J. P., & Erickson M. L. (1982). Contending theories of criminal law: Statutory Penalties versus public preferences. *Journal of Research in Crime and Delinquency, 19*, 25-46.

Wuillemin, D., Richardson, B., & Moore, D. (1986). Ranking of crime seriousness in Papua New Guinea: The effect of urbanization. *Journal of Cross Cultural Psychology, 17*, 29-44.

APPENDICES

APPENDIX A ... Questionnaire and informed consent form used used in Experiment 1.

APPENDIX B ... Proportion matrices, z matrices and successive difference matrices for the groups in Experiment 1.

APPENDIX C ... Crime descriptions and record sheet of demographic variables needed and errors that may have occurred for Experiment 2.

APPENDIX D ... Proportion matrices, z matrices and successive difference matrices for the total sample and female and male samples in Experiment 2.

APPENDIX A

PERCEPTIONS OF CRIME SERIOUSNESS.

The purpose of this study is to ascertain the opinion of several groups of people about crime seriousness.

The following list of crimes has been arranged in pairs. Please decide which one of the pair you think is the more serious and circle it. For example, murder - burglary; you will probably decide that murder is more serious and circle it.

Be sure to circle one of each of all the crime pairs presented.
If you find a pair of crimes that seem equally serious (or equally inoffensive) be sure to circle one of them anyway, even if you have to make a random choice.

Please take your time and think about your answer. Try to be consistent in your decision making. For example, if you circle murder as being more serious than arson, and arson more serious than burglary, then it follows that you should circle murder as more serious than burglary.

To ensure confidentiality, please do NOT write your name on this questionnaire.

1.	ARSON	-	POSSESSION AND USE OF CANNABIS
2.	MANSLAUGHTER	-	KIDNAPPING
3.	BURGLARY	-	MURDER
4.	KIDNAPPING	-	AGGRAVATED ASSAULT
5.	MURDER	-	THEFT (OR STEALING)
6.	BURGLARY	-	RAPE
7.	AGGRAVATED ASSAULT	-	POSSESSION AND USE OF CANNABIS
8.	THEFT (OR STEALING)	-	MANSLAUGHTER
9.	MANSLAUGHTER	-	RECEIVING PROPERTY DISHONESTLY OBTAINED
10.	MURDER	-	KIDNAPPING

11. POSSESSION AND USE OF CANNABIS	-	MURDER
12. MANSLAUGHTER	-	BURGLARY
13. THEFT (OR STEALING)	-	BURGLARY
14. POSSESSION AND USE OF CANNABIS	-	MANSLAUGHTER
15. THEFT (OR STEALING)	-	ARSON
16. RECEIVING PROPERTY DISHONESTLY OBTAINED	-	POSSESSION AND USE OF CANNABIS
17. KIDNAPPING	-	ARSON
18. MURDER	-	RECEIVING PROPERTY DISHONESTLY OBTAINED
19. MANSLAUGHTER	-	ARSON
20. KIDNAPPING	-	POSSESSION AND USE OF CANNABIS
21. RAPE	-	RECEIVING PROPERTY DISHONESTLY OBTAINED
22. ARSON	-	AGGRAVATED ASSAULT
23. THEFT (OR STEALING)	-	AGGRAVATED ASSAULT
24. MURDER	-	RAPE
25. ARSON	-	BURGLARY
26. BURGLARY	-	RECEIVING PROPERTY DISHONESTLY OBTAINED
27. THEFT (OR STEALING)	-	KIDNAPPING
28. AGGRAVATED ASSAULT	-	MANSLAUGHTER
29. KIDNAPPING	-	RECEIVING PROPERTY DISHONESTLY OBTAINED
30. POSSESSION AND USE OF CANNABIS	-	RAPE

31. ARSON	-	RECEIVING PROPERTY DISHONESTLY OBTAINED
32. MANSLAUGHTER	-	RAPE
33. POSSESSION AND USE OF CANNABIS	-	BURGLARY
34. KIDNAPPING	-	RAPE
35. AGGRAVATED ASSAULT	-	RECEIVING PROPERTY DISHONESTLY OBTAINED
36. KIDNAPPING	-	BURGLARY
37. MURDER	-	AGGRAVATED ASSAULT
38. THEFT (OR STEALING)	-	POSSESSION AND USE OF CANNABIS
39. BURGLARY	-	AGGRAVATED ASSAULT
40. MURDER	-	MANSLAUGHTER
41. ARSON	-	RAPE
42. MURDER	-	ARSON
43. THEFT (OR STEALING)	-	RECEIVING PROPERTY DISHONESTLY OBTAINED
44. RAPE	-	AGGRAVATED ASSAULT
45. THEFT (OR STEALING)	-	RAPE

Thank you for completing the questionnaire, please answer the last two questions which are needed for statistical purposes.

A. SEX M / F

B. AGE ☐☐

THANK YOU FOR YOUR CO-OPERATION

MASSEY
UNIVERSITY

Private Bag
Palmerston North
New Zealand
Telephone 0-6-356 9099
Facsimile 0-6-350 5611

FACULTY OF
SOCIAL SCIENCES

DEPARTMENT OF
PSYCHOLOGY

APPENDIX A Cont.

INFORMED CONSENT FORM

Massey University

This is to certify that I hereby agree to participate as a volunteer subject in a scientific investigation by Lucielle Trynes, under the supervision of Dr. John Podd.

