

**AN EVALUATION OF A THREE DAY 'PREVENTION AND
MANAGEMENT OF AGGRESSION' TRAINING PROGRAMME FOR
STUDENT NURSES**

Bernard Frank Beech BSc(Hons), M.A.



**Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy**

March 2005

ABSTRACT

Workplace violence is a serious issue in health care with international surveys revealing disproportionate involvement in certain professional groups, for example, ambulance staff, nurse and student nurses, or settings, for example, mental health and learning disability, elderly care, and A&E units.

Staff training is widely advocated as the appropriate organisational response but there are relatively few published evaluations, and so much remains unknown about training effects or effectiveness. Many published studies are flawed by use of small samples, poor control of extraneous, organisational variables, absence of pre-test or follow-up data, limited range of measures, and weak statistical analysis.

This study examined an existing training programme for student nurses whilst attempting to avoid the limitations identified above.

The effects of training on a number of learning domains, for example, knowledge, self confidence, beliefs and attitudes, and self –assessed skills was investigated using a repeated measures, variable baseline research design, in conjunction with a model of learning.

The likelihood of student nurses involvement in violent incidents, and the power/ ease of use of different change evaluation methods were also investigated.

Repeated administration of a purpose –designed questionnaire at four time points to three consecutive cohorts of student nurses [N=243] provided information about pre-training stability, possible changes on

immediate training completion, and at three-months follow-up, after two clinical placements.

Statistical analysis revealed the Unit to have generally desirable effects on learning domains that were still detectable at three-month follow-up. It also highlighted differential involvement in violent incidents based on placement type, and important differences between evaluation methods in terms of ease of use.

ORIGINALITY AND PUBLICATIONS

All the data presented in this document was obtained during the period of my association as a part-time Post-graduate Student with the Institute of Work, Health and Organisations, University of Nottingham.

At the same time, an attempt to pursue parallel publication of aspects of the study has resulted in the following publications and submitted manuscripts to date:-

- An overview of the study and preliminary analysis of much of the data [aspects of Chapters 1, 2, 3 and 5] has been published as Beech B., & Leather P. (2003) Evaluating a management of aggression unit for student nurses. *Journal of Advanced Nursing* 44(6), 603-612.
- A review of violence training and management issues in mental health nursing [aspects of Chapters 1] has been published as Beech B., & Bowyer D. (2004) Management of aggression and

violence in mental health settings. *Mental Health Practice* 7(7), 31-37.

- A review of workplace violence and training issues in health care [aspects of Chapters 1 and 2] has been accepted for publication, subject to minor amendments, in *Aggression and Violent Behavior*.

ACKNOWLEDGEMENTS

Although this thesis bears only one name, a number of people have played a part in its development over the last five years, and I would like now to publicly acknowledge their contributions.

I would like to express my thanks to staff and fellow PhD students at the Institute of Work, Health and Organisations, University of Nottingham for their advice, suggestions and steady support. In particular I would like to thank Dr Diane Beale for her assistance in creating the questionnaire and, of course, Dr Phil Leather for all aspects of his enthusiastic supervision, and kind words about my writing.

I would like to thank my employer, Keele University, and Head of the School of Nursing and Midwifery, Keele University for providing financial assistance and personal development time. I thank work colleagues for their encouragement, particularly those who have trodden a similar path and so knew the demands. I especially appreciated the contributions of Dr Sue Read and thank her for her efforts in connection with establishing inter-rater reliability.

Most of all, I would like to thank my family – both sets of parents, my wife Clare, and children David, Simon, Rebecca and Philip - for their unflinching encouragement, and certainty that I would complete the task.

I hope that completion of this work and the doctoral process, and any good that may come of it, compensates for some of the sacrifices they have made over the last five years or so.

TABLE OF CONTENTS	
ABSTRACT	ii
ORIGINALITY AND PUBLICATIONS	iii
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vi
LIST OF TABLES	xiv
LIST OF FIGURES	xvii
LIST OF BOXES	xix
CHAPTER 1 – WORKPLACE VIOLENCE, EVALUATED	1
TRAINING AND HYPOTHESES	
1.1 WORKPLACE VIOLENCE	2
1.1.1 Definitions	5
1.1.2 Typology	9
1.1.3 Cause of Violence in Health Care Settings: Attribution and Blame	10
1.1.4 Extent of the Problem: Incidence and Costs	12
1.1.4.1 Incidence	12
1.1.4.2 Costs	22
1.2 MANAGEMENT OF AGGRESSION AND VIOLENCE	
TRAINING	24
1.2.1 Training: Its Advocacy and Prevalence	24
1.2.2 Content of Training	33
1.3 REVIEW OF TRAINING COURSES FROM PUBLISHED	41
LITERATURE	

1.3.1 Summary of Literature Review of Aggression	62
Training Course Evaluations	
1.3.1.1 Care Area	62
1.3.1.2 Course Duration	62
1.3.1.3 Country of Origin	63
1.3.1.4 Content	64
1.3.1.5 Research Design	65
1.3.1.6 Sampling	65
1.3.1.7 Statistics	66
1.3.1.8 Dependent Variables	66
1.3.1.9 Outcomes	67
1.3.2 Conclusions from Review of Aggression Training	69
Course Evaluation Studies	
1.4 RESEARCH QUESTIONS AND HYPOTHESES	71
1.4.1 Research Questions	71
1.4.2 Research Hypotheses	72
CHAPTER 2 - LITERATURE ON EVALUATION OF TRAINING	75
2.1 DEFINITION OF EVALUATION	75
2.2 HISTORICAL DEVELOPMENT AND EVOLUTION OF EVALUATION	81
2.3 EVALUATION: TYPES AND PURPOSE	86
2.4 EVALUATION OF TRAINING COURSES	99
2.4.1 Explaining the Infrequency of Training Evaluation	115
2.5 SUMMARY OF CHAPTER	119

CHAPTER 3 - FOCUS OF THE STUDY - A 3-DAY UNIT FOR	120
STUDENT NURSES IN THE PREVENTION AND	
MANAGEMENT OF AGGRESSION IN HEALTH CARE	
SETTINGS	
3.1 THE DEVELOPMENT OF NURSE EDUCATION	120
3.1.1 Pre-Registration Nursing - Early Years	120
3.1.2 Project 2000	121
3.2 DIPLOMA IN NURSING STUDIES COURSE 1996-2001	124
3.3 TRAINING UNIT: CONCEPTUALIZATION AND	127
DEVELOPMENT	
3.4 TRIMESTER THREE	130
3.5 TRAINING UNIT	132
3.5.1 Structure	132
3.5.2 Goals And Objectives	134
3.5.3 Content	136
3.5.4 Teaching Methods	137
CHAPTER 4 – RESEARCH METHODOLOGY	143
4.1 RESEARCH DESIGN	143
4.1.1 Quasi-Experimentation	144
4.1.2 Time Series Designs	146
4.2 DATA COLLECTION METHODS	149
4.2.1 Questionnaires: Uses and Design	149
4.2.2 Different Types of Question	152

4.2.2.1 Likert scales	152
4.2.2.2 Visual Analogue Scales	153
4.2.2.3 Open Response	155
4.2.2.4 Demographic Data	157
4.3 OTHER ISSUES - BIAS, ERRORS.	158
4.3.1 Ethics and Possible Coercion Problems	158
4.3.2 Response-Shift Bias	160
4.3.3 Internal or External Evaluators?	161
4.4 DESIGN OF STUDY	163
4.4.1 Choice of Methodology and Data Collection Method	163
4.4.2 Overall Questionnaire Design Priorities	166
4.4.3 Process of Creating the Questionnaire	168
4.5 QUESTIONNAIRE – DETAILS OF TYPES OF QUESTIONS CREATED	170
4.5.1 Demographic Details	170
4.5.2 Knowledge	172
4.5.3 Attribution of Blame	179
4.5.4 Attitudes, Competence and Confidence	180
4.6 RESEARCH PROCESS	183
4.6.1 Sample	184
4.6.2 Pattern of Questionnaire Administration	185
4.6.3 Obtaining Access Approval	191
4.6.4 Identify Opportunities to Access Student Nurses	192
4.6.5 Devising Identification Number Coding System	193
4.6.6 Gain Access to Student Nurse Groups, Present	194

Details of Study and Encourage Participation	
4.7 SUMMARY STATEMENT	197
CHAPTER 5 – RESULTS	198
5.1 INTRODUCTION	198
5.1.1. Measures of Reliability	201
5.2 SAMPLE	206
5.2.1 Size And Response Rate	206
5.2.2 Homogeneity Of The Three Cohorts In The Sample	207
5.3 RESULTS - DEMOGRAPHIC DATA	211
5.3.1 Age, Gender, Destined Branch, Previous Training	211
5.3.2 Experience of Violence	213
5.3.3 Number of Incidents	216
5.4 RESULTS - SCENARIOS	218
5.4.1 Number of Risk Factors	218
5.4.2 Attribution of Blame	222
5.4.3 Test-Retest Reliability	224
5.4.4 Bi-variate Tests	225
5.5 RESULTS - 24 LIKERT STATEMENTS	232
5.5.1 Initial Review	233
5.5.2 Factor Analysis	235
5.5.3 Protocol Used for Principal Component Analysis	238
5.5.4 Principal Components Analysis Solution for 24	241
Likert Statements Combined for All Cohorts on All	
Occasions.	

5.5.5 Factor Scores	248
5.5.5.1 Factor 1 [Confidence in maintaining personal safety]	248
5.5.5.2 Factor 2 [Prediction and prevention]	249
5.5.5.3 Factor 3 [Personal practical ability]	249
5.5.5.4 Factor 4 [Self respect and staff rights]	251
5.5.5.5 Factor 5 [Provocative behaviour]	252
5.5.6 Graphic Representation	254
5.5.7 Measures of Reliability	256
5.5.7.1 Test-Retest Reliability	256
5.5.7.2 Internal Consistency	256
5.5.8 Bi-variate Analysis	262
5.6 SECONDARY ANALYSIS OF SCENARIO RISK FACTOR SCORES	266
5.7 SUMMARY OF RESULTS	270
CHAPTER 6 – DISCUSSION	273
6.1 RE-PRESENTATION OF STUDY RESEARCH QUESTIONS AND RESEARCH HYPOTHESES	274
6.2 EXPLANATION OF FINDINGS	276
6.2.1 The Sample: Size, Response Rate, Homogeneity, Generalisability	276
6.2.2 Extent of Student Nurse Experience of Violence	279
6.2.3 Number of Risk Factors Identified in the Scenarios	282
6.2.4 Attribution of Blame	286

6.2.5 Likert Statements	288
6.2.6 Factor Scores	292
6.2.6.1 Factor One [Confidence in maintaining personal safety]	292
6.2.6.2 Factor Two [Prediction and prevention]	294
6.2.6.3 Factor Three [Personal Practical Ability]	296
6.2.6.4 Factor Four [Self-Respect and Staff Rights]	300
6.2.6.5 Factor Five [Provocative Approach]	302
6.2.7 Reliability of Factor Scores	303
6.2.8 Secondary Analysis of Identified Risk Factors	305
6.3 SUMMARY OF FINDINGS IN RELATION TO RESEARCH HYPOTHESES	307
6.4 UTILITY	315
6.4.1 Likert-type questions	316
6.4.2 Visual Analogue Scale	317
6.4.3 Open Response Questions	318
6.5 CRITICAL EVALUATION	321
6.5.1 Issues With Questionnaire Design	324
6.5.2 Questionnaire Administration	326
6.5.3 Findings	327
6.5.4 Interaction of Factors?	330
6.5.5 Additional Measures of Change	332
6.5.6 Reliability	333
6.5.7 Follow-up	334

6.6 IMPLICATIONS FOR PRACTICE	336
6.6.1 Relevance of Aggression Management Training	338
6.6.2 Accuracy of the Level of Incident Reporting	339
6.6.3 Further Training Requirements - Update And Refresher Training	340
6.6.4 Usefulness and Importance of Evaluation Research	341
6.6.5 Low-Level, Soft-Skills Training	343
6.6.6 Utility of Measures	346
6.7 CONCLUDING REMARKS	347
REFERENCES	350
APPENDICES	376
Appendix One - Draft Strategy For Training In Aggression And Violence For Pre-Registration Students (Summary)	377
Appendix Two - Revised violence questionnaire [different layout because of altered page margins]	379
Appendix Three - Completed Request for Access to Students Form	383
Appendix Four - – 100% Stacked for percentage responses to 24 statements at four time points	387

LIST OF TABLES	18
Table 1.1: Reported incidence of violent incidents classified by occupational group and severity [rates per 1000 in the year preceding survey] (from MacKay 1994)	47
Table 1.2: Summary of training courses concerning psychiatric settings	56
Table 1.3: Summary of training courses concerning learning disability, elderly care and mixed settings	206
Table 5.1: Sample size & response rate for each cohort of student nurses	208
Table 5.2: Summaries for each group /cohort on demographic variables N, %	210
Table 5.3: Statistical analysis of homogeneity of the three cohorts on significant demographic variables	212
Table 5.4: Student nurse demographics	214
Table 5.5: Student nurses involvement in verbal or physical violent incidents during three trimesters	215
Table 5.6: Crosstabs breakdown of student nurse involvement in verbal and physical violence in trimester one by placement type	217
Table 5.7: Numbers of incidents [verbal and physical] which students were involved in during placements	219
Table 5.8 : Inter-rater reliability results for number of risk factors identified in scenarios	

LIST OF TABLES -CONTINUED	221
Table 5.9: Number of risk factors identified at four time points for scenarios – mean scores, min, max, range, and paired sample t-test results	223
Table 5.10: Attribution of blame [length of VAS line] identified at four time points for scenarios – mean scores and paired sample t-test results	225
Table 5.11: Test-retest reliability - correlation tests for scenario variables [number of risk factors identified and attribution of blame] at t1 and t2	228
Table 5.12: Selected results for bi-variate analysis of demographic variables and scenario variables- means	231
Table 5.13: Selected results for bi-variate analysis of demographic variables and scenario variables- means 'number of risk factors identified' and 'attribution of blame' for scenarios	242
Table 5.14: Communalities before and after extraction	243
Table 5.15: Initial and rotated eigenvalues, and total variance explained	
Table 5.16:: Rotated component matrix	245
Table 5.17: Questionnaire statements associated with each factor name	247
Table 5.18: Factor scores identified at four time points – mean scores and paired sample t-test results	253
Table 5.19: Test-retest reliability for five factor scores at time points t1 and t2	256

LIST OF TABLES -CONTINUED	258
Table 5.20: Cronbach alpha α for factor scores calculated at time points t1-t4 and mean alpha score	259
Table 5.21: Mean inter-item correlation for factor scores calculated at time points t1-t4	260
Table 5.22: Calculation of cronbach α for factor 3 if item removed at time points t1-t4	261
Table 5.23: Effects on cronbach alpha and mean inter-item correlation of removing item 15 from factor 3 at time points t1-t4	265
Table 5.24: Selected results for bivariate analysis of demographic variables and factor scores	268
Table 5.25: Number of student nurses mentioning non-assailant risk factors for scenarios at four time points	269
Table 5.26: Number of non-assailant risk factors identified at time point t1-t4 - means and paired sample t-test	

LIST OF FIGURES

Figure 1.1: The integrated organisational approach	28
Figure 1.2: Levels of evaluation	43
Figure 1.3: Possible measures of change for violence management training	44
Figure 2.1: Levels of evaluation: a combined framework	114
Figure 3.1: Overview of diploma course structure and placements	126
Figure 3.2: Overview of the structure of trimester three and content of study weeks	131
Figure 3.3: Accommodating increased student numbers in weeks 1-3 of trimester three	132
Figure 3.4: Delivery of the unit to a cohort on different days over a two-week period	133
Figure 3.5: Unit learning outcomes in four learning domains	136
Figure 3.6: Summarized unit content	138
Figure 3.7: Outline timetable for the unit	141
Figure 3.8: Chronology of main developments of the aggression management and prevention unit and its evaluation	142
Figure 4.1: The study design applied to a single student cohort showing trimester structure, placement types and data collection points	188
Figure 4.2: Time scale of data collection for the three cohorts of student nurses included in the study	190
Figure 5.1: Summary of data collection schedule	205

LIST OF FIGURES CONTINUED

Figure 5.2 Scree plot of final solution	244
Figure 5.3: Idealised bar charts for effects of training	254
Figure 5.4: Bar charts of mean scores for factors 1-5 at time points t1, t2, t3, & t4	255

LIST OF BOXES

Box 2.1: Key initial questions and type of evaluation to be performed [adapted from Robson (2000)]	89
Box 2.2: Types of evaluation, (alternative names), associated purpose and activity (Evaluation Research Society 1980:3-4)	91
Box 2.3: The main features of the formative /summative evaluation dichotomy	96
Box 2.4: Stages of program development and related evaluation purpose [adapted from Pancer and Westhues (1989)]	98
Box 2.5: Classification of learning categories, constructs, foci of measurement and potential training evaluation methods [adapted from Kraiger <i>et al</i> 1993]	111
Box 4.1: The advantages and disadvantages of using internal or external evaluators [from Clarke (1999)]	162
Box 4.2: Outline of interrelationship between learning domains, unit learning outcomes and questionnaire sections	168
Box 4.3: Acceptable risk factors in different categories - scenario 1 [young man A&E] [from 250 questionnaires approx]	177
Box 4.4: Acceptable risk factors in different categories - scenario 2 [elderly woman] [from 250 questionnaires approximately]	178
Box 4.5: Relationship between training evaluation model domains, learning outcomes, research hypotheses and the questions on questionnaire	182
Box 4.6: Script for briefing students about the study and gaining their consent for participation	196

CHAPTER 1 – WORKPLACE VIOLENCE, EVALUATED TRAINING AND HYPOTHESES

The problem of violence in the Health Service is associated with its main purpose – the provision of a service to the public. This complex public comprises patients, their relatives and visitors, and others with health service needs on or off site. That some patients may be predisposed towards violence adds a special dimension to the task of its control or prevention, which is an integral part of the management of the service.

Health Services Advisory Committee (1987:1)

In essence, this thesis is concerned with workplace violence and the evaluation of training in the prevention and management of workplace violence. In order to make sense of the choice of the research area it is firstly necessary to review the phenomenon of workplace violence and relate this specifically to staff working in health care settings. It is also necessary to review previously published reports of aggression and violence management training in order to identify gaps in the field and, in so doing, justify the choice of research topic.

At the same time it is essential to establish limits on the range of material to be discussed. Therefore this chapter will *not* consider

theories and models explaining violence in society and will *not* review workplace violence outside of health care settings in any great detail. Furthermore, it will *not* include all aspects of health-care workplace violence; for example, it will not consider bullying or harassment by colleagues, since these aspects tend to be approached separately within a Human Resource framework.

1.1 WORKPLACE VIOLENCE

The development of workplace violence as an international issue has occurred over the last thirty years or so (Bowie 2000). It is truly an international problem (Chappell and Di Martino 2000), as indicated by the literature from Australia (Jackson *et al* 2002, Mayhew & Chappell 2002), Canada (Hesketh *et al* 2003), United States of America (Flannery 1996, Smith-Pittman & McKoy 1999) and Sweden (Arnetz 1998, Nolan *et al* 2001).

Yet, at the same time, one must stay alert to societal differences when identifying the problem and its solution. For example, Whittington (1994) cites an American nurse's advice that "security personnel must remove their guns before entering an inpatient unit", while a more recent National Institute for Occupational Safety and Health (NIOSH) publication reported that in a Detroit hospital a metal detector screening system prevented the entry of 33 hand guns, 1,334 knives and 97 mace-type sprays during a six-month period (NIOSH 2002). Thankfully,

these security measures and advice are still inappropriate in this country.

The concept of workplace violence has overcome the initial stages of 'denial', then lack of data [although under-reporting remains a problem], and now endeavours to agree what is and is not to be included under the definition. Given the evolutionary nature of 'work', it is still difficult to agree what is the 'workplace' (Budd 1999), let alone what is 'violence'. Researchers have investigated the area at the behest of managers and trade unions and made strides in delineating the concept and developing typologies, begun to identify the extent of the problem in various types of work setting and suggested managerial responses to the problem.

Bowie (2000) and Chappell & Di Martino (2000) suggest that early studies of workplace violence were intent on identifying the defining characteristics of the assailant or attacker and generated lists of common features, including feeling aggrieved, irritated or frustrated, the victim of prejudice, being in uncomfortable conditions or having mental instability. More recent studies have identified characteristics of the organisation, for example, environment [crowding, noise privacy] (Cox & Leather 1994, Shepherd & Lavender 1999), or staff team [negative staff attitudes] (Poyner and Warne 1986) as being implicated in determining the rate of violence. These factors suggest that an integrated and multi-factorial model of violence is required if the phenomenon is to be

understood (Poyner and Warne 1986, Cox & Leather 1994, Leather *et al* 1998a, Beale *et al* 1998, Chappell and Di Martino 2000).

These more advanced frameworks offer a systematic problem-solving or 'control cycle' approach (Dickson *et al* 1994) to the problem of workplace violence, in exactly the same manner as any other workplace hazard. They consider the responses to violence in a number of related areas, for example, the effects on assailant and victim, activity before, during and after incidents, at the level of the individual, team and organisation and in relation to over-arching policies, specific procedures to deal with different incidents and the endorsement of best practice and professional behaviours (Beale *et al* 1998).

Chappell & Di Martino (2000:51) assert that the temptation to ascribe violence to a single source should be avoided and instead, "a proper understanding of violence (and ultimately of the means for its control) requires an understanding of the variety and complexity of contributing factors". They further developed an interactive model previously offered by Poyner & Warne (1986) that covered particular aspects of the assailant, the staff, the environment, and the interaction, adding sections on outcomes for victims and organisations. In light of this activity, much is now known and some of this will now be summarized.

1.1.1 Definitions

Several authors suggest that there is still no consensus about the definition of workplace violence (Beale *et al* 1998, Budd 1999, Bowie 2000, Rogers & Chappell 2003), while the Royal College of Nursing (RCN) contend “there is no single definition that is universally applicable to all workplaces, circumstances and occupational groups” (RCN 1998:3). Difficulties exist over the breadth of violence to be included, not merely physical assault or verbal threats, but also the inclusion, or not, of bullying, sexual harassment etc.

The balance of this distinction is important for a number of practical reasons. Firstly, Budd (1999) suggested that victims of threats could be more seriously emotionally affected than victims of assaults. Second, drawing the definitions of workplace violence too narrowly excludes all but the rarest, most serious offences and creates a concept that, thankfully, very few employees can associate with. Alternatively, creating too broad a definition blurs the distinction between workplace violence and ‘general’ violence in society (Perone 1999). Finally, with regard to violence versus bullying /harassment, many organisations, including health services would treat these issues separately, having completely distinct policies for violence and bullying [including harassment].