* The investigation has been explained to me and I understand the explanation.
* I have been given the opportunity to ask questions, and these have been answered to my satisfaction.
* I understand that I am free to deny any answer to specific questions in the questionnaire.
* I understand that any questions will remain confidential with regard to my identity.
* I understand that I will not be informed as to which experimental group I am in until the study has been completed.
* I FURTHER UNDERSTAND THAT I AM FREE TO WITHDRAW MY CONSENT AND TERMINATE MY PARTICIPATION AT ANY TIME.

Participant's signature_____ Date_____

I, the undersigned, have defined and fully explained the investigation to the above participant.

Investigators signature_____ Date_____

APPENDIX B

The following proportion tables are those obtained in Experiment 1 for students and army personnel and for the female and male samples within these two groups. The column offence in all these tables is judged as more serious than the row offence. It is assumed that the obtained proportion for a crime paired with itself is 0.50. Only the top right half of the table is presented as the bottom half is its complement.

Proportion Matrix: Army and Students: Females

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.500	0.775	0.925	0.975	0.975	1.000	1.000	1.000	1.000	1.000
2.Receiving		0.500	0.850	1.000	0.975	1.000	1.000	0.950	1.000	1.000
3.Theft			0.500	0.975	0.950	0.975	1.000	0.925	1.000	1.000
4.Burglary				0.500	0.850	0.975	0.875	0.975	1.000	1.000
5.Arson					0.500	0.625	0.775	0.850	1.000	1.000
6.Kidnapping						0.500	0.650	0.875	0.925	1.000
7.Ag. Assault							0.500	0.825	0.975	1.000
8.Manslaughter								0.500	0.550	1.000
9.Rape									0.500	0.850
10.Murder										0.500

Proportion Matrix: Army and Students: Males

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.500	0.811	0.919	0.946	1.000	0.946	1.000	1.000	1.000	1.000
2.Receiving		0.500	0.946	0.973	0.919	0.973	1.000	0.973	1.000	1.000
3.Theft			0.500	0.892	0.919	0.973	1.000	0.973	1.000	1.000
4.Burglary				0.500	0.811	0.946	1.000	1.000	1.000	1.000
5.Ag. Assault					0.500	0.541	0.622	0.919	0.919	1.000
6.Arson						0.500	0.703	0.919	0.946	1.000
7.Kidnapping							0.500	0.676	0.757	1.000
8.Manslaughter								0.500	0.514	1.000
9.Rape									0.500	0.973
10.Murder										0.500

Proportion Matrix: Total Student Sample

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.500	0.800	0.925	0.975	0.975	1.000	1.000	1.000	1.000	1.000
2.Receiving		0.500	0.850	0.975	0.950	0.975	1.000	0.975	1.000	1.000
3.Theft			0.500	1.000	1.000	0.975	0.975	1.000	1.000	1.000
4.Burglary				0.500	0.850	0.925	0.975	1.000	1.000	1.000
5.Arson					0.500	0.550	0.575	0.900	0.975	1.000
6.Ag.Assault						0.500	0.575	0.900	0.975	1.000
7.Kidnapping							0.500	0.825	0.800	1.000
8.Manslaughter								0.500	0.525	1.000
9.Rape									0.500	0.900
10.Murder										0.500

Proportion Matrix: Female Students

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1.Cannabis	0.500	0.800	0.900	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2.Receiving		0.500	0.800	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3.Theft			0.500	1.000	1.000	0.950	1.000	0.950	1.000	1.000
4.Burglary				0.500	0.900	0.950	0.850	1.000	1.000	1.000
5.Arson					0.500	0.550	0.700	0.950	1.000	1.000
6.Kidnapping						0.500	0.750	0.900	0.900	1.000
7.Ag. Assault							0.500	0.900	1.000	1.000
8.Manslaughter								0.500	0.550	1.000
9.Rape									0.500	0.850
10.Murder										0.500

Proportion Matrix: Male Students

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.500	0.800	0.950	0.950	1.000	0.950	1.000	1.000	1.000	1.000
2.Receiving		0.500	0.900	0.950	0.900	0.950	1.000	0.950	1.000	1.000
3.Theft			0.500	1.000	0.950	1.000	1.000	1.000	1.000	1.000
4.Burglary				0.500	0.850	0.950	1.000	1.000	0.950	1.000
5.Ag. Assault					0.500	0.600	0.600	0.900	0.950	1.000
6.Arson						0.500	0.600	0.850	0.950	1.000
7.Kidnapping							0.500	0.750	0.700	1.000
8.Manslaughter								0.500	0.500	1.000
9.Rape									0.500	0.950
10.Murder										0.500

Proportion Matrix: Army Total

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.500	0.784	0.919	0.946	0.946	1.000	1.000	1.000	1.000	1.000
2.Receiving		0.500	0.946	1.000	0.973	0.973	1.000	0.946	1.000	1.000
3.Theft			0.500	0.865	0.919	0.946	1.000	0.919	1.000	1.000
4.Burglary				0.500	0.865	0.838	1.000	0.973	1.000	1.000
5.Arson					0.500	0.703	0.757	0.865	0.973	1.000
6.Ag. Assault						0.500	0.541	0.838	0.919	1.000
7.Kidnapping							0.500	0.730	0.892	1.000
8.Manslaughter								0.500	0.568	1.000
9.Rape									0.500	0.919
10.Murder										0.500

Proportion Matrix: Army Females

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.500	0.750	0.950	0.950	0.950	1.000	1.000	1.000	1.000	1.000
2.Receiving		0.500	0.900	1.000	0.950	1.000	1.000	0.900	1.000	1.000
3.Theft			0.500	0.950	0.900	1.000	1.000	0.900	1.000	1.000
4.Burglary				0.500	0.800	1.000	0.900	0.950	1.000	1.000
5.Arson					0.500	0.700	0.850	0.750	1.000	1.000
6.Kidnapping						0.500	0.550	0.850	0.950	1.000
7.Ag. Assault							0.500	0.750	0.950	1.000
8.Manslaughter								0.500	0.650	1.000
9.Rape									0.500	0.850
10.Murder										0.500

Proportion Matrix: Army Males

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.500	0.824	0.882	0.941	0.941	1.000	1.000	1.000	1.000	1.000
2.Receiving		0.500	1.000	1.000	1.000	0.941	1.000	1.000	1.000	1.000
3.Theft			0.500	0.765	0.941	0.882	1.000	0.941	1.000	1.000
4.Burglary				0.500	0.882	0.765	1.000	1.000	1.000	1.000
5.Arson					0.500	0.588	0.824	1.000	0.941	1.000
6.Ag. Assault						0.500	0.706	0.941	0.824	1.000
7.Kidnapping							0.500	0.588	0.824	1.000
8.Manslaughter								0.500	0.529	1.000
9.Rape									0.500	1.000
10.Murder										0.500

APPENDIX B Cont.

The following *z* matrices are those obtained in Experiment 1 for students and army personnel and for the female and male samples within these two groups. They are comprised of the *z* values of each proportion from the previous proportion matrices and as such the column offence in all these tables is judged as more serious than the row offence.

z Matrix: Student and Army Females

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.000	0.755	1.440	1.960	1.960					
2.Receiving		0.000	1.036		1.960			1.645		
3.Theft			0.000	1.960	1.645	1.960		1.440		
4.Burglary				0.000	1.036	1.960	1.150	1.960		
5.Arson					0.000	0.318	0.755	1.036		
6.Kidnapping						0.000	0.385	1.150	1.440	
7.Ag. Assault							0.000	0.935	1.960	
8.Manslaughter								0.000	0.125	
9.Rape									0.000	1.036
10.Murder										0.000

z Matrix: Student and Army Males

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.000	0.881	1.399	1.608		1.608				
2.Receiving		0.000	1.608	1.927	1.399	1.927		1.927		
3.Theft			0.000	1.237	1.399	1.927		1.927		
4.Burglary				0.000	0.881	1.608				
5.Ag. Assault					0.000	0.103	0.310	1.399	1.399	
6.Arson						0.000	0.533	1.399	1.608	
7.Kidnapping							0.000	0.456	0.696	
8.Manslaughter								0.000	0.035	
9.Rape									0.000	1.927
10.Murder										0.000

z Matrix: Total Student Sample

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.000	0.841	1.440	1.960	1.960					
2.Receiving		0.000	1.036	1.960	1.645	1.960		1.960		
3.Theft			0.000			1.960	1.960			
4.Burglary				0.000	1.036	1.440	1.960			
5.Arson					0.000	0.125	0.189	1.282	1.960	
6.Ag.Assault						0.000	0.189	1.282	1.960	
7.Kidnapping							0.000	0.935	0.841	
8.Manslaughter								0.000	0.063	
9.Rape									0.000	1.282
10.Murder										0.000

z Matrix: Female Students

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.000	0.841	1.282							
2.Receiving		0.000	0.841							
3.Theft			0.000			1.645		1.645		
4.Burglary				0.000	1.282	1.645	1.036			
5.Arson					0.000	0.125	0.524	1.645		
6.Kidnapping						0.000	0.674	1.282	1.282	
7.Ag. Assault							0.000	1.282		
8.Manslaughter								0.000	0.125	
9.Rape									0.000	1.036
10.Murder										0.000

z Matrix: Male Students

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.000	0.841	1.645	1.645		1.645				
2.Receiving		0.000	1.282	1.645	1.282	1.645		1.645		
3.Theft			0.000		1.645					
4.Burglary				0.000	1.036	1.645			1.645	
5.Ag Assault					0.000	0.253	0.253	1.282	1.645	
6.Arson						0.000	0.253	1.036	1.645	
7.Kidnapping							0.000	0.674	0.524	
8.Manslaughter								0.000	0.000	
9.Rape									0.000	1.645
10.Murder										0.000

z Matrix: Total Army Sample

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.000	0.786	1.399	1.608	1.608					
2.Receiving		0.000	1.608		1.927	1.927		1.608		
3.Theft			0.000	1.103	1.399	1.608		1.399		
4.Burglary				0.000	1.103	0.986		1.927		
5.Arson					0.000	0.533	0.696	1.103	1.927	
6.Ag. Assault						0.000	0.103	0.986	1.399	
7.Kidnapping							0.000	0.612	1.237	
8.Manslaughter								0.000	0.171	
9.Rape									0.000	1.399
10.Murder										0.000

z Matrix: Female Army Sample

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.000	0.674	1.645	1.645	1.645					
2.Receiving		0.000	1.282		1.645			1.282		
3.Theft			0.000	1.645	1.282			1.282		
4.Burglary				0.000	0.841		1.282	1.645		
5.Arson					0.000	0.524	1.036	0.674		
6.Kidnapping						0.000	0.125	1.036	1.645	
7.Ag. Assault							0.000	0.674	1.645	
8.Manslaughter								0.000	0.385	
9.Rape									0.000	1.036
10.Murder										0.000

z Matrix: Male Army Personnel

	1	2	3	4	5	6	7	8	9	10
1.Cannabis	0.000	0.931	1.185	1.564	1.564					
2.Receiving		0.000				1.564				
3.Theft			0.000	0.722	1.564	1.185		1.564		
4.Burglary				0.000	1.185	0.722				
5.Arson					0.000	0.222	0.931		1.564	
6.Ag.Assault						0.000	0.541	1.564	0.931	
7.Kidnapping							0.000	0.222	0.931	
8.Manslaughter								0.000	0.073	
9.Rape									0.000	
10.Murder										0.000

APPENDIX B Cont.