The limits of the term ‘workplace’ are also debated, and this has implications for example, for those workers who ‘work from home’, or

those that are attacked whilst travelling to, or from, or between work sites or in a client's home (Budd 1999, Bowie 2000). In addition, some researchers are primarily interested in violence from 'the public' and exclude violence perpetrated by work colleagues.

Furthermore, within the health care sector most studies that examine *patient* dangerousness and its effects do not tend to use the phrase 'workplace violence', nor position themselves under this category. Instead, workplace violence is reserved for attacks by strangers or colleagues with a grudge. Hatch-Maillette & Scalora (2002:279) suggest that

studies on staff assaults are often found in the nursing or risk assessment literatures pertaining to custodial care of patients and inmates, whereas workplace violence studies are found in literatures focusing on a broader scope of occupations and on staff-on-staff (or "coworker") assault.

Love & Hunter (1996:30) agree and suggest that the recent shift from viewing violence, especially in psychiatry, as a clinical problem to "framing violence as an occupational health concern represents a major paradigm shift" since it brings in to play "powerful external incentives" in the form of Health and Safety legislation. Obviously, all of these issues are important since each will alter the calculation of number and type of incidents reported and recorded in various settings.

Examples of definitions that highlight these issues are as follows:

Jenkins (1996) defined violence at work as “crimes of violence that occur in the workplace or while the victim is at work or on duty”. While this emphasises the broad areas that constitute the ‘workplace’ it seems to emphasise more serious incidents that could be legally described as crimes and will thus exclude some forms of harassment, verbal or emotional abuse and bullying.

In a British study of reported crime Budd (1999:2) defines violence at work as “all assaults or threats which occurred while the victim was working and were perpetrated by members of the public”. Budd (1999) proceeds to explicitly exclude violence from colleagues, arguing that it is likely to have a different nature and pattern to violence involving the public.

In relation to health care settings, the definition adopted by the Health Services Advisory Committee (HSAC) (1997:2) was

any incident in which a person working in the healthcare sector is verbally abused, threatened or assaulted by a patient or member of the public in circumstances relating to his or her employment.

This definition emphasises a “new profile” (Beale *et al* 1998) of a range of behaviours, threats as well as physical damage, but again fails to include the possibility of the assailant being a present or former

colleague. In this day and age any comprehensive definition should include the possibility of violence from 'colleagues'. Farrell (1997, 1999, 2001) has shown that student nurses expect aggression from patients, relatives and even doctors but find aggression from other nurses to be least acceptable and most distressing.

In a study of violence to staff working in the community Beale *et al* (1998:1) adopted a widely used definition of work-related violence which has been accepted by the European Commission DG-V and adapted from Wynne *et al* (1997), namely,

incidents where [staff] are abused, threatened or assaulted in circumstances related to their work, involving an explicit or implicit challenge to their safety, well-being or health".

In a recent 'zero tolerance of violence' campaign, to be discussed later in this chapter (DoH 1999a), the British Government asserted that defining work-related violence is not subjective and proceeded to adopt this same definition. Once again this definition only implicitly includes the possibility of aggression from colleagues.

The RCN (1998:3). defined workplace violence as

any incident in which a health professional experiences abuse, threat, fear or the application of force arising out of the course of their work, whether or not they are on duty.

Here they cover a range of behaviours and also acknowledge that a person can be a victim even in their off-duty time, for example, being abused by an ex-client whilst out shopping.

1.1.2 Typology

Several authors (Beale *et al* 1998, Bowie 2000, Chappell and Di Martino 2000, Hoel *et al* 2001, Mayhew & Chappell 2002) cite the framework devised by the Californian Division of Occupational Health and Safety (1995), which identifies three types of violence. Type 1 involves external perpetrators who have no legitimate relationship with the organisation [for example, robbery or road rage], Type 2 involves aggressive acts by consumers or clients of a service or business [for example, patient or relative aggression], and Type 3 involves aggressive or violent acts by current or former employees, or others with an employment-based relationship with an organisation [for example, bullying]. Mayhew and Chappell (2002) emphasise that while all three types can occur on the same worksite [and to the same unfortunate employee], the perpetrators will have different characteristics and, furthermore, the preventative strategies will be markedly different, as will its control and management (Leather *et al* 1998a).

Baron and Neuman (1998) have studied this Type 3 workplace violence from colleagues and suggested that physical violence is the tip of a workplace aggression iceberg characterised by more subtle, covert

forms of harm-doing which has been exacerbated by recent organisational changes, such as down-sizing and workforce diversity. Many authorities suggest that Type 2 incidents are the most common in health care (International Council of Nurses 1999) but Farrell (1997, 1999) collected evidence suggesting that nurse to nurse aggression was the most common and distressing, accounting for approximately 25% of all incidents. Hoel *et al* (2001) cite a European Foundation (2000) survey of 21,500 face-to-face interviews with workers from 16 European Union states which revealed that 4% of employees are subjected to physical violence from individuals not belonging to the organisation (Type 1 and Type 2) whilst 2% experience violence from their colleagues (Type 3).

1.1.3 Cause of Violence in Health Care Settings: Attribution and Blame

A number of recent studies of aggression and violence in mental health settings have located causes under three conceptual headings, these being psychological /internal to the aggressor; external/environmental; and interactional/interpersonal (Sheridan *et al* 1990, Finnema *et al* 1994, Duxbury & Whittington 2005). Internal causes of aggression are frequently viewed as relating to either personality or mental state. The former tend to be seen by staff as 'controllable' whilst the latter are viewed as 'uncontrollable'. The distinction is crucial since there is evidence that it affects the subsequent treatment or management options considered in any particular case. If the aggression is adjudged

to be internal-controllable then 'punishment' in the form of seclusion or restraint may ensue whilst if the aggression is thought to be internal – uncontrollable then care treatment in the form of medication is more likely (Crichton *et al* 1998, Leggett & Silvester 2003, Ilkiw-Lavelle & Greyner 2003).

In their discourse analysis Benson *et al* (2003) detected a further concern in making an attribution of causation that was described as exoneration from blame. Staff were found to have extricated themselves from possible blame or claim of incompetence by blaming incidents on the internal-controllable personality aspects of the patient aggressor that they were not trained to identify or manage. This tendency has also been recognised by Lanza & Kayne (1995) wherein staff attempted to present themselves and their actions in a good light by giving ego-enhancing explanations for incidents.

Studies of patient's views have tended to demonstrate a different attribution of causation. These studies reveal a tendency for patients to de-emphasise internal factors and, instead highlight the importance of environmental, or more likely, interpersonal or staff communication deficits for the cause of a violent incident (Bensley *et al* 1995, Ilkiw-Lavelle & Greyner 2003, Duxbury & Whittington 2005),

Hence the attribution of the cause of a violent incident and possible ensuing blame is an important area for staff [in terms of perceived

competence] and patients [in terms of treatment or management regime]. Later in this study it will be seen that an attempt was made to gain an understanding of student nurses' attribution of blame for causation of a violent incident in two scenarios. The conceptual categories and choices are also closely related to the previously mentioned Poyner and Warne model wherein blame is attributed to aggressor, environment, staff or task. The connections with the forgoing discussion are self-evident. Aggressor could refer to internal controllable or internal –uncontrollable factors; environment with external/environmental factors; task with interactional factors; and staff with environmental or interactional factors depending on whether aspects of regime or communication style were being emphasised.

1.1.4 Extent of the Problem: Incidence and Costs

1.1.4.1 Incidence

Researchers caution about the interpretation of figures on the incidence of workplace violence for many valid reasons. In addition to previously mentioned definitional issues, there are also legal obligations to report some events and psychological pressures not to report others, including concern for the perpetrator in the case of client-initiated violence (Mayhew & Chappell 2002), willingness to accept low-level aggression as part of the job, avoidance of form-filling as a way of managing the workload and self-conscious awareness about the monitoring and interpretation of incident numbers by the employer (Lion 1981, Dickson

et al 1993, Rosenthal *et al* 1992, Cox & Leather 1994, Beale *et al* 1998).

Chappell and Di Martino (2000:24) assert that the expansion of liability and responsibility for the maintenance of violence-free work environments has prompted efforts to measure incidence and prevalence, and anticipate the closer of any remaining gaps in knowledge. It is widely accepted that the numbers reported are a gross under-estimate of the total number of incidents (Dickson *et al* 1993, Department of Health 1999a, Smith-Pittman & McKoy 1999, Hesketh *et al* 2003) and, for example, in the health care sector a figure of 20% of incidents being reported is often cited (Lion *et al* 1981), literally the tip of the iceberg.

Furthermore it is also generally accepted that the number of incidents of work-place violence is increasing (Dickson *et al* 1993, Flannery 1996, HSAC 1997, Industrial Relations Services [IRS] 1998, RCN 1998, Budd 1999, IRS 2000, Bowie 2000), or at least perceived to be increasing (Cox & Leather 1994) although other effects complicate any interpretation of this trend. These effects include, for example, the possible reduction in tolerance and growing readiness of staff to report incidents following training and awareness raising (Whittington 1994), and concern over the perceived increase of violence in society more generally (Whittington 1997).

Fox *et al* (2002:1) include the following possible explanations for increased workplace violence,

- Higher expectations of service users
- Willingness to complain and demand
- Smoking bans and their enforcement
- Increased stress of complex modern life
- Steady increase in violent crime
- Changing values and beliefs surrounding acceptable behaviour
- Popularisation of violence through television and media
- Negative role models i.e. football and rock stars

There are key staff groups that have an increased risk, including workers in public contact service industries [health, education, personnel], lone workers in community settings [night workers, taxi drivers] staff who handle cash or drugs [convenience stores, pharmacies, petrol stations], security staff and those involved with the legal system [police, lawyers, probation workers] (Flannery 1996, Bowie 2000, Chappell and Di Martino 2000, Mayhew & Chappell 2002).

According to Budd's analysis of the British Crime Survey (1999:15), the police and security staff are at the greatest risk followed by nurses and other health professionals. These relative positions remain the case in the most recent analysis of British Crime Survey data (2000 cited in National Audit Office [NAO] 2003). Budd (1999:14) suggested "those in

the most at risk occupation are 153 more times likely to be a victim of violence at work than those in the least at risk occupation”.

Budd (1999:18) estimated that nurses had “the second highest risk of being assaulted at 5%, that is four times the national average risk” and twice the national average risk of being verbally threatened or intimidated. Furthermore, Budd (1999:25) suggested that when other factors [age and sex, hours worked, occupational status] are controlled, for high-risk groups, “there is something intrinsic in the nature of the work itself which results in high risks”. It may well be that this is what is being referred to in the quotation that opens this chapter. Interacting with members of the public who are in pain, frustrated, receiving bad news that confirms their worst fears, who may have poor impulse or anger control as part of their problem, who are in hospital against their wishes etc. is intrinsically dangerous, and so “it is of no surprise to find that nurses are at particular risk” (RCN 1998:4).

More detailed information is obtained from surveys conducted within the health service sector. Wells & Bowers (2002) provide a useful summary, with an emphasis on the experience of general or ‘adult’ nurses, based on analysis of a literature review. After reviewing the incomplete picture offered by a number of diverse trades union, HSE, Department of Health and small-scale local surveys, the authors suggested that

- nurses appear to be at significantly higher risk than workers generally [more than x 4] or other health carers [more than x 3.5].
- general nurses have excess risk, although less than nurses working in mental health and learning disabilities settings
- notifiable injuries resulting from assault submitted by hospitals are excessive [20% of total notified to HSE]
- conclusions on prevalence of assault or abuse etc. are difficult because of multiple methodological difficulties and limitations (Wells & Bowers 2002).

Of relevance to this study are the authors conclusions, which stressed a need for a greater efforts to manage violence, and mandatory pre-registration training in dealing with violence in all schools of nursing (Wells & Bowers 2002). The clear implication is that this type of training is not currently available for all pre-registration nurses and, therefore, raises the level of interest in this evaluation of such a training course.

Some British surveys, chosen because of their scale or frequency of citation, will now be summarised. Probably the most cited study (Cox and Leather 1994, MacKay 1994, Beale *et al* 1998), now a little dated, is the survey of 5000 health service staff from all clinical areas and disciplines (Health Services Advisory Committee [HSAC] 1987). The 3000 respondents showed that violence was a feature of all health care staff working in all health settings.

The survey collected data on four sorts of violence – verbal threat, minor and major injuries and incidents involving a weapon – and found that overall 1 in 200 had suffered a major injury requiring medical assistance during the last year while 11% had received a minor injury, 4.6% had been threatened with a weapon and 17.5% had received a verbal threat. Higher risk staff groups and settings included ambulance staff and nurses in training. Student nurses were placed equal second in major injuries, first in minor injuries, second in incidents involving a weapon, and second in verbal threats. In light of the emphasis to be placed on student nurses in later chapters of this thesis, a breakdown of the reported incidence of violent incidents classified by occupational group and severity is provided as Table 1.1 (from MacKay 1994).

Other staff with higher risks were those working in Accident and Emergency, Mental Health [1 in 4 minor injury], Learning Disability and Elderly Care settings [1 in 5 minor injury]. At that time only 12% of respondents reported having received any form of training in managing violence, mostly during basic training. A more recent update from HSAC (1997) cited a survey by the NAO that reported 14% of recorded accidents in the NHS involved physical assault making it the third most common type of accident involving staff (NAO 1996).

Table 1.1: Reported incidence of violent incidents classified by occupational group and severity [rates per 1000 in the year preceding survey] (from MacKay 1994)				
	Major	Minor	Weapon	Threat
Hospital doctors	5	59	30	193
GPs	5	5	50	249
Student nurse	16	364	136	402
Staff nurse	0	202	73	337
Charge nurse	16	172	86	242
Ambulance staff	17	174	174	421
Catering	0	11	11	11
Laundry	0	20	0	59
Domestic	6	30	30	42
Porters	0	81	32	210

A study of 105 NHS Trusts by Industrial Relations Services [IRS] revealed that the majority of NHS employers (52%) considered workplace violence a major problem and reported that health workers are four times more likely to suffer work-related violence than the general public, with approximately 1 in 10 suffering a violent incident in the previous year, based on reported incidents (IRS 1998:3). This equates to an overall ratio of 850 incidents per 10,000 workers, although in some settings, including community and mental health settings, a much higher incidence of violence was found, around 1 in 3 being victims of work-related violence. Incidents causing major injuries accounted for around 10% of cases overall and around 20% of

community and mental health cases. Minor injuries accounted for approximately 2/3 of all incidents while weapons were used in 5% of incidents. It is apparently the case that reporting of verbal threats was much less likely to occur.

Reassuringly, as far as faith in survey methodologies goes, this rate figure – 850 per 10,000 - is almost identical to that reported by the Department of Health –7 per 1000 staff per month [840 per 10000 per year] for all NHS trusts - in the annual survey of sickness, absence and violence by the NHS Executive over a contemporaneous period (Department of Health 1999b). This survey of reported incidents of violence from 364 Trusts cited figures of 13 violent incidents per month per trust overall, with 64% of these being against nursing staff. Indeed, across all health care professions nurses are frequently found to face the highest levels of risk of assault. Whittington (1994) found that, in psychiatric settings, from a range of surveys, about 90% of assaults were against nurses who constituted less than 60% of the work force. Occasionally, these assaults prove fatal, for example, the recent murder of a mental health support worker by a patient in his care at Springfield Hospital in London (News item Nursing Times 2003:2).

In the NHS Executive study described above, the incidence rate also varied according to the type of service provided, with mental health/ learning disability services having three times as many incidents as the average for all trusts [24 per 1000 staff per month versus 7 per 1000

staff per month]. This overall figure equates to 65,000 reported violent incidents per year against NHS trust staff.

In October 1999 the UK Government announced an inter-departmental initiative to tackle the rates of workplace violence being experienced in the NHS – the NHS Zero Tolerance Zone Campaign (Department of Health 1999a) - which cited these NHS Executive survey figures, offered advice on good practice and aimed to convince the public, and reassure NHS staff, that violence towards NHS staff is unacceptable and will be tackled. The campaign included specification of 'national improvement targets' - following earlier stated intentions by then Health Secretary Frank Dobson (DoH 1999c) - to force the collation of violence related figures by Trusts using a common definition of violence; to reduce substantially the numbers of incidents over two and four year time periods – by 20% by 2001 and by 30% by 2003; and, publish incident reduction strategies.

Two years after their first report and in light of the zero tolerance campaign the IRS organisation repeated their survey and found evidence of increased rates of incidence, albeit from a much reduced number of NHS Trusts [N=45]. On this occasion average numbers of incidents per trust were reported as 419 in the year to April 1999 and 511 over the next year. This equates to 1200 incidents per 10,000 employees in the year to April 1999 rising to 1400 per 10,000 in year to April 2000 (IRS 2000:4).

Surveys are also conducted by trade unions and professional organizations, and recent results from a regular survey conducted by UNISON (Fidderman 2000:21) reveal that violence to staff has increased between 1995- 2000 and that violence to nurses "increased steadily from 42% [of nurse respondents] in 1995, to 48% in 1999. But, in 2000, this took a massive leap up to 69%".

The very latest figures published in a report from the National Audit Office (NAO 2003) reveal the changes occurring as a result of increased attention being given to the subject. These findings (NAO 2003:2) include:

- Violence and aggression accounted for 40% of the health and safety incidents in the NHS reported to the NAO
- 2000-2001 Department of Health national survey revealing 84,214 reported incidents of violence, an increase of 30% over 1998-1999
- NAO 2001-2002 survey showing a further 13% increase to 95,501 reported incidents and significant variation across regions of the country
- Only 20% of NHS Trusts meeting targets of 20% reduction by April 2002
- "The average number of incidents for mental health and learning disability Trusts is almost two and a half times the average for all

trusts, despite evidence that the staff there are less likely to report incidents of verbal abuse”

- Estimated level of under-reporting at around 39%

1.1.4.2 Costs

Estimating the total cost of violence is difficult since many costs are intangible, for example, loss of morale, problems with recruitment and retention, increased staff fear and subsequent enduring physical or psychological problems (Dickson *et al* 1993). Hence most estimates of costs are regarded as underestimations (Chappell and Di Martino 2000). Hoel *et al* (2001) also highlight theoretical and methodological problems associated with estimating the true costs of workplace violence and emphasise the reservations needed in their reporting of costs to society, organisations and the individual.

Violence is “a major social and industrial issue” (Jackson *et al* 2002:14) and “a serious national health problem” (Flannery 1996:58). The British Crime Survey 1998 estimated that in England and Wales 3.3 million working hours were lost due to workplace violence during 1997 (Budd 1999:36), while a U.S. study cited by Fletcher *et al* (2000) calculated its financial cost as \$55 million per year just in lost wages. The latest NAO report (NAO 2003) highlighted the calculation difficulties before presenting a crude estimation of direct costs of violence to the NHS as at least £69 million per annum, this total taking no account of staff replacement costs, treatment costs and compensation claims.

Internationally, the increase in work place violence has been associated with crises in recruitment and retention of nursing staff (Jackson *et al* 2002, NAO 2003).

Hoel *et al* (2001) identify the U.S. preoccupation with workplace homicides and murders by so-called 'disgruntled workers' even though the former account for only 1 in 650 incidents and the latter for only 4% of homicides. In the USA in the early 1990's approximately 1000 workers annually were killed in work-place violence [homicides] (Jenkins 1996, Elliot 1997, Fletcher *et al* 2000), making it the third most common cause of work-place deaths, although the Occupational Safety and Health Administration (OSHA) (1999) report this figure to have fallen a little in 1997. These homicides are largely related to robbery in retail trades and services rather than health care facilities where non-fatal workplace violence tends to predominate (Elliot 1997).

A further complication of exposure to violence is the possibility of emotional damage and psychological sequelae. As previously mentioned, Budd (1999) identified that verbal threat can have a more serious impact than physical attack. If causes of violence are not tackled or its effects ignored then stress symptoms "are likely to develop into physical illness, psychological disorders, tobacco and alcohol abuse, and so on; they can culminate in occupational accidents, invalidity and even suicide" (Chappell and Di Martino 2000:48).

Several authors (Whittington and Wykes 1992, Wykes and Whittington 1994, Leather *et al* 1997) have investigated the range of anxiety symptoms up to and including those consistent with a diagnosis of post traumatic stress disorder that can follow a physical assault. The recent empirical literature highlights the importance of post-incident de-briefing in order to manage the fear of violence as well as actual exposure (Leather *et al* 1997). The recent NAO report (NAO 2003:4) cited a Nursing Times 2002 survey of 1500 nurses which revealed that of the 581 nurses who had been assaulted whilst on duty only 11% were offered counselling following the incident.

1.2 MANAGEMENT OF AGGRESSION AND VIOLENCE TRAINING

1.2.1 Training: Its Advocacy and Prevalence

Brewer (1999:117) suggests "almost all the published guidance on violence refers to the importance of training as a preventative measure" and, indeed, this has been the case for over a quarter of a century. In March 1976 the then Department of Health and Social Security (DHSS) produced guidelines on 'The management of violent, or potentially violent, hospital patients' (Department of Health and Social Security [DHSS] 1976). This circular asserted that *all* staff, both professional and non-professional, employed in a hospital should receive information and instruction on the principles and practice of dealing with violence, and suggested key content and teaching strategies for induction programmes (DHSS 1976).

Since then numerous other bodies have advocated training of different types to meet the different needs of staff groups. The HSAC study previously referred to asserted that “training in the prevention and management of violence should be available to all staff groups who come into contact with patients and relatives and not only those working in high risk areas” (HSAC 1987:7). Again they proposed a graduated programme of content depending on work area and role that will be reviewed shortly. Ten years later this advice was further developed (HSAC 1997:19) by suggesting that staff managers also need training and specifying content for different levels of training.

Over a decade ago the English National Board for Nursing Midwifery and Health Visiting (ENB) responded to the increased risk of student nurses identified earlier and instructed that all pre-registration programmes must include study related to violence and aggression (ENB 1993). They too indicated appropriate curriculum content along with guidance for trainers’ courses and specifications for ‘control and restraint instructors’. Unfortunately, this requirement is still not the case, as the conclusion by Wells & Bowers (2002) about mandatory pre-registration training in all schools of nursing regarding dealing with violence, referred to earlier in this chapter, confirmed.

Elliott (1997) writing from a U.S. perspective asserted the need for training in basic violence behaviour prevention for all staff along with the need to know the correct emergency response procedures. Once again

different levels are identified for different staff groups and settings and she added that recent Californian law had decreed “mandatory violence training in all work-places for all employees”, with many other U.S. states expected to follow (Elliott 1997:40).

Nineteen ninety-eight was a productive year for advice. The RCN (1998:9) advised employers to provide “appropriate training and education for their staffcommensurate with the degree of risk they face”, as did Beale *et al* (1998) who indicated it should start during the induction and orientation process and be repeated [refreshed] regularly. The highly-valued and well-received Royal College of Psychiatrists guidelines related to mental health settings asserts that “all staff should be trained to recognise warning signs of violence and to monitor their own verbal and non-verbal behaviour” (1998:7).