The following successive difference matrices are those obtained in Experiment 1 for students and army personnel and female and male samples within these two groups.

Matrices of successive differences

Army and student females

STATEMENTS	COLUMN DIFFERENCES								
	C2-C1	C3-C2	C4-C3	C5-C4	C6-C5	C7-C6	C8-C7	C9-C8	C10-C9
1. Cannabis	0.755	0.685	0.521						
2. Receiving	0.755	1.036							
3. Theft	0.403	1.036	1.960	-0.315	0.315				
4. Burglary			1.960	1.036	0.924	-0.810	0.810		
5. Arson		0.315	0.609	1.036	0.318	1.073	0.281		
6. Kidnapping				2.279	0.318	0.385	0.766	0.289	
7. Ag. Assault				0.392	0.370	0.385	0.935	1.026	
8. Manslaughter		0.205	-0.521	0.924	-0.114	0.216	0.935	0.125	
9. Rape						0.766	1.835	0.125	1.036
10.Murder									1.036
SUM	1.913	3.277	4.529	5.352	2.131	2.015	5.562	1.565	2.072
N	3	5	5	6	6	6	6	4	2
MEAN	0.637	0.655	0.906	0.892	0.355	0.336	0.927	0.391	1.036
SCALE VALUES									
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
0.000	0.637	1.292	2.198	3.090	3.445	3.781	4.708	5.099	6.135

Army and student males

STATEMENTS	COLUMN DIFFERENCES								
	C2-C1	C3-C2	C4-C3	C5-C4	C6-C5	C7-C6	C8-C7	C9-C8	C10-C9
1. Cannabis	0.881	0.517	0.209						
2. Receiving	0.881	1.608	0.320	-0.529	0.529				
3. Theft	-0.209	1.608	1.237	0.161	0.529				
4. Burglary	-0.320	0.690	1.237	0.881	0.726				
5. Ag. assault			0.517	0.881	0.103	0.208	1.088		
6. arson	-0.320		0.320	1.505	0.103	0.533	0.866	0.209	
7. Kidnapping					-0.222	0.533	0.456	0.240	
8. Manslaughter						0.943	0.456	0.035	
9. Rape					-0.209	0.911	0.731	0.035	1.927
10.Murder									1.927
SUM	0.913	4.423	3.840	2.900	1.559	3.128	3.597	0.519	3.854
N	5	4	6	5	7	5	5	4	2
MEAN	0.183	1.106	0.640	0.580	0.223	0.625	0.720	0.130	1.927
SCALE VALUES									
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
0.000	0.183	1.289	1.929	2.509	2.732	3.357	4.077	4.207	6.134

Total student sample

STATEMENTS	COLUMN DIFFERENCES								
	C2-C1	C3-C2	C4-C3	C5-C4	C6-C5	C7-C6	C8-C7	C9-C8	C10-C9
1. Cannabis	0.841	0.598	0.521						
2. Receiving	0.841	1.036	0.924	-0.315	0.315				
3. Theft	0.403	1.036							
4. Burglary				1.036	0.403	0.521			
5. Arson	0.315			1.036	0.125	0.063	1.093	0.679	
6. Ag. Assault			0.521	1.314	0.125	0.189	1.093	0.679	
7. Kidnapping				1.772		0.189	0.935	-0.093	
8. Manslaughter						0.779	0.935	0.063	
9. Rape						1.119	0.779	0.063	1.282
10.Murder									1.282
SUM	2.400	2.670	1.966	4.843	0.968	2.860	4.835	1.391	2.564
N	4	3	3	5	4	6	5	5	2
MEAN	0.600	0.890	0.655	0.968	0.242	0.477	0.966	0.278	1.282
SCALE VALUES									
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
0.000	0.600	1.490	2.145	3.113	3.355	3.831	4.798	5.076	6.358

Female student sample

STATEMENTS	COLUMN DIFFERENCES								
	C2-C1	C3-C2	C4-C3	C5-C4	C6-C5	C7-C6	C8-C7	C9-C8	C10-C9
1. Cannabis	0.841	0.440							
2. Receiving	0.841	0.841							
3. Theft	0.440	0.841							
4. Burglary				1.282	0.363	-0.609			
5. Arson				1.282	0.125	0.399	1.121		
6. Kidnapping				1.520	0.125	0.674	0.608		
7. Ag. Assault				0.512	-0.150	0.674	1.282		
8. Manslaughter					0.363		1.282	0.125	
9. Rape								0.125	1.036
10.Murder									1.036
SUM	2.123	2.122	0.000	4.596	0.826	1.138	4.293	0.250	2.072
N	3	3	0	4	5	4	4	2	2
MEAN	0.707	0.707	0.000	1.149	0.165	0.285	1.073	0.125	1.036
SCALE VALUES									
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
0.000	0.707	1.414	1.414						

Note: The last six scale values have been left out as the number of column differences the scale value for burglary was to be based upon were non-existent. This would have resulted in unreliable scale values for the following six crimes.