In their previously-mentioned first report IRS described the provision of staff training as one of a number of important measures trusts can take to prevent violence (IRS 1998). They reinforced this ‘training for all employees’ message in their second progress report (IRS 2000). The second report was prompted by and framed around the then recently released zero tolerance programme (Department of Health 1999a). This programme highlighted the crucial nature of appropriate training in its manager’s guide (Department of Health 1999a).

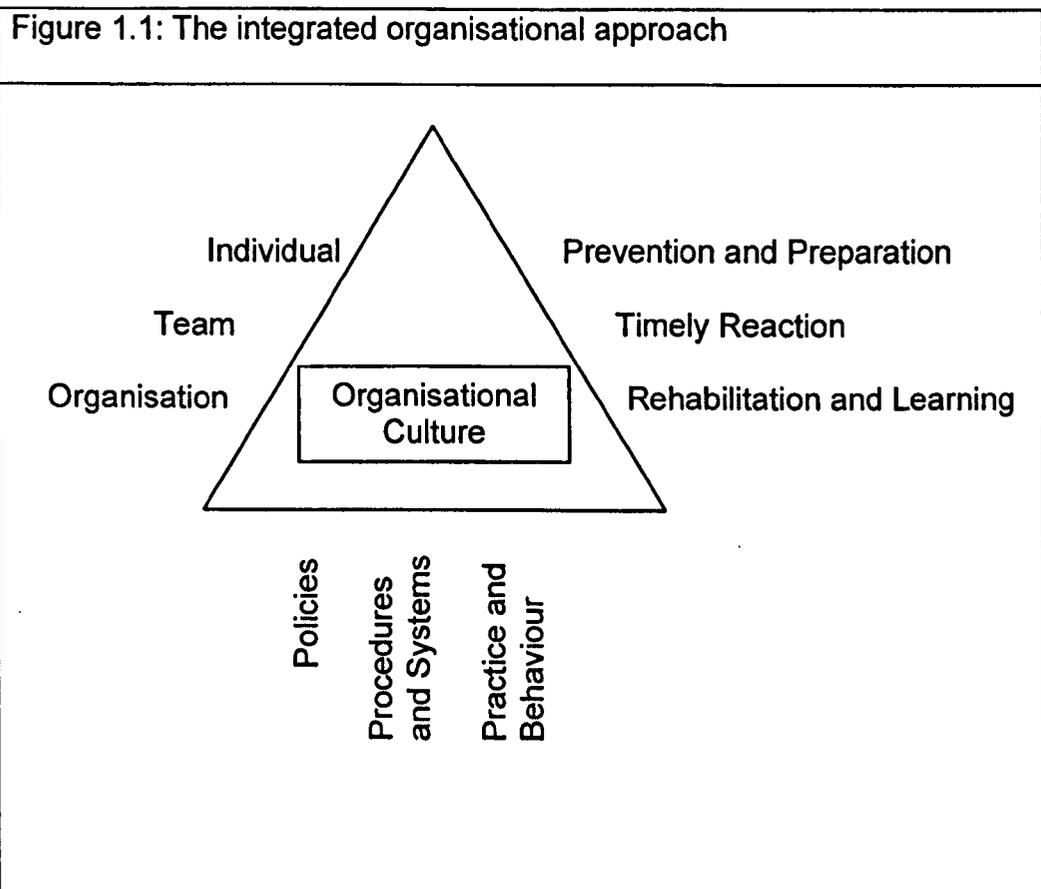
Similarly, Chappell & Di Martino (2000) concluded that regular up-to-date training is essential as part of a battery of 'preventive strategies and measures' that include selection and screening of staff, information and guidance-giving, work organisation and job design, defusing incidents and post-incident de-briefing. Indeed it is the case that many authorities now advocate appropriate staff training not as a 'stand alone solution' but as part of a comprehensive, coordinated health and safety response to the phenomenon of work-place violence (Cembrowicz & Ritter 1994, Cox & Leather 1994, Dickson *et al* 1994, HSAC 1997, Beale *et al* 1998, Royal College of Psychiatrists 1998, Bowie 2000, Hoel *et al* 2001).

The HSAC (1997) suggest a framework that requires three groups of activities to tackle violence in the workplace, namely, research and risk assessment, risk reduction, and monitoring of change. Training is seen as a key aspect of the second activity, along with modifying the working environment, and working practices, increasing security and instituting policies and response strategies. It could be further argued that training pervades all three groups of activities since training could include aspects of risk assessment relevant to the first group and also monitoring and evaluation, part of the third group.

These suggestions resonate with the risk control cycle (Dickson *et al* 1994) and 'total organisational response' model described by Cox & Leather (1994). The 'total organisational response' views the

management and prevention of workplace violence as having a number of inter-related facets; there are obligations for individuals, teams and the organisation, there are preventive interventions to be made prior to incidents, reactive strategies during incidents and rehabilitative and reflective interventions to be made after incidents.

Finally, the model views the setting of over-arching policies, of detailed procedures and protocols and development of safe and professional skills and practice as key activities. Beale *et al* (1998) present the model diagrammatically [See Figure 1.1], making the importance and potential influence of training apparent at all points in the diagram.



The Royal College of Psychiatrists' (1998) guidelines create a multi-factorial response to the management of imminent violence which includes training along with the creation and maintenance of calming environments, varied recreational programmes and activities, generation of policies and protocols, the anticipation of violence through risk assessment and timely de-escalation etc., and medication e.g. rapid tranquillization.

More recently, an extensive consultative survey commissioned by the United Kingdom Central Council for Nursing Midwifery and Health Visiting (UKCC) and conducted by the Institute of Psychiatry (UKCC 2002) asserted that management of the problem of violence needed a multifaceted approach, adding that simply training staff to manage violent behaviour will not, in itself, resolve the overall problem. They advocated a number of strategies to target the problem from a number of angles – organisational, environmental and individual, similar to the list above recommended by the Royal College of Psychiatrists.

[N.B. The latest guidelines from the National Institute for Clinical Excellence –NICE- are due for publication as this thesis is being submitted]

Despite the clear and emphatic endorsement of training offered by eminent authorities summarised above appropriate staff training is still not offered universally or consistently. The HSAC (1987) study previously mentioned reported only 12% of respondents had received

any form of training, the majority of this occurring as part of a basic training programme. In their study of community health care practice Beale *et al* (1998) found wide variations in the type and amount of training received. Beale *et al* (1998) cited an RCN (1994) survey of clinical staff that revealed only one quarter of staff receiving any training regarding violence, and a more recent study (Shacklady 1997) that demonstrated some improvement, finding 51% of respondents had received training in managing violence in the previous five years.

In their 1998 survey IRS showed varying levels of training provision for different staff groups but these itemised figures were poorly presented and difficult to interpret (IRS 1998). In their follow-up report (IRS 2000:9) they reported a similar picture with 93% of responding trusts offering some form of training [82% providing awareness training, 80% breakaway training, 73% restraint training and 25% self-defence techniques].

A recent British survey of over 800 staff working in acute mental health in-patients units conducted as part of a review of the subject on behalf of the UKCC found 88% of this high risk group had received breakaway training [32% reported receiving this during their initial training], and 76% had received restraint training (UKCC 2002). "Very large numbers" had received no training from their Trust since they started work and only a "tiny minority" had received 'refresher training' (UKCC 2002:33).

The report concluded that the background and preparation of trainers, and details of training offered were unclear (UKCC 2002). Additionally, there was “little systematically collected evidence regarding the detailed content or length of training courses” (UKCC 2002:35).

The latest NAO report (2003) found inconsistency between the levels and types of training offered and also the proportions of staff and grades with regard to compulsory or voluntary attendance on training courses. Ambulance and A&E staff were the most likely to have attended specialist training, while only 50% of doctors have received induction training with even fewer junior doctors likely to have attended, and some types of staff [receptionists and porters] rarely receive adequate training despite the emphasis from the zero tolerance campaign that all staff in contact with the public receive training.

This NAO study made several interesting and pertinent observations. Firstly, it identified the limited evidence for safety and effectiveness in health care settings of any taught breakaway, physical control or restraint techniques, despite their popularity (NAO 2003). Secondly, it also found a reactive rather than preventative emphasis in mental health trusts where, despite these trusts having several times the rate of violence of all NHS trusts, a large number of trusts provided higher level diffusion (70%), breakaway (79%) and restraint training (73%) whilst failing to provide training in situation risk assessment (50%) and customer care (36%) (NAO 2003:27). Both these forms of training are

vital since they are preventative and have the potential to reduce both the number and level of violent incidents that confront nursing staff.

The Nursing and Midwifery Council [NMC], the professional and regulatory body for nursing and midwifery, have recently endorsed this latter observation. In a recent newsletter they suggested an over-emphasis on 'hard' restraint skills over 'soft' inter-personal, non-provocative, customer care skills in many forms of training being delivered (NMC 2004a). In a subsequent newsletter the NMC admitted its concern with "the fact that there are currently over one hundred different types of training system being utilised by healthcare services within the UK for managing aggression" (NMC 2004b:5) with no agreed fixed regulation mechanism, although they apparently stated no intention to try to correct this situation.

Lack of risk assessment and inadequate specification of structured training have been associated with the inappropriate hiring of trainers on the basis of informal inquiries or promulgation of out-dated, dangerous or inappropriate training (RCN 1998, Brewer 1999). Beale *et al* (1998) suggest that much consideration should be given to the progression of content and the choice of appropriate trainers for any group of staff, citing Roach (1997) who identified the complexities of managing aggression in a caring service. Roach (1997) claimed that, frequently, inappropriately employed self-defence experts with insufficient knowledge of health care practice will tend to teach what

they know, rather than what is actually required by the particular staff group.

1.2.2 Content of Training

This section will review the *general* guidance on course content, with specific published examples being considered later in the chapter.

Talking broadly about training to prevent workplace violence, Chappell and Di Martino (2000:113) advise that

training involves instilling interpersonal and communication skills which defuse and prevent a potentially threatening situation; developing competence in the particular function to be performed; improving the ability to identify potentially violent situations and people and preparing a 'core' group of mature and specifically competent staff who can take responsibility for more complicated interactions".

From a generic standpoint, Fox *et al* (2002) suggest a mixture of 'soft' and 'physical intervention' skills based on a completed training needs analysis for a particular employer. Identified core 'soft skills' content would include values, definitions, risk assessment and reduction, verbal and non-verbal communication skills, de-escalation, legal issues self-defences and use of force, support and post incident reporting, while core 'physical intervention skills' self-protection and restraint skills in addition to those previously listed.

In relation to healthcare workplace violence, the RCN (1998:9) suggest the following principles for training;

- fit for purpose i.e. reflects local need,
- clear and transparent purpose,
- expressed in learning outcomes,
- based on up-to-date content,
- evidence-based wherever possible,
- delivered by credible staff who value and respect human dignity
- and be responsive to feedback.

Furthermore they suggest a range of courses ranging from half study days on awareness and prevention through principles and practice of personal safety, to short courses on management that include breakaway and restraint techniques to courses to train trainers. The RCN list is similar to that included in the zero tolerance zone campaign literature (Department of Health 1999b), namely,

- up-to-date,
- relevant,
- purposeful,
- backed by evidence,
- given by experts,
- invite feedback
- and, ideally, attended by managers.

Many authorities offer suggestions for appropriate content for staff with different roles and different degrees of patient contact. Many (Beale *et al* 1998, Royal College of Psychiatrists 1998, RCN 1998, IRS 1998, Brewer 1999, Department of Health 1999b, IRS 2000) suggest categories similar to the HSAC (1997) plan of three levels of training.

In its earlier guidelines (HSAC 1987:7) it had advocated material on causes of violence, recognition of warning signs, relevant interpersonal skills and details of management arrangements for “all staff working in areas where the risk of violence has been established”, with the type and depth of training depending on particular roles. Ten years later it advocated (HSAC 1997:19) “good training programmes typically cover:

- theory: understanding aggression and violence in the work place
- prevention: assessing danger and taking precautions
- interaction: with aggressive people
- post-incident action: reporting, investigation, counselling and other follow-up”.

This material was allocated under three levels of training with, for example, basic training for all staff covering the items identified in 1987 [causes etc.], additional training for staff working with violent or potentially violent people also requiring training in de-fusing, de-escalating and avoiding incidents, and breakaway skills and those staff most at risk also requiring restraint skills (HSAC 1997:20). In relation to this content list for good training programmes, analysis of training

content from the IRS survey of Trusts (2000:10) revealed that 97% of employers included assessment of dangerousness and taking precautions, 95% included interaction with aggressive people, 95% included understanding violence and aggression at work, over 88% included reporting and investigation of incidents and 82% included training on counselling and de-briefing. The survey also found that, for higher risk staff, more than 90% of employers provided training in defusing aggression and methods of restraint.

The previously-mentioned ENB guidelines suggested a number of topics for inclusion in pre-registration nurse training, including recognition and prevention of violence, communication skills in relation to impaired perception, assertiveness techniques, defusing, diversion and de-escalation techniques, support, de-briefing and post-incident management, ethical and legal aspects, self-awareness and conflict, understanding violence and aggression as a reaction to circumstances/ conditions, promotion of a positive attitude towards individuals, and physical responses, such as breakaway and escape techniques (ENB 1993). The guidance also specified preparation for suitable instructors. Reference to this list of content will be made again in Chapter 3 when details of the particular Unit under investigation will be described.

McDonnell *et al* (1994) bemoan the lack of a blue print for training and suggested training might include aspects of environmental design and manipulation, de-escalation strategies and simple relatively non-violent

methods of managing incidents. They also include prediction of violent incidents, social skills for defusing incidents, dealing with the physical consequences of violent acts and managing one's own aggression. Finally, they lobby for a comprehensive theory and practice curriculum suggesting, "concentrating on one aspect or another would appear to be only half the answer" (McDonnell *et al* 1994:202).

The Royal College of Psychiatrists' (1998) guidance offers a detailed list of content that includes methods of anticipating, de-escalating or coping with violent behaviour, debriefing and restraint for staff working in mental health services.

In relation to community working, Beale *et al* (1998) offered a list of indicators of good practice. These included emphasis on prevention, calming and negotiating skills rather than confrontation; modular programmes progressing from basic customer care and dealing with difficult clients through to restraint training; material on causes of aggression, reducing risks, anticipating violence occurring, resolving conflict and managing the aftermath of incidents; teaching physical breakaway skills; staff controlling their own feelings; normal and abnormal post trauma reactions; and, familiarity with local arrangements policies.

In relation to acute in-patient mental health services, the report commissioned by the UKCC identified a wide range of courses and

“gained a very strong impression that the detailed curricula for most of the courses is not written down or articulated in any detail in the form of a training manual” (UKCC 2002:35). The report used material from surveys and a consensus exercise to propose essential components of training in the recognition, prevention and therapeutic management of violence, including theoretical aspects of causation, prevention, assessment, and legal and ethical issues, cultural and gender sensitivity; skills of verbal and non-verbal de-escalation, breakaway techniques, and a list of types of restraint techniques (Wright 1999, UKCC 2002).

The Zero Tolerance Zone initiative (DoH 1999a) has subsequently drawn together literature from a number of sources, including some already mentioned [Royal College of Psychiatrists] to give guidance for future development and indication of examples of good practice. With regard to training in Mental Health Services the Zero Tolerance initiative cites the Codes of Practice issued by the British Institute of Learning Disabilities, The Royal College of Nursing and the Mental Health Act 1983 Code of Practice.

These codes set general standards for training and trainers, including training being based on sound theoretical, ethical, and legal principles, physical training being contextual to the service in which the service is provided; training in physical interventions should be provided by people with recognised professional healthcare and teaching

qualifications who have completed recognised courses preferably validated by the RCN or the ENB and have clinical experience in the area where they provide training .

The RCN (1997) published a booklet outlining 'principles of good practice' which includes statements about training within an ethos of caring, and provides an outline curriculum containing the following core aims; courses provide staff with knowledge and understanding of effective and appropriate interventions; explore a range of intrusive and non-intrusive conflict management skills; emphasise safety for service users and workers; and promote understanding of ethical, legal and professional issues. Courses should have an equal balance of theory and practice while indicative content includes relevant psychological aspects of human behaviour and a study of the causes and manifestation of actual and potential aggression.

In September 2002, in consultation with and funded by the Health and Safety Executive, the Employment National Training Organisation [ENTO] produced Standards in Managing Work-Related Violence that consisted of eleven discrete Units and reflected the roles of both managers and front-line workers in preventing and managing violence in the workplace [with the emphasis on prevention]. The Standards are generic, widely applicable, and are anticipated to be useful in determining good practice, identifying staff at risk, identifying training needs and accrediting training programmes. Indeed, it is suggested that

“one of the most important roles for the standards will be as a benchmark for training providers in this area of conflict management. It is clear that there are many such training providers around. Many users of such services, concerned about the legal implications including from the use of physical intervention, recognise that choosing a training provider with the right motivations, experience and understanding, is highly critical” (ENTO 2003).

It is not expected that any one person would complete all eleven units since they equate to different roles and, for the same reason, no qualification is associated with completion of units. Nevertheless, taken together they give a good indication of comprehensive content of a violence prevention and management course.

The Units, each with a small number of elements, have the following titles:

Unit W1 Assess the risk of violence to workers

Unit W2 Develop an effective policy and procedures for minimising the risk of violence to your workers

Unit W3 Implement policy and procedures to reduce the risk of violence to your workers

Unit W4 Develop and maintain an effective management information system

Unit W5 Promote a safe and positive working environment

Unit W6 Ensure your actions contribute to a positive and safe working environment

Unit W7 Protect yourself from the risk of violence at work

Unit W8 Respond to work-related violent incidents

Unit W9 Support individuals involved in violent incidents at work

Unit W10 Investigate and evaluate incidents of violence at work

Unit W11 Ensure effective communications following an incident of violence at work

Fox *et al* (2002) suggest that using the Standards will lead to sharing of best practice within and across sectors, and to better reporting and understanding of the whole area of work-related violence. Fox *et al* (2002) anticipate that the standards will be used, not just by managers but also by other groups, such as employees and trade unions to make requirements of managers, by trainers to structure courses, by professional bodies to specify membership and professional development, and by awarding bodies as a basis for qualifications.

1.3 REVIEW OF TRAINING COURSES FROM PUBLISHED

LITERATURE

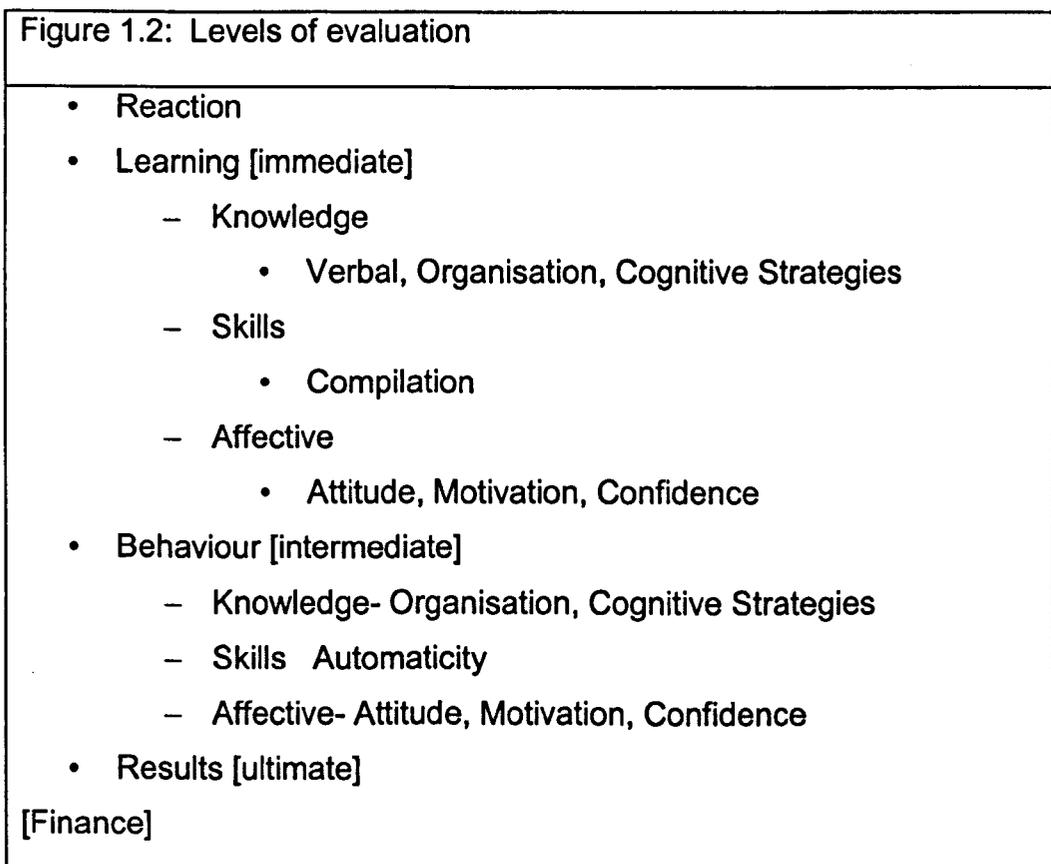
Whilst the calls from the literature to provide training are virtually unanimous there is a less compelling position with regard to the objective justification for training. The benefits of training might be apparently obvious but they become less clear and more difficult to establish on closer examination, not least because so few trainers have

attempted to perform and publish objective evaluations. Hoel *et al* (2001:57) concluded that trainers “failed to undertake a proper evaluation” and from their review of the international literature found “few which included a clear programme description. Even less included information on programme evaluation”. Similarly, in their review, Paterson & McComish (1998:228) concluded that published training course evaluations illustrated “ an extremely wide variation in programme content and duration, together with inadequate descriptions of programme content”, an identical conclusion to that recently reached in the UKCC report (2002).

There are many possible measures of change as a result of attending a training course. The next chapter will consider in detail the process of training evaluation and review models used to guide such a process, including the work of Kirkpatrick (195/9,1976), Warr *et al* (1970), Hamblin (1974), and Kraiger *et al* (1993). Taken together these frameworks clearly highlight a wide range of possible indicators of training effects and potential dependent variables to be used in an evaluation research study. Broadly speaking, changes can be determined in the knowledge and practice of individual course attendees, the functioning of the department in which they work or more broadly in the organisation. As an example of the latter categories, the HSAC (1997) asserted that training can lead to a reduction in the number, seriousness and subsequent psychological effects of

incidents, and an improvement in response to incidents and staff morale.

Whilst Chapter 2 discusses the merits and limitations of each model, it is possible to view them as potentially complimentary, rather than in an either-or competitive manner. Combining the strengths of each into an integrated taxonomy containing elements from several of the popular named models highlights many more ways of determining the effects of training on different aspects of the individual trainee or their employing organisation. Such a taxonomy would look something like Figure 1.2.



For each level and category of change a range of appropriate measures could be used to determine any effects of training. Figure 1.3 identifies a range of these measures of change [potential dependent variables] pertinent to aggression /violence/conflict management type training, aligned with the different levels of evaluation identified in Figure 1.2.

Figure 1.3: Possible measures of change for violence management training	
<ul style="list-style-type: none"> • Reaction • Learning [immediate] <ul style="list-style-type: none"> – Knowledge <ul style="list-style-type: none"> • Verbal, Organisation, Cognitive Strategies – Skills <ul style="list-style-type: none"> • Compilation – Affective <ul style="list-style-type: none"> • Attitude, Motivation, Confidence • Behaviour [intermediate] <ul style="list-style-type: none"> – Knowledge <ul style="list-style-type: none"> • Organisation, Cognitive Strategies – Skills <ul style="list-style-type: none"> • Automaticity – Affective <ul style="list-style-type: none"> • Attitude, Motivation, Confidence • Results [ultimate] • [Finance] 	<p>Subjective satisfaction sheets</p> <p>Knowledge tests, vignettes, scenarios Problem-solving Role plays Skills demonstrations Attitude scales Self-assessment – confidence, morale job satisfaction</p> <p>Manager review Customer satisfaction</p> <p>Events – number type and seriousness Sickness claims Complaints</p> <p>First aid consumables Compensation</p>

This section will summarise the details and specified content of published courses and also highlight the measures used to determine change and the changes evaluated. It is not intended to be a definitive review of all published training courses. Rather it is a summary consisting of

- the most frequently cited evaluations from the psychiatric literature, along with
- some examples from other health care settings – learning disability, elderly care, general hospitals and mixed groups.

In collating the studies from the psychiatric area one obtained a definite sense of circularity and repetition within the literature. Several dated studies were still regularly cited, further endorsing the earlier statements about the relative paucity of studies of training given the importance of the violence problem (McDonnell *et al* 1994, Allen 2001, Hoel *et al* 2001, UKCC 2002, NAO 2003).

Harris (1996) conducted a detailed review of physical restraint for challenging behaviour and devised a helpful framework for summarising the studies. This included the following headings, participant characteristics [number, gender, and age of participants], types of behaviour and treatment interventions or restraint procedures included, research design, and methodology, statistically reliability measures, and main outcomes of the study. More specifically related to workplace

violence, Allen (2001) produced a guide for trainers in physical interventions which included a review of published training studies structured using the following headings to delineate training course characteristics— design [research method, controls, sample size, statistical analysis], focus [training style or allegiance, client group, staff group], measures [dependent variables and measurement tools], outcomes [effects noted].

Due acknowledgement is given to these two researchers and their frameworks have been adapted in summarizing the studies selected for inclusion in this review, with additional information being listed, for example, a summary of content. Details of each selected study are presented in two Tables, prior to observations being discussed more generally. Table 1.2 contains summaries of studies from the psychiatric literature, while Table 1.3 contains summaries of studies located concerning other areas of health care, namely, learning disability, elderly care and mixed staff groups [including student nurses].

Table 1.2: Summary of training courses concerning psychiatric settings									
Authors & Year	Design			Focus			Measures		Outcomes Effects
	Research design/ methods	Sample, Course duration	Allegiance Country/ Staff/ Clients	Therapeutic	Dependent Variables	Data collection tool	Statistical tests used		
St Thomas Psychiatric Hospital 1976	Descriptive year before and year after	N=? 10 hours (5 x 2 Hrs)	Canada Psychiatric Hospital	Therapeutic milieu, observation of changing behaviour non-physical prevention and physical interventions	Number of violent incidents, patient injuries, staff injuries, number of staff man-hours lost due to assault-related injuries	Incident figures Patient accident reports Staff injury claims	Totals percentages	Rates of assault decreased by 9.4%, patient injuries decreased by 12%, staff injuries decreased by 10.4% & lost man- hours reduced by 31%	
Gertz 1980	Descriptive year before and year after	317 staff of treatment teams 2 day workshop	USA Mental health Centre	Prevention guidelines do and don't list Verbal de-escalation skills, breakaway and restraint techniques preventative plans	Confidence patient related accidents	Incident figures	Totals	33% reduction in incidents over 2 year period	

Lehmann Padilla, Clark & Loucks 1983	Within subjects pre/ and immediate post Interview	N=144 5 hrs	USA Veterans Hospital	prevention detection of signs, verbal de- escalation legal issues demonstration practice restraint + breakaways 3 trainers 20 minute presentations	Attitude to violence, objective knowledge confidence comfort detection number of assaults	10 item questionnaire + interview form [8 staff	not given [t-test for interviews	improvements in objective knowledge and significantly more confidence post – training, able to describe precursors & nature of supportive or directive comments, slowed acceleration of rate of incidents
Infantino & Musingo 1985	Between groups Follow-up for 9- 24 months after training trained v. untrained	N=96 [31 attended v. 65 not attended training] care and nursing aides 3 days	USA Florida psychiatric hospital	aggression control techniques local policy, verbal defusing, case vignettes, breakaways, restraint + relocation reporting	Number of attacks number of staff injuries, also follow up group of 56 staff relaxation, confidence, ability to handle situations	Incident records	percentages Chi-squared Yule's Q	Trained staff less likely to be assaulted than untrained, trend towards less likely to be injured during incident, 86% of trained reported increased confidence

Thackrey 1987	Within subject, trained v untrained before v 18 month after	N=125 Community mental health, psychiatric and prison staff 8 hours	USA mental health resource centre, psychiatric and state prison	Therapeutic for aggression training	Staff self- confidence ability preparation comfort in safety effectiveness in self- protection and therapeutic intervention	Thackrey Confidence in coping with patient aggression instrument	Cronbach alpha Reliability	Trained staff showed increased confidence maintained at 18 month follow-up
Carmel & Hunter 1990	Compared high and low training compliance staff groups	N=744 All staff on 27 wards, 16 hours	USA State forensic hospital	Interpersonal skills, didactic and practical instruction in management of violent patients	Compliance with training, rate of employee injury, number of incidents per bed	Staff training compliance records Staff injury records	t-test pearson correlations fisher exact test	Compliance with violence training was associated with lower staff injury rate but not lower rates of incidents
Goodykoo ntz & Herrick 1990	Longitudinal immediately before & 4 months after	N=27 registered staff & aides 4 workshops, 1 session per month	USA University psychiatric unit	verbal de- escalation early recognition signs, self- awareness skills of location, breakaways and restraint, burnout & feelings	Burnout – confidence & control, number of incidents, number of staff injuries	Burnout scale Monthly incident figures Subjective evaluation	Totals Incident rates/month	No overall impact on burnout measures, reduction in number of incidents, anecdotal evidence of fewer injuries

Paterson, Turnbull & Aitken 1992	Within subject pre/and immediate post	N=25 10 day post basic	UK Scotland Mental health + learning disabilities	current practice legal issues safety, prevention verbal and non-verbal, breakaways, restraint (C&R)	Knowledge stress and burnout job satisfaction role conflict skills competence (de-escalation, breakaway and restraint)	questionnaires, factual questions, Likert scales videotaped role-play of skills rated by 'blind' instructors	t-test Wilcoxon test signed rank test Test-retest reliability Inter-rater reliability	Significant changes in knowledge, stress level, role conflict Improvement in role play performance [all except job satisfaction improved]
Collins 1994	Within subject pre/post test + 6 month follow-up	N=31 5 days	UK Scotland Mental Health staff + student nurses	Prevention and management of aggressive behaviour programme Prediction and assessment, De-escalation verbal / nonverbal breakaway restraint skills	Staff attitudes/beliefs towards: violence prediction, patient motivation, patient responsibility for aggression, staff anxiety + fear of assault, need for skilled intervention, staff confidence	12 statement instrument Attitudes Toward Aggressive Behaviour Questionnaire	Likert Category totals	Positive response to all areas except attitudes towards expectation of assault and patient responsibility for violent behaviour

Martin 1995	Descriptive longitudinal 1 year before and 2 years after	'All staff' 1 day workshop + video on de- escalation + competency test +annual update	USA Dept of psychiatry Teaching Hospital	protocols interventions practice of breakdowns, takedowns restraint seclusion, video of verbal de- escalation	Number of incidents level of aggression type of injury number of missed work days cost to department	Incident records Work sheets	totals	Fewer actual aggressive incidents, reduction in injuries in 2 nd year. Dramatic reduction in days missed due to injury and cost of injuries [\$173000 to £2414]
Philips & Rudestram 1995	Between group & control [2 different experiment-al groups [theory + theory and skills],	N=24 allocated to 3 groups of 8 4 hrs	USA 2 state psychiatric hospitals male staff	Violence theories, non- verbal communication warning signs intervention strategies and legal issues non-violent defence (distance, stance pliancy, push-off)	Physical skills competence, staff fear, staff display of aggression number of assaults, number of injuries	Questionnair es, self- completion, independentl y rated pre/post 60 second role play Buss-Durkee hostility-guilt inventory + other self report Likert scale & follow-up questionnair es	3x1 ANCOVA judge rating ANCOVA Newman- Kuels post hoc tests Inter-rater reliability	Knowledge and skills group training had superior rating of competence and lower fear and aggression rating cf. other group or control. Increased competence correlated with change in fear & aggression scores. Later report of fewer incidents and injuries for skills group

Whittington & Wykes 1996	longitudinal design between and within groups. 28 days before v. 28 days after attend v. non-attender	N=155 nurses [47 attended, 108 non attenders 7 hrs	UK England Two psychiatric hospitals 13 wards high compliance v. low compliance wards	Psychological strategies for coping with violent patients risk assessment, prevention and defusion, post assault management, legal issues PTSD verbal & nonverbal, coping strategies	Number of notified assaults number of staff assaulted frequency of assaults on high compliance wards	Contacted wards daily rather than relied on official records	Chi-squared test of association McNemar change test	31% reduction in overall rate of violence no difference between attenders v non-attenders on number of assaults. frequency of assaults on high compliance wards fell by 2/3 while on low compliance rose by 1/2
Parkes 1996	Repeated measures 18 months before and 12 months after training	'All nursing staff' 4 days	UK England Medium secure unit 44 beds	Non-touch de-escalation, 3 breakaways, 3 person restraint	Nature of incident, no of staff involved, no of injuries, staff feelings re. incidents of restraint involving C&R	Rating scale	Totals Chi squared	Overall slight increase in staff injuries during restraint. No change in rating incident seriousness or management difficulty, satisfaction with techniques

Flannery, Hanson, Penk, Goldfinger, Pastva & Navon 1998	Longitudinal multiple baseline, 3 month before scheme v. 1 year after	N=940 Total nursing staff in three hospitals	USA Psychiatric Hospitals	Assaulted staff action programme crisis intervention approach	Rates of assaultive behaviour	Incident figures	Repeated measures analysis of variance Post hoc Dunnett's test	Significant reduction in rates of assaultive behaviour in first quarter after implementation maintained for 1 year study period
Forster, Cavness & Phelps 1999	Longitudinal year before v. year after	All staff with patient contact '1 day +weekly discussion + display of results	USA 83 bed acute psychiatric hospital	hands on 5 point restraint self-defence containment techniques inappropriate use verbal de-escalation role play	Rate of restraint duration of seclusion and restraint per episode staff injuries	Incident & restraint figures	totals rates averages	14% reduction in restraint usage 55% reduction in duration of seclusion/restraint episodes 19% reduction in staff injuries

McGowan, Wynaden, Yassine & Parker 1999	Loose between group and within group study Partial 6 month followup	N=70 28 received new training Psychiatric intensive care staff 1 day & 22.5 hour programme	Australia 2 psychiatric hospitals	Legal & ethical issues, cultural issues effective communicating, managing challenging behaviour, breakaway and restraint techniques	Staff self-confidence ability preparation comfort in safety effectiveness in self protection and therapeutic intervention	Thackrey Confidence in coping with patient aggression instrument 10-item confidence questionnaire	Mann-whitney U Wilcoxon matched-pairs sign-rank	Initially 1 day training programme produced higher confidence in all categories. Later 22.5 hour training raised confidence in 4 categories and surpassed confidence of previously better hospital in other 6 categories
Calabro, Mackey & Williams 2002	Within group pre/and immediate post test	N=118 Nurses aides, therapists 12 hours	USA Acute care psychiatric hospital	Non-violent crisis intervention, prevention and control of disruptive behaviour, self-defence and restraint	Knowledge, attitude self-efficacy, behavioural intention to use techniques learned	34 item scale 5 point Likert scales	Internal consistency Cronbach alpha Paired t-tests	Significant short-term improvement in all four areas studied Anecdotal evidence of reduced number of incidents in 4 months of training implementation

Ilkiw-Lavalle, Grenyer & Graham 2002	Between and within group Pre-and immediately post	N=103 Nurses allied & medical and ancillary 2 days	Australia	INTACT aggression management program Legal issues, characteristics of aggression, prediction, managing aggression, reporting self-care	Knowledge course evaluation	14 item knowledge test requiring short written answers 8 item 10-point Likert programme evaluation	Cronbach alpha Paired t-tests Effect size estimation Analysis of covariance Tukey's HSD	Nurses highest on pre-test. All groups significantly improved post-test with no difference between occupational groups post training Staff without prior training had greatest improvement mainly re. prediction, management and legal aspects of aggression
--------------------------------------	---	--	-----------	--	-----------------------------	--	---	--

Table 1.3: Summary of training courses concerning learning disability, elderly care and mixed settings

Authors & Year	Design		Focus			Measures		Outcomes Effects
	Research design/ methods	Sample, Course duration	Allegiance Country/ Staff/ Clients	Content	Dependent Variables	Data collection tool	Statistical tests used	
LEARNING DISABILITY SETTINGS								
Allen McDonald, Dunn & Doyle 1997	Descriptive No pre measure no control 5 year study	Not stated 3 days	Wales UK 6 bed residential unit	Understanding aggression, risk factors, verbal & non-verbal defusion, self-defence and minimal restraint post incident support	Number and type of incident, level of response reduced staff and client injuries	Clinical records Spirrison & Grosskopf scale of incident severity Institutional records of injury	Percentages Linear regression lines using least-squares method, Pearson r correlation	Trend towards decreased use of major reactive strategies and decrease in staff and resident injuries
McDonnell 1997	Longitudinal pre-& immediately post	N=21 Care staff trained & untrained 3 days	UK Learning Disability service	Law, causes, defusion, management, breakaway, non-violent restraint methods	Knowledge of incidence and causes, management of behaviour and recording confidence	Violent Incident Knowledge Test [20 item]. [15 item] Managing Challenging Behaviour Confidence Scale Restraint Role Play Test	Cronbach alpha factor analysis t-tests	Increase in confidence was most significant effect. Also increase in knowledge and skills of restraint

Allen & Tynan 2000	Longitudinal Between and within group	N=109 [51 previously trained, 58 not] Staff working in community settings 3 days	Wales UK Community learning disability service	Management of Aggression Training Programme: risk factors, distraction & defusion strategies, non- aggressive physical intervention, breakaways	Knowledge of reactive behaviour management confidence	Thackrey Confidence in coping with patient aggression instrument [10 item] 11 point scale Reactive Strategies Questionnaire [20 multi choice]	Chi -squared Cronbach alpha Mann- Whitney U Wilcoxon signed -rank	Training significantly increased confidence and knowledge raised scores to level of those staff previously trained
Baker & Bissmire 2000	Longitudinal Within group pre and 3 months post No control	N=17 Care staff [mostly untrained] 2 days	UK Independent residential learning disability	Strategies in Crisis Intervention and Prevention: Values, understanding challenging behaviour, prevention, early non-aversive intervention, health and safety, legalities	Confidence level of support number of incidents type of intervention	Questionnaire 5 point Likert Client records	Mann- Whitney Wilcoxon Rank Sum Spearman rank correlation	Improvement in level of confidence in handling & preventing incidents. Staff felt more supported. No difference in number of incidents per month but greater tendency to respond physically after training

McKenzie, Paxton, Patrick, Matheson & Murray 2000	Longitudinal Between groups Pre/ post [immediately or 3-12 months after]	N=59 trained and N=73 control Residential, day care and health staff 1 day	Scotland UK Learning disability service	Duty of care, assessing behaviours challenging behaviour Reactive strategies Behavioural & positive programming approaches	Knowledge; definition of LD, management of challenging behaviour, client choice and duty of care	Questionnaire, Two scenarios measuring responses to challenging behaviours Training evaluation	Kappa t-tests ANOVA Tukey best	Trained significantly increased knowledge in all areas, except those important in managing challenging behaviour. Untrained group no change. Knowledge not faded over 12 months
Gentry, Icceton & Milne 2001	Longitudinal Pre & immediate post	N=101 Community residential respite and day care 3 days	UK Learning disability service	Interactive staff training whole team approach :challenging behaviour, assessment, functional analysis guidelines, non- aversive management of behaviour	Knowledge skills attitudes	Knowledge questionnaire, individual competence and clarity rating Role play Training evaluation	t-tests percentages	Highly significant improvement in knowledge, skills developing, attitudes improved re both job clarity and confidence in managing challenging behaviour Very favourable review of course content and whole team approach

ELDERLY CARE SETTINGS

Hagen & Sayers 1995	Interrupted time series 8 days before and after training	N=171 Nurses, aides recreation therapists 3x30 minute modules in 3 months	Canada Long-term care facility	Aggression and dementia, risk factors, prevention strategies, de-escalation strategies, protective interventions	Number and type of aggressive incident	Incidence and nature of aggressive incident form	Percentages Paired t-test	Reports of aggression dropped by 50% after training. Significant in each comparable shift Qualitative approval of training
Maxfield Lewis & Cannon 1996	Between & within group pre/post Variable baseline	N=97 Nurses and assistants 2x 1hour session +individual on-job assistance	Elderly care hospital	Appraisal of approaches theories self care, therapeutic intervention, methods and demonstration all related to bathing and grooming	Knowledge number of incidents connected with bathing & grooming prn medication staff injuries	21-item test Hospital incident report forms	Cronbach alpha ANOVA Bonferroni method of multiple group comparisons Means and percentages	Meaningful gains in knowledge, qualified nurses scored highest pre and post 50% Reduction in incidents and prn one month before and after, injuries drooped to zero

Wilkinson 1999	Descriptive Longitudinal 12 months Pre and 11 months during training	N=32 Nurses and assistants 1 day work- shop +ongoing training	USA Elderly care hospital	Individual treatment risk factors, assertive communication, reporting, physical protective techniques	Number and type of aggressive incident	Hospital incident report forms	frequency	Decrease in patient injuries but almost doubling of number of incidents during training of year before. Query accuracy of recording prior to training
MIXED SETTINGS								
Hurlebaus 1994	longitudinal within and between groups pre & immediate post + control	N=42 Conven-ience sampling Trained nurses 2 hrs	UK general hospital	Theories, identifying and preventing aggressive behaviour, verbal de-escalation, physical defence techniques	Objective knowledge safety & confidence	15 multi-choice question test visual analogue	t-test	Significant change in knowledge, no change in safety or confidence
Hurlebaus & Link 1997	longitudinal within and between groups pre & immediate post test + control	N=32 Trained general nurses and assistants 4 hrs	USA inner city teaching hospital	'Managing Aggressive Behavior': statistics, signs of agitation, de-escalation, physical defence and restraints	Objective knowledge safety & confidence	15 multi-choice question test visual analogue scales	correlations t-test	Significant change in knowledge, no change in safety or confidence

Beech 1999	Longitudinal within group pre & immediately post	N=58 Student nurses 3 days	UK Students destined for different branches	Theories, risk factors, self awareness, risk factors, assessment, management, de- escalation, legalities breakaway	Knowledge attitudes confidence	20 statement attitude questionnaire Course evaluation	Percentages Wilcoxon signed-rank matched pairs	Increase in positive attitudes regarding confidence, risk identification General endorsement of topic and course as appropriate for student nurses
Arnetz & Arnetz 2000	Longitudinal between groups Pre/ post Over two years	Approx 1500 Staff at 47 health care work places [24 intervention & 23 control] ongoing	Sweden Emergency department, elderly care, psychiatric and home health services	Use of a Violence Incident Form VIF Structured programme of discussion at intervention sites	Number of incidents awareness of risk situations knowledge of management	19 item questionnaire regarding previous violence experience VIF recording all violent incidents Discussion & feedback at intervention	Chi square ANOVA Logistical regression	Reduction in violence more so for control group Intervention encouraged reporting of violence so that it could be discussed

1.3.1 Summary of Literature Review of Aggression Training Course

Evaluations

Notwithstanding the difficulties identified earlier, some conclusions will now be attempted.

1.3.1.1 Care Area

Whilst not intended to be exhaustive [as stated above], most studies in Tables 1.2 and 1.3 related to the area of psychiatric care [18] with learning disability being the next most prevalent area [6]. Allen (2001) [writing from a learning disability perspective] found similarly with mental health and learning disability making up respectively 44% and 39% of total reviewed studies in that report.

Learning disability courses tended to have a broader target area, usually being directed at the management of 'challenging behaviour' rather than aggression and violence, the latter being viewed as a subset of the former.

1.3.1.2 Course Duration

There was a surprisingly large range of time for the duration of courses, ranging from two hours through to ten days. Generally, this related closely with the ambition of the learning objectives but one obvious exception is the study by Lehmann *et al* (1983), which sought to deliver in five hours the range of theory and practical skills that most other

trainers attempted in 5-10 days. Most courses were for a block of time [hours or days] but there were some examples of short workshops spread regularly over several months.

Published courses for staff in psychiatric settings showed the biggest range, being skewed towards the shorter end of the time scale, in total thirteen were for 3 days or less and only 3 were for four days or more [but one of these was for ten days]. This short duration may not be typical of actual courses being delivered since, for example, a survey by Lee *et al* (2001) of psychiatric intensive care unit and regional secure unit staff in England and Wales found an even larger range of course duration [0.5 days to 21 days] with a mean average course length of 6 days.

In learning disability areas two-thirds of included courses lasted three days and the remainder lasted 1-2 days. In elderly care settings the maximum length of course located was for one day, the other two lasting up to 2 hours.

1.3.1.3 Country of Origin

Four countries were identified in the published studies. Within the psychiatric studies in Table 1.2, North America was the most frequent originator [11 from United States of America, one from Canada], followed by UK [4 studies] and Australia [2 studies].

Chronologically, the first seven courses in the review were North American, highlighting their early interest in the topic, and it is only in the last few years that Australian studies have been located. It is apparent that aggression management training courses have been evaluated for a longer period of time in psychiatric settings [and presumably delivered for a longer period of time], being, comparatively, much more recent in learning disability, elderly care and general settings.

1.3.1.4 Content

When reviewing course content related to legal and ethical aspects Beech (2001) identified that many training courses, especially early ones had a very practical orientation and made no explicit mention of ethical issues. Legal issues fared better being more likely to have explicit mention.

Generally there appears to be a clear unofficial 'core curriculum' that closely mirrors the suggestions mentioned earlier (ENB 1993, McDonnell 1994, Royal College of Psychiatrists 1998, Beale 1998, UKCC 2002).