Male student sample

STATEMENTS	COLUMN DIFFERENCES								
	C2-C1	C3-C2	C4-C3	C5-C4	C6-C5	C7-C6	C8-C7	C9-C8	C10-C9
1. Cannabis	0.841	0.804							
2. Receiving	0.841	1.282	0.363	-0.363	0.363				
3. Theft	0.363	1.282							
4. Burglary				1.036	0.609				
5. Ag. Assault		-0.363	0.609	1.036	0.253		1.029	0.363	
6. Arson				1.392	0.253	0.253	0.783	0.609	
7. Kidnapping						0.253	0.674	-0.150	
8. Manslaughter					0.245	0.362	0.674		
9. Rape						1.121	0.524		1.645
10.Murder									1.645
SUM	2.045	3.005	0.972	3.101	1.723	1.989	3.684	0.822	3.290
N	3	4	2	4	5	4	5	3	2
MEAN	0.682	0.751	0.486	0.775	0.345	0.497	0.773	0.274	1.645
SCALE VALUES									
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
0.000	0.682	1.433	1.919	2.694	3.039	3.536	4.237	4.547	6.192

Total army sample

STATEMENTS	COLUMN DIFFERENCES								
	C2-C1	C3-C2	C4-C3	C5-C4	C6-C5	C7-C6	C8-C7	C9-C8	C10-C9
1. Cannabis	0.786	0.613	0.209						
2. Receiving	0.786	1.608							
3. Theft	-0.209	1.608	1.103	0.296	0.209				
4. Burglary			1.103	1.103	-0.117				
5. Arson	-0.320	0.529	0.296	1.103	0.533	0.164	1.231	0.824	
6. Ag. Assault		0.320	0.621	0.454	0.533	0.103	0.884	0.412	
7. Kidnapping					0.594	0.103	0.612	0.625	
8. Manslaughter		0.209	0.529	0.824	0.117	0.374	0.612	0.171	
9. Rape					0.529	0.161	1.066	0.171	1.399
10.Murder									1.399
SUM	1.043	4.887	3.861	3.780	2.398	0.905	4.405	2.203	2.798
N	4	6	6	5	7	5	5	5	2
MEAN	0.261	0.814	0.643	0.756	0.342	0.181	0.881	0.441	1.399
SCALE VALUES									
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
0.000	0.261	1.075	1.718	2.474	2.816	2.997	3.878	4.319	5.718

Female army sample

STATEMENTS	COLUMN DIFFERENCES								
	C2-C1	C3-C2	C4-C3	C5-C4	C6-C5	C7-C6	C8-C7	C9-C8	C10-C9
1. Cannabis	0.674	0.971							
2. Receiving	0.674	1.282							
3. Theft	0.363	1.282	1.645	-0.363					
4. Burglary			1.645	0.841			0.363		
5. Arson		0.363	0.440	0.841	0.524	0.512	-0.362		
6. Kidnapping					0.524	0.125	0.911	0.609	
7. Ag. Assault				0.245	0.911	0.125	0.674	0.971	
8. Manslaughter			-0.363	0.971	-0.362	0.362	0.674	0.385	
9. Rape							0.289	0.385	1.036
10.Murder									1.036
SUM	1.711	3.898	3.367	2.535	1.597	1.124	2.550	2.350	2.072
N	3	4	4	5	4	4	6	4	2
MEAN	0.570	0.974	0.842	0.507	0.399	0.281	0.425	0.587	1.036
SCALE VALUES									
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
0.000	0.570	1.544	2.386	2.893	3.282	3.573	3.998	4.585	5.621

Male army sample

STATEMENTS	COLUMN DIFFERENCES								
	C2-C1	C3-C2	C4-C3	C5-C4	C6-C5	C7-C6	C8-C7	C9-C8	C10-C9
1. Cannabis	0.931	0.255	0.378						
2. Receiving	0.931								
3. Theft			0.722	0.841	-0.378				
4. Burglary			0.722	1.185	-0.463				
5. Arson			0.339	1.185	0.222	0.709			
6. Ag. Assault		0.378	0.463	0.500	0.222	0.541	1.022	-0.633	
7. Kidnapping					0.389	0.541	0.222	0.709	
8. Manslaughter						1.342	0.222	0.073	
9. Rape					0.633		0.858	0.073	
10.Murder									
SUM	1.861	0.633	2.664	3.711	0.625	3.133	2.324	0.222	0.000
N	2	2	5	4	6	4	4	4	0
MEAN	0.931	0.316	0.532	0.928	0.104	0.783	0.581	0.055	0.000
SCALE VALUES									
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
0.000	0.931	1.247	1.779	2.707	2.811	3.594	4.175	4.230	

Note : The scale value for murder has been left out as the column differences it was to be based upon were non-existent. This would have resulted in an unreliable and unstable scale end-point.

APPENDIX C

CRIME DESCRIPTIONS USED IN EXPERIMENT 2.

HIGH SERIOUSNESS DESCRIPTIONS

1. POSSESSION AND USE OF CANNABIS: The offender is found in a public bar in possession of 28 grams of cannabis, well in excess of the amount required for personal use.

2.RECEIVING STOLEN PROPERTY: The offender receives a tape deck from a friend worth $1,000 knowing that it had been stolen.

3. THEFT: The offender takes a cheque book and credit cards and obtains $1,500 worth of stolen cash and goods.

4.BURGLARY: The offender enters a house and takes $500 in cash, $5,000 worth of jewelery and uses the victim's car to escape.

5. ARSON: The offender set fire to a house. The victim escapes but the house and it's contents are destroyed.

6. AGGRAVATED ASSAULT: The offender repeatedly hits the victim about the head and body until stopped. The victim requires hospitalization.

7. KIDNAPPING: The offender kidnaps a person. After $20,000 is paid the victim is released but requires hospitalization.

8. MANSLAUGHTER: The offender, while carrying out his duties as a train driver, falls asleep. The train derails and 60 passengers are killed.