Two interesting recent exceptions are the work of Flannery (1998) that demonstrated a reduction in violent incidents following introduction of a crisis intervention / staff support scheme rather than a training course, and the study by Arnetz & Arnetz (2000) that focused on widespread

consistent use of a particular report form and subsequent discussion and reflection on incidents which produced a similar effect on the numbers of incidents.

1.3.1.5 Research Design

A tendency is noted for many studies to be descriptive observations of violent incident rate over time [approximately 13 examples]. A large number of studies were quasi-experimental in using within-subject design to measure pre- and [usually immediately] post-training scores [approx 18], whilst a small number attempted between subject design using no-training controls or staff trained in a different manner or at an earlier time or in a different hospital [approx 9].

When evaluation of course attendees was performed it was predominantly pre- and immediately post-training, with one study measuring only post-training, and little evidence of following up any enduring effects or subsequent change [only 4 studies included follow-up].

1.3.1.6 Sampling

Four psychiatric studies didn't specify numbers in sample. Of the remainder, the range of participants spanned from 24 to 940. However, this top figure related to Flannery's study (1998) of crisis intervention without training and if this top figure is removed then the largest sample becomes 744 (Carmel & Hunter 1990).

Yet, doubts remain even here when the authors talk about non-compliance with training so the reliable top figure for trained staff is 317 (Gertz 1980). Over one third of the studies had less than 31 experimental participants and half the studies had less than 100 trained staff in the sample.

1.3.1.7 Statistics

Only a few studies contained any mention of establishing the psychometric qualities, for example, reliability, of data collection measures. Seven mentioned Cronbach alpha, often in connection with use of a pre-existing scale such as Thackrey's "Confidence in coping with patient aggression scale", while two studies mentioned inter-rater reliability.

Otherwise, chi-squared, Wilcoxon sign-rank, matched pair test, and t-test were used but to a large extent there was a strong reliance on percentages or raw frequency totals.

1.3.1.8 Dependent Variables

Earlier in this section a range of potential effects of training [training outcomes] were identified [see Figures 1.2 and 1.3]. Allen (2001) identified direct [effects on person] and indirect outcomes [consequences on service]. Many measures used in the studies are indirect and often acknowledged as being unreliable, for example, violent incident rates.

With regard to the training courses in psychiatric settings [Table 1.2], The most frequently measured dependent variables were number of incidents [12 studies], number of injuries [9] [usually staff but sometimes including patient injuries], and staff confidence [8]. Less frequently self-efficacy [6], staff attitudes [6], knowledge [4], skills [3] or costs [2] were recorded.

With respect to other settings [Table 1.3], a slightly different picture was obtained. Knowledge was the most frequently measured dependent variable [9 studies], with confidence and number of incidents each being recorded in 6 included studies. Attitudes were measured in three studies and skills and number of injuries were each recorded in two studies. If the studies in the two Tables are combined, the most frequently utilised dependent variables are number of incidents [18], confidence [14], knowledge [13], number of injuries [11] and attitudes [9].

1.3.1.9 Outcomes

The selected evaluation studies demonstrated desirable trends in a number of these areas but the results were inconsistent, and indeed, sometimes contradictory.

Knowledge was usually tested via completion of short answers or multiple choice-type questions. Trained staff tended to be more knowledgeable than untrained, a difference maintained at follow-up.

In relation to confidence measures, similar trends are noted, trained staff were more confident following training, with differences detectable after up to 18 months, but impact may be low and less so in females (Allen 2001). Confidence was assessed via Likert-type statements or validated confidence scales, for example, Thackrey's "Confidence in coping with patient aggression scale".

With regard to skills, studies showed that staff could develop competence in a range of physical interventions and demonstrate this new ability in role-plays, yet role-play generates high stress levels, may be more difficult for older staff, and can be inconsistently rated by observers. Staff may freeze in real situations or even behave more aggressively. Improvement in staff stress and motivation but not job satisfaction was also supported by a small number of studies (Allen 2001).

With regard to indirect measures, most studies demonstrated that the numbers of incidents, staff and client injuries all tended to decrease following training but some studies report no change and several showed increases in number of incidents or staff injuries. As Allen (2001:27) concludes, "unfortunately, the research indicates that none of the above outcomes can be guaranteed from the training, and negative results have also been observed in each of the above areas".

1.3.2 Conclusions from Review of Aggression Training Course

Evaluation Studies

Given the 25 year time-scale, seriousness, and cost [financial, human resource and personal] of the topic, the paucity of studies and relative dearth of thorough, rigorously designed studies is a great disappointment.

Those studies identified are limited and deficient in many differing respects, including

- poor research design, at best usually pre- and immediate post-training testing
- poor consideration of reliability and other psychometric properties,
- inadequate consideration of possible contemporaneous changes in service organisation or context
- infrequent follow-up of interaction with practice [learning transfer]
- small sample
- limited range of outcomes measures to determine learning (Kirkpatrick 1976) - knowledge , confidence, attitudes
- emphasis on subjective reaction level (Kirkpatrick 1976) or easily obtained but unreliable measures of behaviour change (Kirkpatrick 1976), for example, number of incidents reported
- very little research on student nurses, despite their acknowledged prominence in the high risk categories of health

workers professional groups (HSAC 1987) for all types of aggression

- often basic data analysis techniques, for example, some of the most frequently cited studies restricted data analysis to raw tally counts and percentages

In summary, there is currently insufficient (a) rigorous and systematically designed evaluation research (b) which includes reliability ratings and (c) long-term follow-up data (d) with sufficient depth and detail in the types of data collected and its analysis (e) related to an acknowledged high-risk and vulnerable health staff group.

Some of these limitations can be understood. For example the uncertainty about the actual levels of violence and numbers of incidents or the ethical difficulty or practical impossibility of creating strict control and experimental groups. Others are more easily corrected, for example, better, more rigorous research design, more sophisticated measurement of training outcomes, incorporating more complex models of learning (Kraiger *et al* 1993), or more robust statistical analysis of data.

The study reported in this thesis was conceived and completed in response to the existing inadequacies listed above,. It aimed to address some of the currently existing deficits and omissions in the area and build upon the author's existing interests and earlier attempts at training

evaluation. The study attempted to use a rigorous research design to obtain more detailed and precise data about the effects of attending one particular, established aggression prevention and management training course. Details of the course, its conception and rationale, learning outcomes/ objectives, content, duration, position within an existing curriculum etc. will be provided in Chapter 3.

1.4 RESEARCH QUESTIONS AND HYPOTHESES:

A number of researchable questions and hypotheses were educed.

1.4.1 Research Questions

The study endeavoured to answer a number of research questions:-

Question 1: What is the extent of the exposure of student nurses to aggression and violence during the first year of their training course?

Question 2: Does a relatively short, three-day unit on the prevention and management of aggression in health care settings have a positive effect on its student nurse course participants?

Question 3: Can desired changes in learning domains, including knowledge, attitudes, confidence, and self-assessed competence be unambiguously detected following attendance on the course?

Question 4: If changes in different learning domains are detected can these be monitored and subtly explored in detail and depth?

Question 5: Do any immediately detected changes [identified in question 4] remain, increase or deteriorate over time when the student nurses return to clinical practice?

Question 6: Do different student nurse sub-groups respond differentially to the training course?

Question 7: How do a range of techniques for measuring training effects differ in sensitivity, usefulness and ease of use?

1.4.2 Research Hypotheses:

From these research questions a number of hypotheses were specified for investigation, namely:-

1. Student nurses are involved in substantial numbers of violent incidents during their training.
2. After attending a three-day training unit as part of their first year training programme student nurses will demonstrate increased knowledge about causation and prevention of violent incidents in health care settings.

3. After attending a three-day training unit as part of their first year training programme student nurses will demonstrate increased confidence in their ability to remain safe while interacting with aggressive clients.
4. After attending a three-day training unit as part of their first year training programme student nurses will more broadly attribute the blame for causation of violent incidents and not always blame the assailant.
5. After attending a three-day training unit as part of their first year training programme student nurses will demonstrate more adaptive and realistic attitudes about workplace violence and the possibility of reducing its incidence.
6. After attending a three-day training unit as part of their first year training programme student nurses will demonstrate more adaptive and realistic attitudes about the role of the student nurse in the management of violent incidents.
7. After attending a three-day training unit as part of their first year training programme student nurses will assess themselves as being more competent in interacting with aggressive clients.

8. Desired changes observed on completion of the unit will still be detected three-months later following two short clinical placements.

Before setting out the strategy adopted for testing the above hypotheses it is firstly necessary to inform about the general nature of training evaluation, and also to describe the particular circumstances of the nurse training course to which the evaluation was applied.

Therefore, Chapter 2 will cover the field of training evaluation and evaluation research, while Chapter 3 will present an overview of the specified nurse training curriculum and course. Only then will the focus move on to methodology in Chapter 4.

CHAPTER 2 - LITERATURE ON EVALUATION OF TRAINING

The previous Chapter described the context of the training being studied and indicated the relative merits of similar previous studies. However little was said about the broader issues associated with training evaluation or the specific frameworks and models used to conduct evaluation research.

This Chapter will firstly review the concepts of *evaluation* and the more recent development, '*evaluation research*' by considering its definition, historical development and evolution, types, and processes. This material will then be applied more specifically to the evaluation of training courses including models of training evaluation. It is hoped that the consideration of various models and framework options in this Chapter, aligned against the cited limitations and shortfalls of previous studies detailed at the end of Chapter 1 will assist in the comprehension and clarification of the research design selected for this study [to be detailed in Chapter 4].

2.1 DEFINITION OF EVALUATION

Evaluation is something that we all do all of the time. The term 'evaluation' has a clear usage within everyday language and can be seen as "an elastic word that stretches to cover judgements of many kinds" (Weiss, 1972: 1). An apparently straightforward statement such as 'United played well on Saturday', or a more personal one like 'that

jumper suits you', each belie the mental effort and processes that contributed to their formulation. Yet, it is possible to readily identify many of the main concepts associated with evaluation when one examines each of these seemingly trivial utterances a little closer.

For example, they imply that specific relevant information has been gathered for a particular purpose, that it has been collated and that judgments have been made in relation to some 'standard'. The first statement suggests that the football team played well in comparison to their usual level of performance or to that of their competitors. The second statement implies that the overall appearance of the jumper is more attractive than others that the person wears.

Furthermore, the examples illustrate the possible wider implications of, and pressures on, evaluation. In the first case, particularly if issued by an influential sports reporter, there could be major 'business' implications in terms of share prices, future crowd attendance numbers, club income, and team and individual player valuations. In the second case one would need to consider the motives, intentions and risks taken by the evaluator. In other words, evaluation can be a 'political' activity that impinges on a number of 'stake-holders' and can have enduring implications for managers, workers, shareholders and the researchers themselves.

Patton (1981: 35) observed “no single-sentence definition will suffice to fully capture the practice of evaluation”. This caution has not stopped many authors attempting definitions, and this section will review some of the definitions of evaluation/ evaluation research. However, the intention is to dissect and emphasise the concepts that they contain rather than list many verbatim definitions, hence a couple are thought to suffice. In so doing this section will draw on the comprehensive review performed by Clarke (1999).

It is immediately apparent that a large proportion of the leading texts are North American in origin, where evaluation has a longer history of specialist study. For example,

Evaluation research is the systematic application of social research procedures for assessing the conceptualization, design, implementation, and utility of social intervention programs. (Rossi and Freeman 1993: 5)

and,

Program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgements about the program, improve program effectiveness, and/ or inform decisions about future programming. (Patton 2002: 10)

The definitions emphasize different aspects of evaluation through the use of the terms evaluation, evaluation research and programme evaluation. The main characteristics of the evaluation process can be readily identified from these definitions.

- Evaluation is seen as being applicable to a wide range of topics and as having a wide range of purposes (Robson 1993).
- Evaluation is seen as a systematic approach that has an associated scientific rigour within its overall design and data collection methodologies. Rossi and Freeman assert that a 'commitment to the "rules" of social research is at the core of our perspective on evaluation' (1993: 6). More specifically, it is viewed as a form of 'applied social research'. Indeed Clarke (1999) describes applied social science as the dominant paradigm in evaluation studies.
- Evaluation is concerned with merit judgments and determining the 'worth or value' of something against previously specified goals or criteria. This could present a tension for the evaluator with the previous characteristic of scientific rigour within which neutrality is sought. In theory the evaluator remains value neutral whilst interacting with interested parties who may be directly affected by the results of the evaluation. Some definitions restrict themselves to measuring expected goals or outcomes whilst

others also include an explicit consideration of the *process* of goal achievement or of *unexpected* outcomes or consequences.

- The judgement characteristic also introduces the idea that evaluation is a political activity (Rossi and Freeman, 1993) since it raises the issue of vested interests (Clarke 1999) and begs the question as to who defines the goals, criteria, and standards that will be evaluated. More fundamentally it questions the researcher-researched relationship and leads to a consideration of which questions, from the many potential ones that could be asked about a programme, will [possibly] be answered and hence the wider concept of 'stake-holders'(Guba and Lincoln 1989; Pawson and Tilley 1997; Rossi *et al* 1999).
- The primary purpose of evaluation is to produce information that will be used by someone ((Rossi and Freeman, 1993; Robson 1993;). Often the evaluation considers the implementation or effectiveness of a particular educational or social programme. The commissioners of the evaluation may then subsequently use the results to inform practice and determine the future development [or termination] of that programme.
- The definitions do not mention a purpose of evaluation as being to generate new knowledge, as is the case with basic research. This fundamental distinction between basic research and

evaluation research has been highlighted by several authors. For example, Stufflebeam and Shinkfield [cited by Clarke 1999] asserted that 'the most important purpose of evaluation is not to prove but to improve' (1985: 151). Similarly, according to Weiss (1997: 516) 'evaluation, unlike the basic sciences, does not aim for "truth" or certainty. Its aim is to help improve programming and policymaking'.

Whilst some authors use the terms 'evaluation' and 'evaluation research' interchangeably, others suggest that *evaluation* refers to the specific task of determining the worth of an intervention or programme, whereas *evaluative research* is more to do with using scientific methods to explore possible cause-effect relationships and determine how a course works. For example, Patton (1986) makes a distinction between *evaluation* and *evaluation research* in that evaluators use research methods to gather data on programmes with the intention of furnishing course managers with information for making decisions about the future development of those programmes, an emphasis on "utility, relevance, practicality, and meeting the information needs of specific decision makers" (1986: 15). By way of contrast, in *evaluation research*, according to Patton, "there is relatively greater emphasis on generalizability, causality, and credibility within the research community" (1986: 15).

Earlier Weiss (1972) had compared evaluation and research and identified differences and similarities. Differences not already mentioned included 'programme derived questions (Weiss 1972:6), that is to say the questions are derived by the decision-maker rather than the evaluator, 'role conflicts', a reference to the various stake-holders involved in a programme and 'publication', wherein basic research is published and dissemination is unquestioned but "probably the majority of study reports go unpublished" (Weiss 1972:7). Weiss (1972:8) also notes important similarities, for example, evaluators using the whole gamut of research methods and designs.

2.2 HISTORICAL DEVELOPMENT AND EVOLUTION OF EVALUATION

Several high profile authors have attempted to document the long-term historical development of evaluation and its more recently evolution from "commonsense program evaluation into 'evaluation research', a heterogeneous mix of substantive issues and procedures" (Berk and Rossi 1990:8). Guba and Lincoln (1989: 22) make the point that evaluation, as we understand it, did not appear one day but is the result of a developmental process of "construction and reconstruction that involves a number of interacting influences". In their review they identify a total of four major developmental phases or 'generations' during the last century or so.

Several reviews of evaluation commence at the beginning of the Twentieth Century with educational evaluation of literacy in school children and the creation of the IQ test by Binet, or health evaluation of control of infectious diseases. Others offer a longer lineage, tracing the first attempts at what would eventually be called 'evaluation research' to the origins of modern science over three centuries ago (Rossi *et al* 1999). At this time Thomas Hobbes and his contemporaries attempted to create numerical measures for assessing social conditions and studying the causes of mortality, morbidity, and social disorganization.

It is suggested that social experiments, "the most technically challenging form of contemporary evaluation research" (Rossi *et al* 1999:4), can be traced back over a similar time period. An early example concerned the 18th century British ship's captain who observed both the lack of scurvy among sailors serving on Mediterranean naval ships, and the citrus fruit included their rations. Subsequently, he experimented, half his crew consume limes while the rest ate as usual, and successfully demonstrated a preventive effect.

Most authors agree that, unsurprisingly, the evolution of evaluation research in the last two century has mirrored contemporary forces and developments in both science and politics, for example,

- the phenomenal rise in the profile of social science driven by the earlier call of John Stuart Mill for social science to follow the

success of the natural sciences, and illustrated by the impetus to education offered by the thesis of Charles Darwin about small differences having significant functional differences over time.

- the emergence of the scientific management movement in business and industry. This 'time and motion' approach with its emphasis on efficiency, effective management of the major human resource and testing in production spread to the field of education where pupils were seen as raw material to be processed by the educational system.
- In the 1930s several social science disciplines were involved in programme evaluation research. Within psychology, during this pre-World War II era, landmarks include the seminal work of Lewin on 'action research', Lippitt and White's work on leadership styles and studies on worker productivity at Western Electric which led to the identification of the 'Hawthorne effect' (Rossi *et al* 1999).
- World War II provided a substantial boost to the opportunities for program evaluation and noted examples are the work of Stouffer in developing procedures for monitoring soldier and civilian morale, and evaluating personnel policies and propaganda techniques (Stouffer *et al* 1949). Furthermore smaller studies

considered the efficacy of price controls and media campaigns to modify eating habits.

- The 'boom period for evaluation research', according to Rossi *et al* (1999), was post World War II to the 1980s. During this time major social programmes were launched, both nationally [urban development and housing, technological and multi-racial education] and internationally [family planning in Asia, agricultural and community development on Africa]. All these programmes demanded evaluation with increasingly sophisticated sampling and survey designs and statistical procedures.

The 1960s saw an increase in academic output, and it is during this time that the numbers of publications, books, journals and journal articles grew rapidly. Rossi *et al* (1999) include the works of Suchman (1967) and Campbell (1969) as being of special merit. The 1970s saw evaluation research emerge as a separate specialist field in the social sciences with its own textbooks, for example Weiss (1972).

Rossi *et al* (1999) acknowledge that the period of rapid growth is now over but, nevertheless, assert the continuing importance and necessity of specialist evaluation within the social sciences research panoply.

A number of challenges remain for evaluation research, including

- Accommodating the recent sea change involving the significant influence of consumers and other stakeholders in determining the scope of any evaluation, and
- the political nature of the evaluation process and use of results (Guba and Lincoln 1989, Rossi *et al* 1999)

Pawson and Tilley (1997) provide an alternative review of the historical development of evaluation research but this time from a more cynical British sociological perspective. Their “history of evaluation in 28 ½ pages” is an intellectual analysis that commences in the 1960s but recognises many of the trends already mentioned above. They argue that this evaluation movement failed to deliver the expected goods to policy-makers and managers, not least because pursuit of the supposed strengths of experimental design and objectivity stripped the studies of cumulative, contextual generalisability. As a consequence, major reviews of social programmes -for example, Martinson's much-cited review of over 200 penal reform programmes (Martinson 1974) and Ford's review of mental health professional studies between 1960 and 1978 (Ford 1979) - reported disappointing, equivocal, non-cumulative findings.

Dissatisfaction eventually produced a series of alternative developments, although these are viewed as predominantly confined to the U.S., whilst in the U.K. “until very recently, there has been significantly less movement from the quasi-experimental orthodoxy”.

Pawson and Tilley (1997:11). Attempted accommodation of the qualitative, interactionist, phenomenological approaches in a form that Pawson and Tilley (1997) call 'constructivism' is summarily dismissed by Pawson and Tilley (1997:20) not merely because of its "misplaced pseudo-scholarship" but, moreover, because it indicates "a more deep-seated air of unreality about the evaluations-as-negotiations perspective... (and)... failure to appreciate the asymmetries of power".

Unsurprisingly, Pawson and Tilley use the assertion of the pendulum swinging too far in one direction as the springboard for their own vision – realistic evaluation. This is a perspective which combines "the rigour of experimentation with the practical nous on policy making of the pragmatists, and with the empathy for the views of the stakeholders of the constructivists" (Pawson and Tilley 1997:24).

2.3 EVALUATION: TYPES AND PURPOSE

The intention in this section is to, firstly, provide an overview of different types and purposes of evaluation and, secondly, locate the chosen evaluation research design within these frameworks. Many different types of evaluation have been identified, indeed Robson (1993) cites Patton (1981) as listing over a hundred types. Each main *type* looks at different evaluation *questions*, tends to focus on different *aspects* of evaluation (Robson 1993), and utilises *research methods* and *data collection techniques* from the various research paradigms.

A popular way of breaking into this complex inter-relational arrangement is to ask questions at the outset of the evaluation that are designed to reveal its intended scope, and the questions it is, in turn, required to answer. Several authors offer suggestions for these sorts of questions. Within the evaluation of training context Reid and Barrington (1997) suggest asking five questions:

1. Why is the evaluation required?
2. Who should do it?
3. What aspects should be evaluated and when should this be done?
4. What kind of measurement will be used?
5. When will it be done?

These questions are of fundamental importance and, certainly, questions 3 and 4 will be explored in much greater detail later [question 3 in this chapter when models are discussed, and question 4 in the chapter on research methods]. A more elaborate version is offered by Lee and Sampson (1990) which includes 10 questions, as follows,

1. What is the programme to be evaluated?
2. Why is the programme being evaluated
3. How are people to be prepared for the evaluation?
4. What are the main issues/questions with which the evaluation is to deal?

5. Who will do what?
6. What are the resources for the evaluation?
7. What data need to be collected?
8. How will the data be analysed?
9. What will be the reporting procedure?
10. How will the report be implemented?

The list gives a clear indication of the points of interaction of course design and evaluation and illustrates the point that evaluation considerations are best built in to the original design of a programme and considered from the outset. Robson (2000) offers nine questions and illustrates more clearly how the questions can be linked to four broad types of evaluation, namely, needs, processes, outcomes and efficiency. Box 2.1 summarizes these connections.

Bee and Bee (1994:174) suggest that the purposes of training evaluation can be summarized under four groupings, namely,

- To improve the quality of training, in terms of the delivery – methods, length of training, content, level
- To assess the effectiveness of – the overall course, trainer, training methods
- To justify the course – prove the benefits outweigh the costs
- To justify the role of training – for budget purposes, in cutback situations.