9. RAPE: The offender beats a woman into submission with his fists and rapes her. She requires hospitalization.

10. MURDER: The offender kills the victim by repeated stabbings (20 times) before fleeing.

LOW SERIOUSNESS DESCRIPTIONS

1. POSSESSION AND USE OF CANNABIS: The offender is found in a public bar in possession of 4 grams of cannabis, which the offender says, is for personal use only.

2. RECEIVING STOLEN PROPERTY: The offender receives a tape deck from a friend worth $50 knowing it had been stolen.

3. THEFT: The offender takes a cheque book from a bag and cashes a cheque to the value of $50.

4. BURGLARY: The offender enters a back yard while walking home at night. A bicycle is stolen which the offender uses to ride home.

5. ARSON: The offender sets fire to an empty garage, well away from the victim's house. Minimal damage is incurred.

6. AGGRAVATED ASSAULT: The offender punches the victim twice. The victim's injuries include a nose bleed and a black eye.

7. KIDNAPPING: The offender kidnaps a person. After $10, 000 in ransom is paid the victim is released with no physical injury.

8. MANSLAUGHTER: The offender, while carrying out his duties as a train driver, falls asleep. The train derails and one passenger is killed.

9. RAPE: The offender rapes a woman, but he does not inflict any further injuries upon her.

10. MURDER: The offender kills the victim with a single stab and then flees.

APPENDIX C

SUBJECT CODE, SUBJECT DEMOGRAPHIC DATA AND RECORD SHEET OF ERRORS MADE

YOUR SUBJECT CODE IS:_____

PLEASE STATE YOUR SEX: M / F

PLEASE STATE YOUR AGE:_____

Please write down the trial number and the answer you would have put if an error in your answering occurs.

Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____
Trial number _____	Answer_____

APPENDIX D

The follwing *z* matrices are those obtained in Experiment 2 for the total sample and for the female and male samples within this group.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 Low Cannabis	0.000	1.185	1.007	1.293	2.054		1.763	1.564	2.054		1.564			1.564						
2Low Receiving		0.000	0.856	0.931		0.173	0.722	1.896		2.054	1.293	2.054	1.896	1.564						
3 Low Burglary			0.000	0.123	0.722	0.073	1.007	1.419			1.419			1.763		1.763				
4.Low Theft				0.000	0.323	0.431	0.893	1.094		1.763	1.419			1.007	2.054	2.054			2.054	
5 High Receiving					0.000	0.458	0.856	0.786	1.564	2.054	0.856	2.054		1.185						
6 High Cannabis						0.000	0.274	0.541	0.722	1.007	0.722	1.185	1.419	1.419	1.293	2.054	2.054	1.896	2.054	
7 Low Ag. Assault							0.000	0.274	0.173	0.722	0.274	0.856	1.007	1.094		1.763		2.054		
8 Low Arson								0.000	0.025	1.237	0.722	1.419		1.094	2.054					
9 High Theft									0.000	2.054	0.377		2.054	0.931	2.054	2.054				
10High Burglary										0.000	0.377	1.419	1.293	0.541	1.185	2.054	1.419	2.054	1.564	
11Low Manslaughter											0.000	0.025	0.025		0.722	0.484	0.931	1.185	1.564	
12Low Kidnapping												0.000	0.073	0.274	1.108		1.763			
13High Arson													0.000	0.274	0.431	0.856	1.094	1.419	1.419	2.054
14High Manslaughter														0.000	0.123	0.173	0.377	0.659	1.007	1.007
15High Ag. Assault															0.000	0.173	0.856	2.054	1.564	2.054
16High Kidnapping																0.000	1.007	1.419	1.763	1.564
17Low Rape																	0.000		0.856	1.293
18HighRape																		0.000	0.274	1.293
19Low Murder																			0.000	
20High Murder																				0.000

z Matrix: Total Students

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 Low Cannabis	0.000	1.221	1.221	1.221			1.787	1.787		1.447				1.447						
2 Low Receiving		0.000	1.045		0.896	0.431	1.447		0.765	1.045			1.787	1.787						
3 Low Burglary			0.000	0.431	0.235	0.235	1.221		0.896	1.221				0.896	1.787					
4 High Receiving				0.000	0.235	0.235	1.045		1.045	1.045				1.221						
5 Low Theft					0.000	0.646	1.045		1.045	1.045	1.787	1.787		0.896	1.787		1.447			
6.High Cannabis						0.000	0.646	0.765	0.331	0.646	1.045	1.045	1.447	1.221	1.426	1.787				
7 Low Arson							0.000	0.048	0.536	0.431	1.787	1.221		0.896		1.787				
8 High Theft								0.000	0.235	0.431	1.787			0.896						
9 Low Ag. Assault									0.000	0.048	0.431	0.765	0.646	1.045				1.787		
10Low Manslaughter										0.000	0.331	0.235	0.235		0.431	0.765	0.896	1.045	1.787	
11High Burglary											0.000	1.787	1.447	0.431	0.431	1.787	1.045	1.787	1.787	
12Low Kidnapping												0.000	0.141	0.235		1.221	1.447			
13High Arson													0.000	0.235	1.045	1.045	1.221	1.447	1.447	
14High Manslaughter														0.000	0.048	0.235	0.536	0.765	1.221	0.896
15High Kidnapping															0.000	0.048	1.221	1.045	1.787	1.221
16High Ag. Assault																0.000	0.896			1.787
17Low Rape																	0.000		0.536	1.447
18High Rape																		0.000	0.235	1.221
19.Low Murder																			0.000	
20High Murder																				0.000