Box 2.1: Key initial questions and type of evaluation to be performed
 [adapted from Robson (2000)]

Question	Type of evaluation
1. What is needed? [relates to proposed course]	Needs approach central
2. Does what is provided meet clients needs?	Needs approach central. Processes, outcomes and efficiency could be relevant
3. What happens when it is in operation?	Processes central. Needs could be relevant
4. Does it attain its goals or objectives?	Outcomes approach central. Efficiency could be relevant
5. What are its outcomes?	Outcomes approach central. Efficiency could be relevant
6. How do costs and benefits compare?	Efficiency approach is central. Outcomes is relevant
7. Does it meet required standards?	Any or all of approaches could be relevant depending on circumstances
8. Should it continue?	Any or all of approaches could be relevant depending on circumstances
9. How can it be improved?	Process approach likely to be central but any or all of approaches could be relevant depending on circumstances

Bee and Bee (1994:175) further suggest that all these purposes are interlinked and that "to some extent any evaluation addresses all four

purposes". However it is important to be focused and clear about the *primary* purpose so that the best research approach can be adopted. Likewise, Robson (1993:178) suggests that an evaluation is likely to be concerned with several purposes. Robson (1993) emphasises the relationship between purposes and types of research and cites an American classification constructed by the Evaluation Research Society (1980:3-4) [See Box 2.2] that associates various types of evaluation with purpose and the kinds of activity typically performed in the evaluation.

Robson (1993) suggests that this list has to be viewed as a broad classification of ideal types and possibilities and adds that more specific evaluation activities can be listed within each of these six categories. These include awareness evaluation, cost-benefit evaluation, cost-effectiveness evaluation, criterion-referenced evaluation and quality assurance. Not all authors agree about multi-purpose evaluations. For example, writing much earlier Weiss (1972:15) suggests that the "all-purpose evaluation is a myth", that some purposes of evaluation are incompatible with others and concluded that the evaluator has to make choices.

Box 2.2: Types of evaluation, (alternative names), associated purpose and activity (Evaluation Research Society 1980:3-4)	
Types of evaluation, (alternative names),	Associated purpose and activity
Front-end analysis (Pre-installation, context, feasibility analysis)	Takes place before programme starts, to provide guidance in its planning and implementation
Evaluability assessment	Assesses feasibility of evaluation approaches and methods
Formative evaluation (Developmental, process)	Provides information for programme improvement, modification and management
Impact evaluation (Summative, outcome, effectiveness)	Determines programme results and effectiveness, especially for deciding about programme continuation, expansion, reduction, funding
Programme monitoring	Checks for compliance with policy, tracking of services delivered, counting of clients
Evaluation of evaluation (Secondary evaluation, meta-evaluation, evaluation audit)	Critiques of evaluation reports, re-analysis of data, external reviews of internal evaluations

With regard to training course evaluation, Sanderson (1992:126) presents a more comprehensive set of purposes for evaluation, as follows,

- To determine whether the objectives of training were met
- To determine whether the objectives of training were the right ones
- To improve current and future programmes
- To improve trainers
- To establish the cost-effectiveness or cost benefit of programmes
- To establish the contribution of the training function
- To provide marketing data
- To determine unmet training needs.

Patrick (1992) is another author who indicates the interrelationships between purposes, methods and types when suggesting that the four main aims or purposes underlying training evaluation are

- training programme improvement [covering content, design]
- decisions about trainees [further training needs, debriefing]
- administrative and organisational decisions [quality and value]
- training research [which manipulations or variables improve training]

and that each of these is associated with particular approaches or research designs. For example a systems, scientific research or naturalistic approach is associated with the first category, whilst a cost effectiveness or statistical approach is most obviously linked with the

third purpose. Patrick (1992:524) asserts that the first purpose [training course improvement] is the most important, citing the authority of Cronbach (1969:364) “the greatest service evaluation can perform is to identify aspects of the course where revision is desirable”.

Thackwray (1997) also highlights the interrelatedness of research purpose, research strategy and type, and data collection techniques. Writing about the field of higher education Thackwray (1997:42) concludes that responsive evaluation asks the key question “why should we evaluate?” and may find that several purposes are detected, some open, for example, to measure change in knowledge, attitudes or skills or personal effectiveness, and some less so.

Thackwray offers a brief classification of purposes of evaluation, created by Easterby-Smith (1994). Easterby-Smith (1994) developed four purposes of evaluation:

1. Proving: demonstrating something is happening because of training actions
2. Improving: something becomes better than it currently is
3. Learning: evaluation is an integral part of the development process
4. Control: relating training actions to organisational objectives

In the forgoing discussion the purposes of research have been assumed to be legitimate and overt. However, Suchman (1972) cited by Rutman (1977) provides a cautionary list of covert purposes for commissioning or under-taking evaluation research that relates to the previous 'political' characteristic and has a surprisingly contemporary feel. According to Rutman (1977:26) Suchman termed such studies "pseudo-evaluations" and identified

- Eyewash – a deliberate focus on surface appearance to make it look good
- Whitewash – an attempt to cover up programme failures during the investigation
- Submarine – political use of research to destroy a programme
- Posture – a ritual having little substance [perhaps undertaken as a condition of funding]
- Postponement – using evaluation to postpone needed action

Obviously, not all types can be covered here but some of the most influential ones will now be briefly considered. Several authors (Weiss 1972; Patton 1986; Clarke 1999) credit Scriven (1967) with devising the most popular and enduring classification of types of evaluation within the dichotomous terms *formative* and *summative* evaluation. Although originating in the categorisation of educational curricula, the terms now apply more generally to programme evaluation.

Essentially, formative evaluation is used in-house during the early stages of course development to provide trainers with feedback and “support the process of improvement” (Scriven 1991:20). The concern is to ascertain the perceptions and considerations of educators and course attendees on course design and implementation so that current strengths and weaknesses can be identified, and improvement and development enacted.

In the case of summative evaluation, the intention is to “determine the overall effectiveness or impact of a programme or project, with a view to recommending whether or not it should continue to run” (Clarke 1999:8). Patton observed, 'summative evaluations tend to be conclusion-oriented whereas formative evaluations tend to be action-oriented' (1986: 66).

It is possible to detail the essential elements of these two evaluation approaches as two ideal types located at opposite ends of a continuum. Clarke (1999) summarizes these extremes in a format adapted from Herman et al. (1987) and reproduced below [see Box 2.3]. It can be seen that the type of research can have major implications for the role and relationship of the researcher to others, the overall research methodology, data collection techniques, and the frequency and scope of any summary reports. A detailed discussion of the issue of methodology will be reserved for later [Chapter 4] but Box 2.3 highlights a clear distinction with summative evaluations generally expected to be

quantitative. However, of late, no superiority of summative over formative is suggested, and, selection would be dependent on intended purpose.

Box 2.3: The main features of the formative /summative evaluation dichotomy		
	Formative	Summative
Target audience	Programme managers/ practitioners	Policy-makers, funders, the public
Focus of data collection	Clarification of goals, nature of implementation, identifying outcomes	Implementation issues, outcome measures
Role of evaluator	Interactive	Independent
Methodology	Quantitative and qualitative (emphasis on the latter)	Emphasis on quantitative
Frequency of data collection	Continuous monitoring	Limited
Reporting procedures	Informal via discussion groups and meetings	Formal reports
Frequency of reporting	Throughout period of observation/ study	On completion of evaluation

Whilst the simplistic delineation of the main features of formative and summative evaluation presented in Box 2.3 is helpful, it is important to remember the middle ground not identified within the representation. Indeed, aspects of both forms of evaluation could be combined in a single study (Clarke 1999) and it will later be seen that the evaluation study to be detailed and reported within this work does not sit entirely

within either column but contains features of both formative and summative evaluation.

Furthermore the study will also go on to present knowledge that could be useful in addressing some of the more basic questions to do with measuring changes underlying training programme evaluation. This sort of activity has previously been identified by Patton (1996) as 'knowledge-generation for conceptual use' and clearly falls outside Scriven's formative/summative division.

Rossi *et al* (1999) provide a helpful exhibit containing work by Pancer and Westhues (1989) that has been adapted in Box 2.4. The listing shows clearly the inter-relationship between the stages of developing a programme [social or educational], the key questions to be asked, the purpose of the evaluation and the appropriate type of evaluation design. Within this framework Scriven's *formative* evaluation would relate to Stages 1-6, wherein the programme is being created and initialised. In Box 2.4, Stage 7 relates to *summative* evaluation that occurs some time after the programme has been implemented.

Box 2.4: Stages of program development and related evaluation purpose [adapted from Pancer and Westhues (1989)]		
Stage of programme development	Question to be asked	Evaluation purpose
1. Assessment of problems and needs	To what extent are needs and standards met/ unmet	Needs assessment; problem description
2. Determination of goals	What must be done to meet needs & standards	Needs assessment; service needs
3. Design of programme alternatives	What service could produce the desired changes	Assessment of programme logic or theory
4. Selection of alternative	Which of the possible programme approaches is best	Feasibility study
5. Programme implementation	How should the programme be implemented	Implementation assessment
6. Programme operation	Is the programme operated as planned	Process evaluation; programme monitoring
7. Programme outcomes	Is the programme having the desired effects	Outcome evaluation
8. Programme efficiency	Are programme effects attained at a reasonable cost	Cost-benefit or cost-effectiveness analysis

As previously mentioned the primary purpose of any evaluation largely determines the research methodology adopted. Robson (2000:54) asserts that "outcome evaluations tend to be quantitative. The evidence which many find most acceptable about outcomes is in numerical form".

Notwithstanding the dissatisfaction voiced by some evaluators of late about quantitative methods and the growth in popularity of qualitative approaches [see the historical review earlier in this chapter], the experimental design is still highly esteemed and remains the only method of attempting to prove causality, that is to say that a particular aspect of learning was caused by attendance on the course and nothing else. The primary intention is to achieve and maintain validity in the experimental design. Campbell and Stanley (1966) and Cook and Campbell (1979) provide authoritative accounts of the causes of invalidity in designs that will be expanded in Chapter 4, along with the other terms and concepts outlined here.

The goals and objectives of a programme offer clear guidance about what should be measured but process evaluation, in the form of observation or conversation/informal interview with protagonists, may well shed light on possible unintended outcomes. Hence a variety of research methodologies and data collection methods could be appropriate. This emphasis on unintended outcomes is given even greater importance in the light of frequent findings of only marginal change or equivocal outcomes in studies or meta-analyses discussed earlier.

2.4 EVALUATION OF TRAINING COURSES

This section will review some of the key influential models of training evaluation devised over the last forty years or so. After doing so,

damning findings from several published studies and associated statements will be presented in order to illustrate how little evaluation of training is routinely performed within organisations.

Without doubt Kirkpatrick's model of training evaluation (1958/9; 1976) is still the most influential and commonly used (Kraiger *et al* 1993; Kraiger and Jung 1997; Quinones 1997; Thackwray 1997). Credited with revolutionising the thinking on training course evaluation, it remains the only model that many organisations and training departments are aware of (Thackwray 1997). For these reasons it will be considered at some length.

The model is structured around four levels, each measuring complementary aspects of a training course. The levels are entitled Reaction; Learning; Behaviour; and Results. The Reaction level considers the immediate subjective opinions of participants about a course, what they liked/disliked about a course and is equivalent to measuring their feelings. Crucially, it should be noted that this level does not measure any learning that has taken place. Reactions are comparatively easy to measure so tend to be the first [and too often only] resort.

Training managers design anonymous self-completion 'happy sheets' and obtain feedback on specific aspects - what was liked, useful, etc. - along with spaces for additional comments. At their worst they may

merely collect information about the standard of catering or accommodation. At their best they could provide illuminatory material about the interaction between trainers and trainees and the trainee's experience of completing the course, along the lines of process evaluation previously identified. Bee and Bee (1994) suggest that, subject to course participant fatigue, these questionnaires can be administered at various intervals- after each session, daily, weekly, at the end of a module etc. depending on the stage of development of the particular course [harking back to the formative-summative discussion considered earlier].

The Learning level relates to the extent to which trainees achieved learning objectives and absorbed the knowledge and skills delivered in the course, although 'learning' is poorly defined (Thackwray 1997; Kraiger and Jung 1997). Kirkpatrick suggests that many aspects of experimental research design are useful at this level. For example, using a before-and-after approach, measuring each learner in an objective manner to produce quantitative results, using a control group where possible, using statistical analysis to identify correlation or levels of confidence. Behavioural tests are advocated for skills, and multi choice tests for knowledge.

The Behaviour level refers to 'training transfer' - the results of training in terms of on-the-job performance back at work. It is acknowledged that this is more demanding than the previous stages. There are difficulties

in attempting measurement at this level, ranging from methodological issues to the necessity of certain personality characteristics in course attendees. Indeed, Kirkpatrick identified the following personality characteristics:- individuals wanting to improve; recognising their own weaknesses; working in a permissive environment; support from interested others; and opportunity to try out new ideas. The same components of experimental research design as previously mentioned were recommended for designing evaluation at this level.

The final level is 'Results' and this refers to the impact of training on the department or organisation in terms of performance or profitability.

Kirkpatrick (1959) acknowledged this to be the most difficult form of evaluation, given the many extraneous influences on organisations, and advocated a participative approach with the incorporation of peer and self-assessment.

The model encourages evaluators to progress through the levels in the specified order until unsatisfactory results are revealed. Apparently, the underlying intentions were to raise the aspirations of training managers, to increase their efforts, and to encourage them to "gradually progress from a simple subjective reaction sheet to a research design that measures tangible results" (Kirkpatrick 1976:26). Kirkpatrick (1994) has continued to expound the model and publish case studies of its use, however, critics of the model suggest that it is now decidedly dated. It has been useful in guiding thought about how to evaluate, and indeed,

still remains a “valuable heuristic for evaluating training outcomes” (Quinones 1997:179) but is restricted in helping make decisions about what to evaluate or how to convert evaluation results into decisions about future training (Kraiger and Jung 1997).

Thackwray (1997) suggests several specific criticisms related to this point. For example, that the four level model does not necessarily produce results of use over the longer-term, that it implies a ‘standardised process’ which is clearly not the case, and that it over-emphasises evaluation as a *fact*-producing process instead of generating information on achievement of targets useful to management. The suggestion of a hierarchy of levels with some levels displaying superiority over others has never fully been refuted.

Another identified limitation of the model (Kraiger *et al* 1993; Kraiger and Jung 1997) concerns its assumption that the levels are causally or sequentially linked, although this is also a criticism of other models [see later]. For example, learning being “conceptualised both as a causal result of positive reactions to training and as a causal determinant of changes in trainee behaviour” (Kraiger *et al* 1993:311). The belief that we have to ‘like’ a course or a presenter in order to learn the most from it/them is also unproven, seemingly suggesting that a slick, visual entertaining presentation of empty, unmeaning content can lead to learning.

The model also lacks clarity about operationalising the different levels of measurement, i.e. how objectives would be measured, and fails to suggest appropriately varied methods for evaluating different levels (Kraiger *et al* 1993). Kraiger *et al* (1993) suggest that it is unclear as to whether Kirkpatrick viewed learning skills and knowledge facts as synonymous and, therefore, whether the same assessment tools were appropriate for both. Furthermore, it is indicative of the behavioural stimulus-response type models of learning which were popular in the 1950s and 1960s and lacks modern ideas or psychological theories about how people learn or acquire skills (Kraiger & Jung 1997), or take account of modern teaching methods, for example, computerized technologies.

Notwithstanding the above criticisms and limitations, the model remains highly influential in organisations. Thackwray (1997:33) concludes that “if all organisations in the U.K. at least followed Kirkpatrick, billions of pounds would be saved each year”.

There have been several developments of Kirkpatrick's model and two more will now be summarised. Warr *et al* (1970) also devised a model with four levels that incorporated the features evident in Kirkpatrick's model but extended the evaluation to also include a consideration of the context of the evaluation and the inputs or resources available.

Sanderson (1992) suggests this model offers a broader perspective, which views evaluation as a continuous process commencing with the

needs analysis, and inter-related with the subsequent stages of course design and programme delivery.

Warr *et al's* model has the acronym CIRO, being the first letters of each of the levels - Context evaluation, Input evaluation, Reaction evaluation and Outcome evaluation. 'Reaction' has the same meaning as described in Kirkpatrick's model, whilst 'Outcome' is subdivided into 3 sub-levels that correspond to Kirkpatrick's other three levels but in this case labelled *immediate* [knowledge, skills and attitude development at the end of training], *intermediate* [changes in on-the-job performance] and *ultimate* [desired changes in the organisation] (Sanderson 1992). 'Context' refers to obtaining information on the operational situation in order to clearly identify needs and hence learning objectives. 'Input' concerns the best method of delivery taking in to account time scales, in-house resources, level and types of input, financial resources available (Thackwray 1997).

Illustrating the hierarchical bias that is identified in the models of Kirkpatrick and Warr *et al*, Sanderson (1992:129) concludes "both frameworks view the last level of evaluation [results or ultimate] as the most difficult, the least often done and the most valuable. Reaction evaluation is the easiest, the least useful and the most frequently used method". Similarly, Reid and Barrington (1997) suggest that unless evaluation is completed at lower levels it will not be possible to identify the cause of failure identified by evaluation performed at a higher level.

Given the multitude of possible extraneous factors that impinge on departments and organisations and potentially confound any predicted change, Sanderson (1992) cites sensible advice from Warr *et al* (1970). They conclude that it is prudent to concentrate on the lower levels [context, input, reaction, immediate outcome] and assume that if these have been completed properly then intermediate and ultimate level outcomes are likely to be successfully accomplished.

A possible exception may be where most of, or an entire, department undergoes training development. In this case, evaluation of intermediate level outcomes may be practicable. Similarly, the importance and centrality of a training programme in terms of likely costs, number of trainees, number of repetitions, length of training, support offered by line managers etc. may also make intermediate level evaluation of a training programme desirable.

Hamblin (1974) further developed the ideas of Kirkpatrick and Warr *et al* [reactions, immediate or learning, intermediate or job behaviour, ultimate or results] creating a fifth level by effectively dividing the fourth level into two. This allows the evaluator to distinguish between the outcomes for the organisation in terms of productivity, sales, absenteeism etc., and the effects on costs in terms of a cost-benefit or cost effectiveness analysis (Bee and Bee 1994).

Each of Hamblin's five levels [reactions, learning, job behaviour, organisation, ultimate value] contains both objectives and effects and it is suggested that lower-level effects are pre-requisites for higher levels ones in the now familiar hierarchical formation. The levels are viewed as forming the links in a chain that can snap at any point (Patrick 1992). Hence learning has to occur before behaviour can change but learning may occur without any resultant behaviour change. The model incorporates a consideration of indirect factors at each level [accommodation, cuisine, conversations with others] but is least convincing in its pre-requisite linking of reaction with learning, and organisation with ultimate effects. Once again, evaluation is easiest at the lowest levels (1 and 2) and becomes increasingly difficult as the levels rise (Reid and Barrington 1997).

Finally, the calculation of cost-benefit or cost-effectiveness of training is difficult and there are few good examples in the literature. Patrick (1992) suggests that there are practical and technical difficulties in performing such evaluations. For example, collecting cost or effectiveness data solely attributable to the training is difficult, and it could be that the cost of performing the evaluation exceeds the financial return from the training.

Within the higher education sector the costs and benefits of training are notoriously difficult to quantify. Notwithstanding the difficulties associated with separating out the effects of formal training or education

from other sources, it could be argued that in, say, the 'higher education sector', the benefits could be spread over a lifetime in terms of the direct or indirect consequences for society, clearer thinking, and more considerate social behaviour (Patrick 1992).

Reid and Barrington (1997) assert that most training in the private and public sector occurs in busy working environments. In this 'real world' a strict scientific approach with pre-and post-training tests, control and experimental groups etc. is impracticable and yet it is important to know how effective training has been. According to Reid and Barrington (1997) the solution to this dilemma involves the following three steps:-

1. Set clear training objectives, expressed as far as possible in behavioural terms or competences, specifying the performance evidence and range
2. Include objectives for each level of evaluation
3. Evaluate systematically at as many levels as practicable to obtain the total picture.

One further model will now be presented, that of Kraiger *et al* (1993).

This model is not viewed as a linear development of Kirkpatrick and all that has gone before. In this case its emphasis is rooted in the contemporary psychology of learning. Given that the model forms a key part of the research design and will be used throughout the study, it will be discussed in some depth now.

Kraiger *et al* (1993) offer “a theoretically driven definition of learning along with a preliminary classification scheme for selecting evaluation measures given knowledge of learning outcomes” (Kraiger and Jung (1997:153). The model places a key emphasis on ‘learning outcomes’, which, according to Gagne(1984), illustrate the sort of categories that describe the anticipated results of training. In this way the model guides the evaluator in determining how to measure changes and thus overcomes one of the major criticisms of Kirkpatrick’s model. Furthermore, the model criticises the “simplistic and uni-dimensional” approach of Kirkpatrick and others (Kraiger *et al* 1993:312) and instead incorporates theoretical and research-based work from a number of diverse sources, including Bloom (1956), Krathwohl, *et al* 1964 and Gagne (1984) in developing more expansive taxonomies of learning outcomes.

This development extends the range of possible learning outcomes that can be evaluated by proposing three types or categories of learning outcomes [cognitive, skill-based and affective] and sub-dividing each of these in to a number of categories and constructs. For example, cognitive outcomes are sub-divided in to verbal knowledge, knowledge organisation and cognitive strategies, whilst skill-based outcomes include compilation and automaticity. Affective learning outcomes are represented by attitudinal and motivational outcomes, which include aspects of self-efficacy and goal-setting.

These differing constructs reflect the complexity or stages of development of each category. The proposed model explores each of these categories and presents a cogent review of relevant educational psychology theory. Furthermore it also identifies the likely foci for measurement and suggests appropriate training evaluation methods. Box 2.5 [adapted from Kraiger *et al* 1993] summarises this classification.

Obviously the model cannot be expounded in minute detail but, as an illustration, Kraiger *et al* (1993) suggest that cognitive outcomes can be constructed around the recall of facts and declarative knowledge, around the internal structuring and organisation of knowledge [as in the terms 'mental models' or 'mental maps'] and cognitive strategies " the extent to which knowledge can be accessed or applied more rapidly or more fluidly" (Kraiger *et al* 1993:315). This last category is obviously at the highest level and distinguishes 'meta-cognition', the ability to accurately monitor our own thinking and know when, say, problem solving is unlikely to work, or when we need to revise a particular way of working.