z Matrix: Female students

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 Low Cannabis	0.000	1.150	0.813	1.728		1.385	1.385	1.385			1.728					1.728				
2.Low Receiving		0.000	0.674	0.966	0.105			0.674		1.728	1.728	1.728	1.728			1.728				
3 Low Burglaty			0.000	0.000	0.105	1.150	1.728	1.150			1.728				1.728	1.728				
4 Low Theft				0.000	0.209	0.431	0.966	0.966		1.728				1.728		1.150			1.728	
5.High Cannabis					0.000	0.431	0.431	0.209	0.674	0.966	0.813	1.385	1.385	1.150		1.728	1.728	1.385	1.728	
6.High Receiving						0.000	0.547	0.647	1.150	1.728	0.674	1.728				1.385				1.728
7 Low Arson							0.000	0.000	0.000	1.150	1.150	1.728				1.385				
8 Low Ag. Assault								0.000	0.146	1.150	0.674	0.966	1.385		1.385	1.385	1.960			
9 High Theft									0.000		0.318		1.728	1.728	1.728	0.966				
10High Burglary										0.000	0.431	1.150	1.150	0.813		0.674			1.385	
11Low Manslaughter											0.000	0.209	0.318	0.674	0.547		0.966	1.385	1.385	
12Low Kidnapping												0.000	0.318	0.966		0.318				
13High Arson													0.000	0.105	0.674	0.318	0.966	1.385	1.385	1.728
14High Ag. Assault														0.000	0.431	0.547	0.813	1.728	1.150	
15High Kidnapping															0.000	0.318	0.813		1.728	
16High Manslaughter																0.000	0.209	0.547	0.813	1.150
17Low Rape																	0.000		1.385	1.150
18High Rape																		0.000	0.318	1.385
19Low Murder																			0.000	
20High Murder																				0.000

z Matrix: Male students

APPENDIX D Cont.

The following successive difference matrices are those obtained in Experiment 2 for the total sample and female and male samples.

	c2-c1	c3-c2	c4-c3	c5-c4	c6-c5	c7-c6	c8-c7	c9-c8	c10-c9	c11-c10	c12-c11	c13-c12	c14-c13	c15-c14	c16-c15	c17-c16	c18-c17	c19-c18	c20-c19
1 L Can	1.185	-0.178	0.286	0.761			-0.199	0.490											
2 L Rec	1.185	0.856	0.075			0.549	1.174			-0.761	0.761	-0.158	-0.332						
3 L Bur	0.151	0.856	0.123	0.599	-0.650	0.934	0.412												
4 L The	0.363	0.808	0.123	0.323	0.108	0.461	0.201			-0.344				1.047					
5 H Rec			0.399	0.323	0.458	0.398	-0.070	0.778	0.491	-1.198	1.198								
6 H Can		0.101	-0.359	-0.027	0.458	0.274	0.268	0.181	0.285	-0.285	0.463	0.234		-0.126	0.761		-0.158	0.158	
7 L Ass	1.041	-0.285	0.114	0.037	0.582	0.274	0.274	-0.100	0.549	-0.449	0.582	0.151	0.087						
8 L Ars	-0.333	0.477	0.325	0.308	0.244	0.268	0.274	0.025	1.212	-0.515	0.697			0.960					
9 H The					0.841	0.549	0.148	0.025	2.054	-1.677			-1.124	1.124					
10H Bur				-0.291	1.047	0.285	-0.515	-0.817	2.054	0.377	1.042	-0.126	-0.752	0.644	0.869	-0.635	0.635	-0.491	
11L Man	0.270	-0.126		0.563	0.134	0.449	-0.449	0.345		0.377	0.025				-0.238	0.447	0.255	0.378	
12L Kid					0.869	0.329	-0.563			1.394	0.025	0.073	0.201	0.834					
13H Ars						0.412			0.761	1.268	-0.048	0.073	0.274	0.158	0.425	0.238	0.325		0.635
14H Man		-0.199	0.756	-0.178	-0.234	0.325		0.163	0.389				0.274	0.123	0.051	0.203	0.282	0.348	
15H Ass									0.869	0.463	-0.386	0.677	0.308	0.123	0.173	0.682	1.198	-0.491	0.491
16H Kid			-0.291			0.291				1.570			0.682		0.173	1.007	0.412	0.344	-0.199
17L Rap										0.488	-0.832	0.669	0.717	-0.479	-0.151	1.007			0.437
18H Rap						-0.158				0.869			0.760	-1.396	0.635			0.274	1.020
19L Mur													0.412	-0.557	-0.199	0.907	0.582	0.274	
20H Mur													1.047	-1.047	0.491	0.270			
SUM	3.862	2.310	1.551	2.418	3.857	5.640	0.955	1.090	8.664	1.577	3.527	1.592	2.555	1.408	2.990	4.126	3.531	0.795	2.384
N	7	9	10	10	11	15	12	9	9	15	11	8	13	13	11	9	8	8	5
MEAN	0.552	0.257	0.155	0.242	0.350	0.376	0.080	0.121	0.963	0.105	0.321	0.199	0.196	0.108	0.272	0.458	0.441	0.099	0.477
SCALE VALUES																			
S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
0.000	0.552	0.809	0.964	1.206	1.556	1.932	2.012	2.133	3.096	3.201	3.522	3.721	3.917	4.025	4.297	4.755	5.196	5.295	5.772