Box 2.5: Classification of learning categories, constructs, foci of measurement and potential training evaluation methods [adapted from Kraiger *et al* 1993]

Category	Learning constructs	Focus of measurement	Potential evaluation methods
Cognitive			
		Outcomes	
Verbal knowledge	Declarative knowledge	Amount of knowledge. Recall accuracy Speed of access to knowledge	Recognition and recall Power tests Speed tests
Knowledge organisation	Mental models	Similarity to ideal. Interrelationships of elements. Hierarchical ordering	Free sorts Structural assessment
Cognitive strategies	Self-insight Meta-cognitive skills	Self-awareness Self-regulation	Probed protocol analysis Self report Readiness for testing
Skill-based			
		Outcomes	
Compilation	Composition Proceduralization	Speed and Fluidity of performance Errors. Chunking Generalisation Discrimination Strengthening	Targeted behavioural observation Hands-on testing Structural Situational interviews
Automaticity	Automatic processing Tuning	Attentional requirements Available cognitive resources	Secondary task performance Interference problems Embedded measurement
Affective			
		Outcomes	
Attitudinal	Targeted object [e.g. safety awareness] Attitude strength	Attitude direction Attitude strength Accessibility Centrality Conviction	Self-report measures

Motivation	Motivational disposition	Mastery versus performance orientations Appropriateness of orientation	Self-report measures
	Self-efficacy	Perceived Performance capability	Self-report measures
	Goal setting	Level of goals Complexity of goal structures Goal commitment	Self-report measures Free recall measures Free sorts

With regard to skills-based learning, Kraiger *et al* (1993) suggest that we need to pass through a number of developmental stages, including goal orientation and linking behaviours sequentially and hierarchically. Skilful performance is recognised by being smooth, rapid and error-free and in addition, the performer is able to maintain parallel activities [do other things] and detect appropriate situational circumstances for varying the skill.

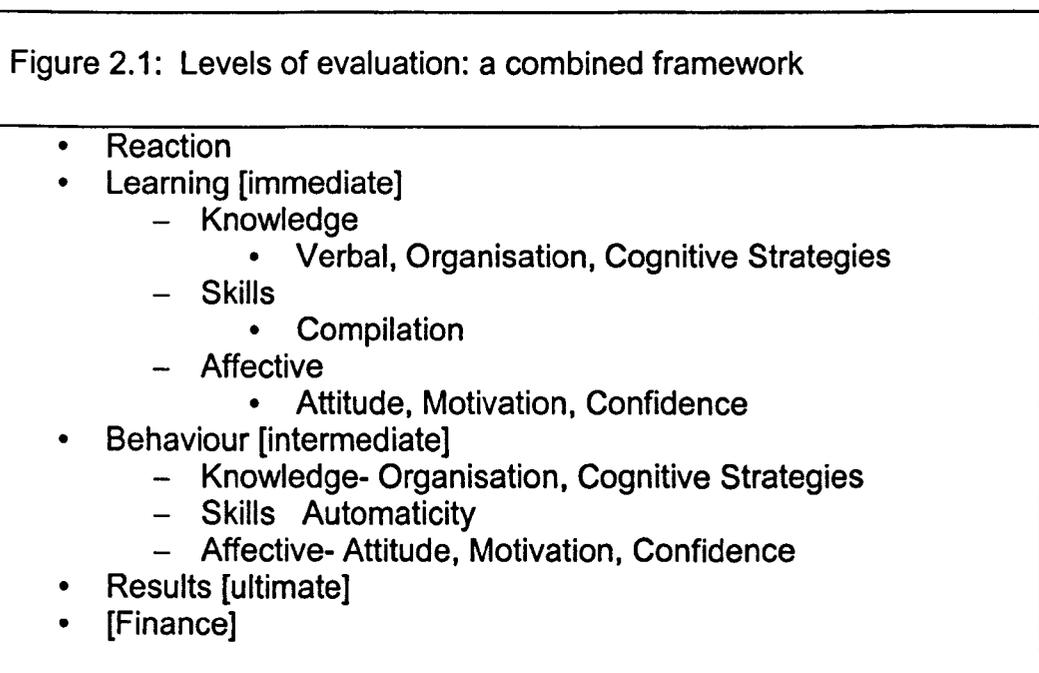
The third category consists of aspects of attitude and motivation, an area of learning different to 'reaction' and completely ignored by Kirkpatrick (Kraiger *et al* 1993:318). It includes elements of self-efficacy, a term which originates from the social-learning work of Bandura (1977) and refers to 'one's perceived performance capabilities for a specific activity'. Self-efficacy is theorised as determining whether an individual is likely to engage with and persist with a particular activity (Bandura 1977).

Hence the model is supposed to overcome many of the deficits identified in the work of Kirkpatrick and followers. It certainly uses the published material derived by learning theorists and researchers, offers clear direction in terms of operationalising change in different categories, and suggests methods by which this change can be measured. Furthermore, the model's emphasis on learning outcomes has helped in refocusing the attention of course managers and trainers on to learning outcomes at every stage of the course design process.

However, the model gives little guidance on how to identify training outcomes given a set of more specific learning objectives (Kraiger and Jung 1997). It also gives no guidance on determining the financial value or cost effectiveness of training, as highlighted in some of the other models previously discussed. Furthermore, it emphasises evaluative aspects of the 'trainee' rather more than the training course [or the employing organisation], underplays the possible delays between training and on-the-job performance improvement, and offers little opportunity to collect and incorporate the immediate subjective reactions and views of trainees [or trainers] in to the evaluation.

In Chapter 1 it was suggested that viewing Kraiger *et al*'s and Kirkpatrick's models as complimentary, rather than competing, creates a framework that could prove helpful to training evaluators in determining a broader range of areas of change and associated measures. The strengths of each model were assembled in to an

integrated taxonomy [Figure 1.2]. A comprehensive range of possibilities are offered, although it is not suggested here that any study would be able to measure changes at all these levels. Furthermore it was suggested that a range of appropriate measures could be used to determine the effects of training at each level and category and in Chapter 1 the range of ways in which the effects of aggression management training could be evaluated were indicated in Figure 1.3 In light of the fore-going discussion on the merits and limitations, the strengths and weaknesses of each individual model, the figure is now re-presented below as Figure 2.1.



2.4.1 Explaining the Infrequency of Training Evaluation

Finally this section will present a summary of evidence that shows evaluation of training remains a distant aspiration in the mind of many training and personnel officers and a lowly priority for managers.

Patrick (1992) cites an early [1977] survey of management training reported by Phillips (1990) in which only one quarter of 3100 executives reported evaluating staff training courses by measurement of subsequent job behaviour, while just over half relied on trainee feedback and a mere 1.8% performed cost-benefit calculations.

Bee and Bee (1994) refer to a more recent British survey that suggested that only 15% of organisations attempted an evaluation of training, with only 2.5% performing a cost-benefit analysis (Training Agency 1989).

In a similar vein, Thackwray (1997) highlights a 1994 survey of a varied range of 467 companies by the Industrial Society that addressed a number of questions to their personnel/training professionals. The survey revealed that almost one fifth of companies did not attempt any systematic evaluation. Eighty percent of the ones that did used reaction sheets, while only 14% used a follow-up line-management questionnaire sometime later.

The main problems identified by respondents were difficulty in establishing measurable results, lack of time, lack of knowledge of evaluation techniques, unclear training objectives, and lack of senior

management support. These findings are consistent with many others in showing that training is evaluated too infrequently, and that, even when it does occur, it tends to be confined to measuring trainee reactions.

The numerous practical difficulties – for example, surrounding the dispersal of trainees to different settings immediately after completion of training, the detection of subtle changes in skill level and job behaviour, assessing the value of different attitudes involves many assumptions, and calculating the costs of aspects of training - can be dealt with by well-established evaluation methods (Patrick 1992). Unfortunately, too often, organisations may lack the political will to evaluate training, and view evaluation negatively so that it is possibly perceived “more as a weed-killer than a plant fertilizer” (Patrick 1992:514).

Adding to this rather depressing list Sanderson ((1992) suggests a number of other reasons commonly expressed to explain the reluctance to undertake evaluation. She identifies

- Evaluation being unnecessary - benefits of training are obvious
- Threat to trainers, revealing inadequacies or ineffectiveness
- Only worthwhile to perform rigorous and scientific evaluation.

Unfortunately these methods are difficult or impossible in the real world so nothing is done

- Lack of skills and incentives by trainers
- Reaction sheets on final day being sufficient

- Uses up scarce monies from training budget that could provide more training
- Evaluation requiring coordination between trainers, managers and administrators that is not possible in practice.

Ghodsian *et al* (1997) suggest that, whilst post-training measures of on-the job performance are potentially the most valid measures of the effectiveness of training, in practice they present several difficulties, and proceed to identify examples of both practical difficulties and institutional impediments to this type of evaluation. For example, practical difficulties include geographical or administrative distance between trainers/evaluators and trainees after course completion, time lag between course completion and receiving evaluation data, lack of necessary in-house skills, and potential sensitivity of findings with consequent apprehension. Institutional impediments identified by Ghodsian *et al* (1997) include active resistance to evaluation driven by fear of liability, and the potential costs of having to retrain personnel if they are shown to be 'substandard'.

Additional, sometimes valid points made include the fact that any formative or on-going review of a course tends to result in changes to content, methods and even some of the objectives before any summative evaluation is completed (Reid and Barrington 1997). Once delivered it is history, a past event, and evaluation is seldom seen as a chance to influence future programmes.

Several authors also state that a thorough needs analysis is necessary prior to the design of a training course but rarely performed. Instead, managers either buy in on the basis of a glossy brochure or personal recommendation (Dipboye 1997), or a knee-jerk response from the training department. As a consequence “many training courses are doomed to failure because trainers are more interested in conducting training programs than in assessing needs of the organisation” (Goldstein 1993:37).

Dipboye (1997) asserts that true or rigorous quasi-experiments are rare, the predominant design [when any evaluation occurs] being simple pre-post or just post-training collection of data. Dipboye (1997:37) adds “organisations typically use only ‘happy sheets’ and ignore whether training has had an impact on learning, behaviour, and the performance of the trainee on the job”.

It was noted in the last chapter that many of these criticisms apply to the very limited number of published course evaluations in the area of aggression management training. Overcoming these practical difficulties, rationalisations and excuses presents a major challenge to anyone interested in improving the quality of training provided in an organisation.

2.5 SUMMARY OF CHAPTER

The focus of the thesis is the evaluation of a management of aggression training unit for student nurses. This Chapter has sought to highlight the issues pertinent to training evaluation and illustrate the range of emphases within a number of established models.

It will become evident that a rigorous evaluation research design can be integrated with the models of Kirkpatrick and Kraiger *et al* to produce evidence of training effectiveness which builds on the author's previous efforts in this area. However, before proceeding with this endeavour in Chapter 4 and subsequent Chapters, the intention is to firstly review the aggression management training unit that became the focus of this study. Chapter 3 will detail this material.

CHAPTER 3: A 3-DAY UNIT FOR STUDENT NURSES IN THE PREVENTION AND MANAGEMENT OF AGGRESSION IN HEALTH CARE SETTINGS

This Chapter presents details on the specific unit offering training on preventing and managing aggression in health care settings that served as the focus of this study. The intention here is to provide an indication of its conceptualization and development along with more precise details of its structure, content, learning outcomes, and teaching methods. Firstly, however, it is necessary to contextualize the Unit, both with regard to the development of nurse education in general, and in relation to the particular course within which the unit was embedded.

3.1 THE DEVELOPMENT OF NURSE EDUCATION

3.1.1 Pre-Registration Nursing – Early Years

Nurse education is big business and a vital component in the maintenance of the National Health Service. It is reported that in 1999 there were 310,000 nurses, midwives and health visitors working in NHS hospitals and community services (NAO 2001). The NAO report also identified a budget of £705 million in 1999-2000 spent by the NHS on training and bursaries in 73 higher education institutions for some 50,000 and midwifery students and 14,000 health professional students [physiotherapists, radiographers etc.] trained (NAO 2001).

However, it would be inaccurate to characterise all the current types of nursing as having the same tradition and identical evolution. In the UK, it is only since 1989 that nurses trained for adult, mental health, learning disability etc. have shared parts of a common course [common foundation programme], or even the same hospital building. Prior to that time the development of each speciality, in terms of education and registration, has taken very different paths.

Formal, nationwide psychiatric nursing training commenced with a common syllabus in 1890, almost thirty years prior to its achievement in general [adult] nursing (Nolan 1990). Common to each of the different specialisms were the training methods, which consisted predominantly of practical work in the clinical setting interspersed with lectures from medical staff and nurses, and the system of assessment. Until the late 1980s training continued to be delivered in NHS-run Colleges of Nursing and Midwifery wherein nursing and midwifery students were considered NHS employees and made a direct contribution to patient care (NAO 2001).

3.1.2 Project 2000

In 1989 a new era of nurse education dawned with the advent of 'Project 2000'. The scheme "represented the most radical overhaul of training in the history of nursing" (Nolan 1993:144). It was necessary in order to address growing concerns about the future of nurse education, for example, the high attrition rate and looming 'demographic time-

bomb' of reduced numbers of suitably qualified school-leaver entrants (NAO 2001), and included a number of new innovations that had been demanded within the profession for some time.

According to the UKCC (1986), listed by Nolan (1993), the scheme was designed to:

- Win, for nursing students, the status and educational opportunities of other professional groups undertaking a vocational education;
- Terminate the practice of immersing students in hospital culture and ward routines;
- Establish links with Higher Education so that nursing education might receive academic validation;
- Improve morale in the profession so that recruitment of the 30,000 new nurses required each year in the UK might be assured and their services retained;
- Place greater emphasis on health promotion and disease prevention than hitherto had been the case.

The realisation of these intentions resulted in a number of developments that pertain to this study, for example,

- Combining the theoretical training of student nurses destined for various different speciality branches for the first half of the three-year course [common foundation programme]

- Offering education as well as training, and substantiated this by presenting a widely-recognised, tertiary education qualification to students who successfully completed the training – a Level Two Higher Education Diploma
- Granting 'super-numerary status' to students for part of the course and allocated a greater proportion of the course to theoretical training in order to produce 'the knowledgeable doer'
- 'Widening the entry gate' by encouraging more men, older women and people from ethnic minorities to apply, and broadening the range of acceptable qualifications in order to facilitate this

Training courses were still guided by National Board syllabi and validated by the National Boards but now con-jointly with Higher Education Institutions, and were still delivered in health services premises by staff employed by regional health service boards.

In the five years that followed there was a consolidation and amalgamation of nurse training institutions associated with a transfer of training into higher education settings. By 1996 all NHS Schools of Nursing had transferred in to higher education (NAO 2001).

Training staff [Tutors] became employees of universities and other higher education institutions [Lecturers], and student nurses became higher education students. Subsequently, more students were offered the opportunity to study nursing at first-degree level, that is, Higher Education Level Three.

3.2 DIPLOMA IN NURSING STUDIES COURSE 1996-2001

The Diploma in Nursing Studies course was a three-year, Higher Education Level Two course which trained nurses in four specialities or 'branches', namely, adult nursing, children's nursing, mental health nursing, and learning disability nursing. It replaced the original Diploma in Nursing Studies course offered by the Department of Nursing, Keele University, which had commenced in 1989 as one of the original pilot areas for the radical Project 2000 training initiative. The course was itself succeeded in January 2001 by a third diploma curriculum, still current, and further developed in September 2001 when a Level Three full-time degree option was introduced.

There were two intakes or cohorts of students per year, at the end of September and end of January, although not all branches had students in both cohorts. The number of students recruited in each cohort was determined at a regional level by a NHS Education and Training Consortium that included representatives from health authorities, NHS Trusts and Primary Care Trusts, social services and other health care employers in the designated geographical area. There had been a reduction in student nurse numbers in recent years prior to the commencement of this particular curriculum [such that some branches were under threat of discontinuation], although during the duration of this curriculum the numbers of students in each cohort increased considerably.

Nationally, there was a 50% increase in the number of entrants to nursing and midwifery pre-registration courses, from 12,480 in 1994-95 to 18,707 in 1999-2000 (NAO 2001), with further increases in 2000-01 when 19,604 students commenced training (ENB 2001). Over time this increase in student numbers necessitated the sub-division of some cohorts within the CFP and consequent duplication of lectures.

The course was divided into two equal parts with an 18-month 'Common Foundation Programme' [CFP] being followed by an 18-month 'Branch programme'. The 3-year programme was divided into nine fifteen-week trimesters, three per year, each with a theme and each trimester having theory and associated clinical placements.

As the name suggests, the introductory CFP was common to all student nurses, irrespective of destined branch, and concentrated on basic biology, psychology, and sociology applied to health care, health studies, inter-personal and communication skills, nursing theory and skills, each subject being taught in large groups [25-50] for much of the time. During the CFP student nurses had two placements in their chosen speciality area [in trimesters 1 and 5], and also shorter placements in other specialty areas. Figure 3.1 shows an overview of the course structure and placements.

Figure 3.1: Overview of diploma course, structure and placements				
-----3 x15 week trimesters per year-----				
Year 1	Trimester 1 Speciality area placement	Trimester 2 Child and midwifery area placements		Trimester 3 Mental health & learning disability area placements
Year 2	Trimester 4 Adult and community area placements	Trimester 5a Speciality area placement	Trimester 5b Speciality area placement	Trimester 6 Speciality area placement
Year 3	Trimester 7 Speciality area placement	Trimester 8 Speciality area placement		Trimester 9 Speciality area placement

	Common Foundation Programme
	Branch Programme

Practicalities associated with course delivery should be borne in mind.

Given the numbers of student nurses on a three year programme at any time [600+], the intricate web of clinical placements being used, each with their own specified maximum quota of students at any time, and the competing demands for lecturer time on other post-registration full- and part-time specialist training courses with the department, any change in a curriculum has to be considered and introduced very carefully.

3.3 TRAINING UNIT: CONCEPTUALIZATION AND DEVELOPMENT

As previously mentioned in Chapter 1, in 1993 the English National Board for Nursing, Midwifery and Health Visiting (ENB), [which published syllabi, controlled curricula and accredited training], published guidelines which specified that all pre-registration training courses for nurses and midwives must include instruction in the theory of aggression and violence and helpful interventions. It further specified that suitably qualified instructors should provide training. In the years that followed these guidelines were interpreted broadly. Nationally, within different University Departments of Nursing, there was wide variation in the quality and quantity of material provided and the status of teachers/ trainers delivering the material.

Within the Department of Nursing and Midwifery at Keele University at that time there was piecemeal inclusion of material throughout the foundation year and specialist two-year branch programmes of the existing Diploma course. All student nurses received guidance on the role of the nurse in aggressive situations and indications of what to do and what not to do, prior to mental health placements. Furthermore, some theory was offered on models of aggression and the context of challenging behaviour. Mental health branch students usually received some instruction and practice of breakaway techniques.

Earlier in 1997 the Department had recognized the importance of the area and supported the training of one of its lecturer's [the author] to

become an accredited trainer with the RCN Institute in the prevention and management of aggression and violence, training which also satisfied the ENB criteria for trainers (ENB 1993). Following completion of the training course a draft strategy document was published in December 1997 that highlighted the current position and proposed a “comprehensive and better coordinated unit of instruction on aggression and violence that includes theories of aggression and violence, health care settings and societal contexts, and that incorporates different levels of intervention skills”. Appendix One provides a summary of this strategy.

Given the aforementioned comments on the possible implications and difficulties of changing a curriculum ‘in mid flight’ a place had to be found for the inclusion of the module. A good argument could easily be made for inclusion of material on managing aggression in trimester 1 of the course prior to the first placement. Unfortunately, the same could be said for many other topics, and, given that the first placement commenced after only three weeks of the course and that the initial weeks were already very full, reluctantly, a different place was sought.

The stereotype of aggression in health care is that it is located in A&E departments, mental health and learning disability, and elderly care settings. The reported incidence figures presented in Chapter 1 bore this out. Figure 3.1 showed that student placements

- for trimester 1 were chosen specialities [each in own destined speciality],
- for trimester 2 child and midwifery, and
- for trimester 3 mental health and learning disabilities.

It was impractical to incorporate the unit in to trimester 1, even though some students would have had placements in mental health, learning disability, and elderly care settings but not A&E departments.

Furthermore, the placements in module 2 were relatively brief [and relatively safe] child and midwifery settings. Hence, it was decided to deliver the unit at the start of trimester 3 as part of the preparation for the mental health and learning disability placements. This arrangement ran the risk of reinforcing the stereotypical view that violence is a phenomenon of only mental health and learning disability settings but it was one of the objectives of the unit to correct this common misconception.

Consequently, the strategy proposed three days of instruction within Trimester 3 of the Diploma course for all student nurses prior to their mental health and learning disability placements. It was anticipated that group numbers would vary on different days between up to fifty and 25 or less, depending on the content, teaching methods and safety issues. More will be said about structure, content and teaching methods shortly.

The strategy was presented to representatives from the relevant NHS Trusts and received overwhelming approval. It was incorporated into Trimester 3 of the existing curriculum and delivered for the first time in June 1998. Not surprisingly, minor changes differences from the strategy occurred during the realisation of the unit.

3.4 TRIMESTER THREE

Trimester three was the point in the Diploma in Nursing Studies course where all student nurses sampled the work of mental health and learning disabilities nurses via a three-week placement in each setting and correspondingly relevant theoretical sessions. Figure 3.2 gives an overview of the typical structure of Trimester 3 following inclusion of the aggression management unit, along with an indication of taught content.

In weeks 1-3 the students were prepared for their forthcoming placements with sessions on meanings, manifestations and effects of mental health and learning disability; interpersonal skills, communication and counselling; biology. These sessions still had to be accommodated, and so it was planned that time previously allocated as 'directed study time' would be utilised to deliver the aggression unit. The two placements and directed study period followed and the trimester concluded with a further three weeks of theory, assessment and evaluation.

Figure 3.2: Overview of the structure of trimester three and content of study weeks	
<p>Weeks 1-3</p> <p>Introduction, placement preparation & aggression management unit</p>	<p>-Introduction to mental health and learning disability , Attending and listening skills, Observation skills, Endocrine system, Counselling skills, Non-verbal communication, Touch, Preparation for placements, [Aggression Unit], Sexual abuse, Challenging behaviour, Mental health assessment, Epilepsy, Manual handling skills update</p>
<p>Weeks 4-6</p> <p>Placement 1</p> <p>Mental health/ Learning disability</p>	<p>A placement on a mental health or learning disability ward/unit</p> <p>May be an elderly care ward or day hospital, an admission ward, a community residential or day facility, a group home or rehabilitation facility Work 15 shifts supervised by qualified staff</p>
<p>Weeks 7-9</p> <p>Guided study using work books</p>	<p>Complete a guide study book</p> <p>Consider aspects of mental health/ learning disability care: Stigma Community-care Institutionalisation, Stressors, Treatments, Diagnosis, Skills teaching, Communication</p> <p>Prepare project presentations</p>
<p>Weeks 10-12</p> <p>Placement 2</p> <p>Learning disability/ Mental health</p>	<p>A placement on a mental health or learning disability ward/unit</p> <p>May be an elderly care ward or day hospital, an admission ward, a community residential or day facility, a group home or rehabilitation facility Work 15 shifts supervised by qualified staff</p>
<p>Weeks 13-15</p> <p>Theory, Consolidation, assessment, and trimester evaluation</p>	<ul style="list-style-type: none"> -Stress -Nervous system -Examination -Dual diagnosis -Loss and change -Placement feedback and presentations -Teaching skills -Attitudes stigma, and labelling -Learning disability and the family -Misuse of drugs -Legal and ethical issues

Over the course of several years the number of student nurses in cohorts increased from approximately 50 to over 80. It became necessary to divide the cohorts in half for many sessions in order to accommodate this increase. In Trimester 3 this arrangement resulted in the material on aggression delivered in the aggression management unit being repeated on consecutive weeks during the preparatory phase of the trimester. Figure 3.3 shows this arrangement.