Total student sample

	c2-c1	c3-c2	c4-c3	c5-c4	c6-c5	c7-c6	c8-c7	c9-c8	c10-c9	c11-c10	c12-c11	c13-c12	c14-c13	c15-c14	c16-c15	c17-c16	c18-c17	c19-c18	c20-c19
1 L Can	1.221																		
2 L Rec	1.221	1.045			-0.465	1.016			0.280										
3 L Bur	0.176	1.045	0.431	-0.196		0.986			0.325					0.891					
4 H Rec			0.431	0.235		0.810													
5 L The		0.661		0.235	0.646	0.399				0.742				0.891					
6 H Can		0.196		-0.411	0.646	0.646	0.119	-0.434	0.315	0.399		0.402	-0.226	0.204	0.361				
7 L Ars	0.340	0.226	0.176		0.399	0.646	0.048	0.488	-0.104	1.356	-0.566								
8 H The						0.718	0.048	0.235	0.196	1.356									
9 L Ass		-0.131	-0.149		0.714	-0.204	0.301	0.235	0.048	0.384	0.334	-0.119	0.399						
10LMan	0.402	-0.176	0.176		0.399	0.215		0.384	0.048	0.331	-0.097				0.344	0.131	0.149	0.742	
11HBur					0.742	-0.742		1.356	0.100	0.331	1.787	-0.340	-1.016		1.356	-0.742	0.742		
12LKid					0.742	-0.176			0.530	-1.552	1.787	0.141	0.094			0.226			
13HArs									0.411	-1.212	1.306	0.141	0.235	0.810		0.176	0.226		
14HMan	-0.340	0.891	-0.325	0.325	-0.325	0.325		-0.149			0.196		0.235	0.048	0.187	0.301	0.230	0.456	-0.325
15HKid					0.361								0.998	0.048	0.048	1.174	-0.176	0.742	-0.566
16HAss										-1.022	0.566	0.176	0.810	0.187	0.048	0.896			
17LRap										-0.149	-0.402	0.226	0.686	-0.686	0.325	0.896			0.911
18HRap									0.742	-0.742			0.682	-0.280				0.235	0.986
19LMur													0.226	-0.566			0.301	0.235	
20HMur														-0.325	-0.566	0.340	0.226		
SUM	3.020	3.757	0.740	0.188	3.859	4.639	0.515	2.115	2.891	0.222	4.911	0.627	3.123	1.222	2.103	3.398	1.698	2.410	1.007
N	6	8	6	5	10	12	4	7	11	12	9	7	11	11	8	9	7	5	4
MEAN	0.503	0.470	0.123	0.037	0.386	0.386	0.129	0.302	0.263	0.018	0.546	0.089	0.284	0.111	0.262	0.377	0.242	0.482	0.252
SCALE VALUES																			
S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
0.000	0.503	0.973	1.096	1.133	1.519	1.905	2.034	2.336	2.599	2.617	3.163	3.252	3.536	3.647	3.909	4.286	4.528	5.010	5.262

Female students

	c2-c1	c3-c2	c4-c3	c5-c4	c6-c5	c7-c6	c8-c7	c9-c8	c10-c9	c11-c10	c12-c11	c13-c12	c14-c13	c15-c14	c16-c15	c17-c16	c18-c17	c19-c18	c20-c19
1 L Can	1.150	-0.337	0.915																
2 L Rec	1.150	0.674	0.292	-0.861															
3 L Bur	0.139	0.674		0.105	1.045	0.578	-0.578												
4 L Thef	0.762	0.966		0.209	0.222	0.535													
5 H Can			-0.104	0.209	0.431		-0.222	0.465	0.292	-0.153	0.572		-0.235				0.343	0.343	
6 H Rec			0.719		0.431	0.547	0.127	0.476	0.578	-1.054	1.054								
7 L Ars			0.762	0.535	-0.116	0.547			1.150		0.578								
8 L Ass	0.711	-0.476	0.184	0.757	-0.465	0.674		0.146	1.005	-0.476	0.292	0.419				0.575			
9H Thef					-0.476	1.150	-0.146	0.146							-0.762				
10H Bur				0.762	-0.762	0.578				0.431	0.719		-0.337						
11LMan					0.139	-0.476	0.476	0.356	-0.113	0.431	0.209	0.109	0.356	-0.127			0.419		
12LKid					-0.343		0.762			0.941	0.209	0.318	0.648						
13HArs								-0.343	0.578	0.832		0.318	0.105	0.569	-0.356	0.648	0.419		0.343
14HAss				0.578					0.915	0.139	-0.292	0.861	0.105	0.431	0.116	0.266	0.915	-0.578	
15HKid								-0.343					0.243	0.431	0.318	0.495			
16HMan			0.578	-0.578	0.343			0.419	0.292				-0.229	0.229	0.318	0.209	0.338	0.266	0.337
17LRap													0.153		0.604	0.209			-0.235
18HRap													-0.343					0.318	1.067
19LMur													0.235	-0.578	0.915	-0.572	1.067	0.318	
20HMur																	-0.235		
SUM	3.912	1.501	3.346	1.716	0.449	4.133	0.419	1.322	4.697	1.092	3.341	2.025	0.701	0.955	1.153	1.830	3.266	0.667	1.512
N	5	5	7	9	11	8	6	8	8	8	8	5	11	6	7	7	7	5	4
MEAN	0.782	0.300	0.478	0.191	0.040	0.517	0.070	0.165	0.587	0.136	0.418	0.405	0.064	0.159	0.165	0.261	0.368	0.133	0.378
SCALE VALUES																			
S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
0.000	0.782	1.082	1.560	1.751	1.791	2.308	2.078	2.543	3.130	3.266	3.684	4.089	4.153	4.312	4.477	4.738	5.106	5.239	5.617

Male students