Figure 3.3: Accommodating increased student numbers in weeks 1-3 of trimester 3			
	Week 1	Week 2	Week 3
Group 1	Introduction, mental health & learning disability	Aggression management unit	Theory – interpersonal communication
Group 2	Introduction, mental health & learning disability	Theory – interpersonal communication	Aggression management unit

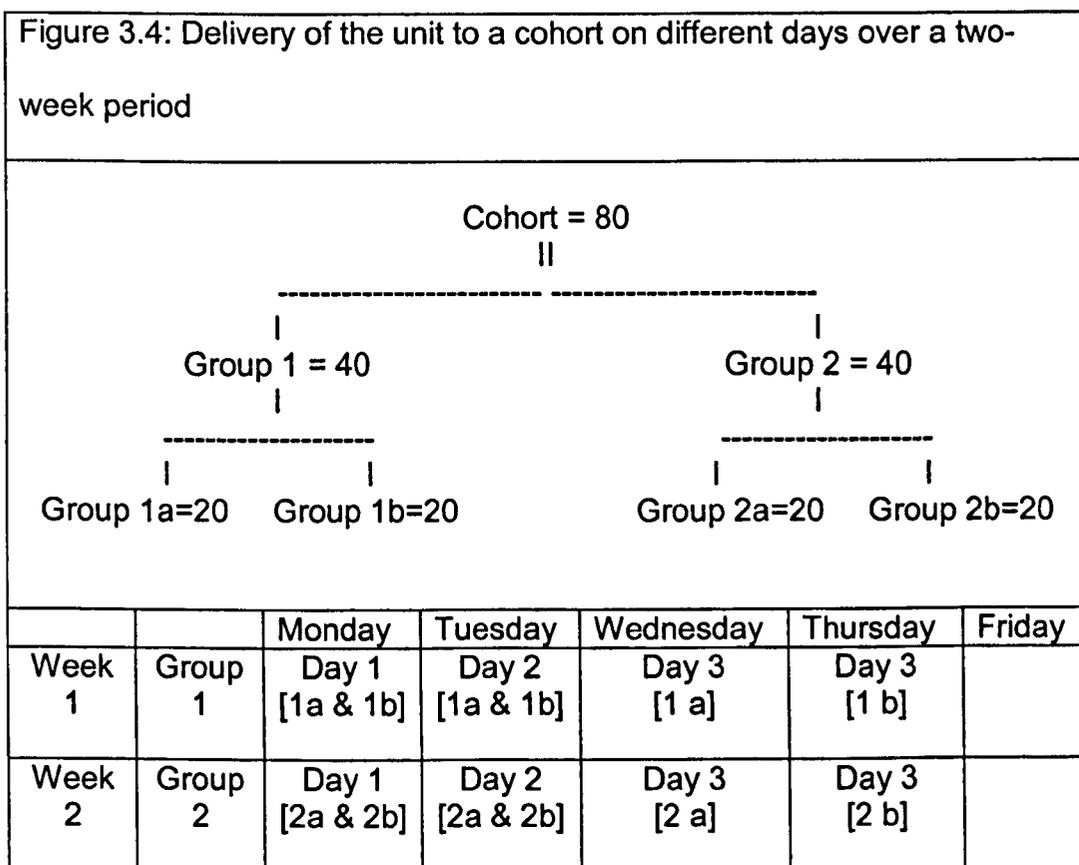
3.5 TRAINING UNIT

3.5.1 Structure

From an individual student nurse's perspective the training unit lasted three days. However, different session content, teaching methods and associated learning objectives required different levels of student participation and supervision, and different staff–student ratios.

Consequently, the optimum numbers of students differed for particular sessions and so a more complicated time-tabling arrangement resulted.

In fact *eight* days were required to delivery the *three*-day unit to a typical group of 80 student nurses. This arrangement is important since the time needed for delivery and utility, or ease with which this timing can be organised within a complicated timetable are important points when the overall evaluation of a unit or course is being made. Figure 3.4 illustrates how the three-day unit would be delivered to different numbers of students on different days over a typical two-week period.



There was an immediate intention on the part of this author to monitor and evaluate the course and disseminate any findings. For example, the first two deliveries of the unit had been evaluated in terms of

Kirkpatrick's (1976) immediate reaction and also in terms of attitude change pre- and immediately post-unit. This material had been presented at one nurse education conference, was accepted as a poster presentation at a major U.K. nurse research conference and was in the process of being written up for subsequent publication (Beech 1999). Indeed, it was also during this year that possibility of post-graduate study was investigated and later commenced [December 1999].

3.5.2 Goals And Objectives

Goals and learning objectives for the unit were agreed in consultation with the course leader, other course lecturers and following perusal of ENB guidelines and very limited published nurse education literature.

Goals:

1. To reduce the risks that student nurses and health professionals face by increasing their awareness of risk factors, improving their inter-personal and psycho-motor skills and, as a consequence,
2. to reduce the number and seriousness of violent incidents that they face.

Learning Objectives:

At the end of the Unit the student will be able to:-

- Briefly describe 3 psychological theories of aggression

- List the 4 components of Poyner and Warne's model of workplace violence (1986)
- Identify work-related factors associated with each of Poyner and Warne's components
- Use the model to better anticipate potentially dangerous situations and reduce associated risks
- Describe the role of the student nurse in any aggressive/ violent incident
- Demonstrate a non-provocative approach and stance and provide a rationale for actions
- Describe the main principles of breakaway techniques
- Describe the legal basis for using reasonable force against another person and the two decisions which have to be made [intent and potential]
- Express increased confidence in their own abilities to remain safe
- Express more reasonable attitudes and beliefs about:-
 - own safety,
 - relative risk and its assessment,
 - people with mental illness and learning disabilities,
 - the predictability of violence.

Subsequently, these learning objectives would present a number of learning outcomes, many of which would be evaluated within this research study [see Figure 3.5].

Figure 3.5: Unit learning outcomes in four learning domains	
Knowledge and Awareness	of psychological theories of Poyner and Warne's model of health service statistics of legal issues of principle of breakaway techniques of accepted risk factors of assessment, audit & prediction
Beliefs and Attitudes	about patients and staff rights to safety dangerousness and predictability mental illness & learning disability role of student nurse in management of violent incidents
Self Esteem and Confidence	to maintain own safety to be non-provocative to recognise early signs of aggression and risk factors in environment, task, self and others
Skills	-ability to take a non-provocative stance before another person -develop verbal and non-verbal inter-personal skills -practice and develop breakaway skills

3.5.3 Content

Reviewing and revising the content of training courses designed for permanent staff that the author was familiar with determined specific items of content. In addition valuable guidance about the particular needs of student nurses was obtained from ENB guidance and articles published in nurse education literature, particularly the work of Whitley *et al* (1996). As indicated in Chapter 1, the ENB guidance suggested

inclusion of the following topics:- recognition and prevention of violence; communication skills and assertiveness techniques; defusing and de-escalation techniques; support, de-briefing and post-incident management; ethical and legal aspects; self-awareness; understanding violence and aggression as a reaction to circumstances/conditions; promotion of a positive attitude towards individuals; and physical responses, such as breakaway and escape techniques.

Whitley *et al* (1996) suggest a range of student nurse-appropriate content, including a preventative philosophy, assessment of self, environment, and client, the role of the student nurse in an incident, and techniques for de-escalation. Consideration of this guidance, the numbers of student nurses to be included, the teaching resources available, other teaching commitments and the finite period of time that could be made available within a pre-existing curriculum resulted in the summarized list of content, presented in Figure 3.6.

3.5.4 Teaching Methods

Most material was delivered in a lecture discussion manner. The theoretical material was delivered by this author to larger groups of students [approximately 30-50 students] whilst demonstration and practice of de-escalation skills and self-protective breakaway skills was delivered in smaller groups [maximum 25 students] on Day 3 as in Figure 3.4.

Figure 3.6: Summarized unit content	
➤ Biological & Psychological theories of aggression/violence;	○ Genetic, hormonal or neurotransmitter imbalance, toxic substances, disease or structural abnormality, instinct, social learning/culture, aversive stimuli (frustration)
➤ Health service statistics;	○ Surveys by, for example, health services advisory committee
➤ Possible effects of aggression on the individual;	○ Short, medium and long-term effects Mezey and Shepherd (1994)
➤ Integrated model of workplace violence;	○ (Poyner & Warne 1986) assailant; victim; task; environment
➤ Self awareness of own response to aggression	○ Completion of self-assessment forms and discussion of results
➤ Risk factors, risk assessment, scanning	○ Examples of published risk assessments, risk management process, audit, personal surveillance strategy
➤ Risk reduction	○ Measures to reduce degree of personal risk
➤ Student role;	○ Guided by trained staff, in 'relative background'
➤ Non-provocative approach and stance;	○ Advice about what to do and not to do, including verbal and non-verbal skills (Turnbull <i>et al</i> 1990)
➤ Legal issues;	○ health and safety legislation, reasonable force, duty of care
➤ Breakaway skills;	○ Demonstration and practice of skills to breakaway from various wrist, arm and clothing grabs, hair-grabs from front and rear, strangles from front and rear
➤ Interpretation of videoed scenarios and problem solving;	○ Observation analysis and discussion of brief, commercially-produced, health care setting-based, fictitious, acted scenarios of actual or potentially violent incidents. Emphasis on identification of risk, explanatory model, key risk factors, good and bad practice, problem-solving, alternative approaches etc. Used to consolidate the unit.

The analysis and discussion of the videoed scenarios was also performed in smaller groups on Day 3 in order to consolidate the unit, demonstrate the applicability of principles to the health-care setting and allow greater student involvement and contribution. Every student was encouraged to contribute to the analysis of at least one videoed scenario.

Self-completion questionnaires were used to stimulate discussion and self-awareness of one's own tolerance levels, aggression potential and weakness (Breakwell 1997). Handouts included:- the Dangerousness Checklist (Breakwell 1997), effects of assault on the victim (Mezey and Shepherd (1994), health service statistics (Health Services Advisory Committee 1987), dos and don'ts (Turnbull *et al* 1990) and lawful excuse and reasonable force (Gostin 1986, Smith & Hogan 1988, Martin 1990, Dimond 1990, Lyon 1994, Paterson *et al* 1997).

Small group work activities were designed around the HSE Integrated Model of Workplace Violence (Poyner & Warne 1986). The breakaway training was delivered by two trainers, this author and one other accredited trainer [in total five different trainers were used at some time] in accordance with published RCN Institute standards for training (Royal College of Nursing 1997).

Beech (1999) presented an outline timetable for the unit and this is reproduced in Figure 3.7. It should be stated that, unsurprisingly, the

unit evolved slightly over time, in response to student feedback and reflection. Later versions of the unit merged some of the original day 2 material into day 1, creating a shorter day 2.

From Figure 3.7 it can be seen that the content identified there is very similar to that suggested or included by others. For example, Beale *et al* (1998), include prevention and calming measures, the multi factorial nature of aggression, self –awareness of own potential, logical progression from causes of aggression, risk assessment and reduction, interpersonal skills, training in non-aggressive breakaways, consideration of the aftermath and possible effects of aggression on the individual. The training also shares many similarities with a course described by Taylor (2000) for student nurses with five sessions [24 hours] which includes content on personal safety, theories of violence and aggression, legislation, violence at work and coping with violence, breakaway training and escape techniques.

Figure 3.7: Outline timetable for the unit
<p>Day 1 [Large group of students]</p> <p>Definitions, theories and models of aggression and violence</p> <p>Effects of aggression on the individual</p> <p>Incidence of aggression in health settings</p> <p>Self-awareness: tolerance, response to aggression and provocation</p>
<p>Day 2 [Large group of students]</p> <p>Violence equation</p> <p>Risk factors, aggressor, staff, environment, task</p> <p>Scanning and audit</p> <p>Prediction and management of risk, assessment of dangerousness</p> <p>Dos and don'ts of verbal and non-verbal interaction (theory)</p>
<p>Day 3 [smaller group less than 25 students]</p> <p>Dos and don'ts of verbal and non-verbal interaction (practice)</p> <p>Videoed scenarios of dangerous/difficult care situations</p> <p>Legal issues; use of reasonable force</p> <p>Breakaway skills training</p>

Finally, Figure 3.8 summarises chronologically the evolution of the training unit over time, early attempts at its evaluation and the logical development of this work in to a PhD study.

Figure 3.8: Chronology of main developments of the aggression management and prevention unit and its evaluation	
1996	New curriculum commences [training in prevention of aggression ad hoc]
April –October 1997	Author completes RCN Institute Course 'Trainer in prevention and management of actual and potential violence and aggression'
December 1997	Presentation of strategy for organised training for all student nurses
December 1997- March 1998	Consultation with NHS Trusts and gaining of approval
June 1998	First delivery of unit. 58 students in cohort. Evaluation via feedback sheet and attitudes questionnaire
July 1998	Evaluation of feedback
September 1998	Presentation of data at Nurse Education Tomorrow conference, Durham
October 1998	Second delivery of unit and collection of evaluation data
January 1999	Presentation of progress report to Strategic Course Management Meeting
April 1999	Poster presentation at RCN Research Society conference, Keele University
Summer 1999	Publication of article in Nurse Education Today (Beech 1999)
December 1999	Commencement of part-time PhD study at Institute of Work, Health and Organisations, Nottingham University
June 2000 - September 2001	Delivery of unit and collection of data for PhD [last three consecutive cohorts on curriculum]

CHAPTER 4 – RESEARCH METHODOLOGY

It was identified in Chapter 2 that evaluation research is a form of applied research with no exclusive methodology of its own. Rather, its distinguishing feature is the purpose for which data are collected (Clarke 1999).

This chapter presents a detailed description and justification of the research design used in the study being reported. It will describe the methodology utilised, the creation, testing and administration of data collection instruments, and indicate the methods used to identify the sample and obtain access. The chapter will also incorporate issues related to choice of research designs and data collection methods, including advantages and disadvantages of various designs and tools, possible sources of bias and best advice about methods to control or eliminate bias.

4.1 RESEARCH DESIGN

There is no such thing as the perfect piece of research. All research involves making decisions and compromises and good research design attempts to minimise the negative effects of such decisions. Hoel *et al* (2001) acknowledged this reality, in relation to research evaluating interventions to manage violence.

As identified in Chapter 2, the tradition of programme evaluation is not strong in this area. Reviewing intervention programmes for workplace stress and violence, Hoel *et al* (2001) found that most failed to withstand rigorous scientific review, based on chosen research design, for example, lack of a randomised control group.

Research design is the skill of turning research questions into do-able projects (Robson 1993), thus the principle becomes one of selecting appropriate research strategies, data collection methods and protocols to achieve answers to the questions asked. The research design selected for this study was a *quasi-experimental one, specifically an interrupted time series design*. The data collection method selected was the use of a *purposely-designed questionnaire* and the main researcher was the person who delivered the unit. Some of the issues associated with this research design and data collection method will now be highlighted. Like all methodological options the approach selected for this study has advantages and disadvantages but it provides an effective way of answering the Research Questions identified at the end of Chapter 1.

4.1.1 Quasi-Experimentation

Difficulties arise when the pure experimental model is transferred from the laboratory to the real world. Maybe, as in case, it is not considered possible to randomly allocate participants to different treatment or control group, or perhaps it would be unethical to do so [again as in this

case] (Stecher & Davis 1987, Fife-Schaw 2000a, Robson 2000).

Possibly adequate control over extraneous variables may be lacking, or mask the effects of treatment variables, or bias assessment procedures (Robson 1993).

The intervention or programme may not be delivered consistently over time due to ideological or personnel changes or data collection methods may not be adhered to consistently (Clarke 1999). Furthermore, failure to isolate interaction between subjects allocated to different groups may neutralise their assumed independence (Robson 2000). This is a considered a real possibility when, as in this case, studies take place in single, geographically and socially close settings (Robson 2000).

These difficulties are often denied or overlooked in any interpretation of results but a more honest approach pioneered by Campbell & Stanley (1963), Cook & Campbell (1979), would be to acknowledge them and, instead, devise 'quasi-experimentation' – a range of designs more easily realizable outside the laboratory, for example, possibly using a whole existing school class rather than randomly selected samples.

Essentially, a quasi-experiment involves using an experimental approach without random assignment to treatment and comparison groups (Campbell & Stanley 1963).

Validity is the main potential casualty resulting from lack of random allocation and control groups. Cook & Campbell (1979) added four suggestions to a highly influential, earlier list of eight threats to internal validity devised by Campbell & Stanley (1963). The possibility of each threat has to be assessed and managed in quasi-experimental research approaches. The list will not be considered in detail but note is made of the need for researcher vigilance in preventing possible bias in the interpretation of scores as a consequence of participant history, maturation, mortality [drop out], familiarity with test instruments or change in interpretation of test scores, regression to the mean, compensation effects in the name of fairness etc. Three strong quasi-experimental designs are referred to – the pre-test post-test non-equivalent group design, the [interrupted] time series design and the regression-discontinuity design (Robson 1993), although, once again, aspects of several can be combined (Fife-Schaw 2000a).

4.1.2 Time Series Designs

Robson (1993:105) describes this as the 'interrupted time series design' wherein "there is just one experimental group, and a series of observations or tests before and after an experimental treatment" and suggests that, with regard to the total number of observation made, there are advantages from including even one additional pre- and/or post-test [but preferably both] . The benefit means that the researcher starts to collect information about possible trends in the data, rather

than describing an isolated point before and after the intervention, and hence counter several threats to internal validity and reliability. Fife-Schaw (2000a) makes a similar point about increasing the number of observation points but cautions about the possible negative effects on participant fatigue, boredom and irritation.

As stated earlier, the chosen research design for the study being reported was an interrupted time series design. Fitz-Gibbon & Morris (1987) identify this design as appropriate for course evaluation. They suggest five stages to the conducting of this type of research (Fitz-Gibbon & Morris 1987), namely,

1. prepare/select an outcome measure which can be used repeatedly
2. decide on composition of experimental group; same people measured repeatedly or successive groups of different people [this study incorporates both]
3. collect at least three measurements made prior to programme made at regular intervals [this study managed two]
4. check programme implementation
5. collection of measurements continued at same regular intervals after programme conclusion.

Purported advantages of this approach include it being relatively less susceptible to history effects since organisational changes or historical events are unlikely to coincide with the treatment effect or have an

enduring effect over time (Fife-Schaw 2000a) and, in addition, small differences between pre and post treatment scores are unlikely to be maintained if the treatment is really ineffective. Also, any maturation effects are likely to be reflected in gradual score trends rather than sudden changes occurring at the time of the intervention (Fife-Schaw 2000a).

Potential problems are likely to be linked with testing effects associated to the growing familiarity with the test instrument/ procedure, and reduced test anxiety that accompanies repeated application (Fife-Schaw 2000a). If any experimental effect was small in magnitude it might be hidden by this tendency for people to slowly improve without any experimental intervention. Instrumentation effects are another possible source of bias or error, wherein possible changes in the manner of administration or interpretation might occur. In addition, given a prolonged timescale of participant involvement, there are potential problems with participant mortality, in that participants may indeed die but are more likely to drop out because of boredom or relocation. Any risk of a possibly biased sample emerging could be monitored by careful pre- and post-intervention mapping of the sample scores (Fife-Schaw 2000a).

Fife-Schaw (2000a: 78) is bullish in asserting that “quasi-experiments should not be seen, however, as always inferior to true experiments”. Indeed there are many occasions when the choice would be preferable

to a true experiment performed in a controlled laboratory setting, for example, as an aid to generalisability or to increase external validity.

4.2 DATA COLLECTION METHODS

4.2.1 Questionnaires: Uses and Design

In the study being reported the chosen data collection method was a questionnaire. Several authors advocate the use of questionnaires to gather data on personal attributes and characteristics, behaviour and events, beliefs and knowledge, and attitudes and opinions (Parahoo 1993, Clarke 1999, McColl *et al* 2001). Questionnaires are very widely used in small-scale evaluations, for example, in higher education student evaluation of courses, possibly because of, the apparently deceptive simplicity of creating questionnaires and the speed of completion and ease of incorporation into teaching programme timetables (Robson 2000).

Other advantages including making the least demand on the time of the participant, avoiding observer/interviewer bias effects, and allowing quick analysis if properly designed (Bee & Bee 1994). It is also versatile and comparatively low cost (Fife-Schaw 2000b). McColl *et al* (2001:24) suggest that, although self-completion questionnaires are most commonly delivered and returned through the mail, they can also be administered, as in this case, via supervised, [so-called 'captive audience'] self completion, wherein "respondents complete questionnaires in the presence of a researcher, who is available to

provide some assistance or explanation and who may also check questionnaires for completeness of response". They add that the technique is appropriate for collecting data from students in a class room (McColl *et al* 2001), a point also made by Parahoo (1993). At the same time it is acknowledged that collection of data from a 'captive audience' by a member of staff of that organisation raised several ethical questions related to power and coercion, which will be covered shortly.

Supervised completion also overcomes some of the identified administrative disadvantages of postal surveys, for example, low response rate leading to possibly biased responses (Bee & Bee 1994), not knowing who completed the returned questionnaire or whether they consulted with others or jointly completed it (Kumar 1999). Limitations associated with questionnaires would include the gap between the accuracy of peoples' reported intentions and their actions, memory inaccuracies in recalling past events or beliefs, and possible exaggeration to give socially acceptable answers or to protest against or offer support to a scheme (Parahoo 1993). Finally, self-completion and particularly postal questionnaires fail to provide additional information from interpretation of body language and do not allow the opportunity to probe for further information.

Many authors offer detailed advice and guidance on aspects of questionnaire design, for example, structure, layout, question type and

response format, wording, sequence, length (Robson 1993, Oppenheim 1992, McColl 1993, Bee & Bee 1994, Robson 2000, Fife-Schaw 2000b). McColl *et al* (2001) reviewed the evidence and made recommendations for best practice in relation to the use of questionnaires in surveys of health service staff and patients. The pertinent recommendations for practice they made include the following:

- Using open-ended questions sparingly
- Exercise caution in the use of negatively phrased attitudinal items
- Possibly place demographic items at the end of the questionnaire [limited evidence]
- Include the middle response category for attitude/opinion questions since it does not necessarily represent a neutral position
- Maintain same question ordering over time in longitudinal studies
- Avoid excessively long questionnaire especially if saliency is low
- Avoid crowding or reducing 'white space'
- Avoid splitting a question over two pages
- Use a font of at least 10 points, a distinct type face and avoid excessive use of italics or upper case characters
- Use a vertical response format for closed-ended questions and a horizontal response format for rating scales
- Place instructions directly at the point that they are needed
- Use coded [numbered and identifiable] questionnaires to facilitate follow-up and record linkage (McColl *et al* 2001)

4.2.2 Different Types of Question

The changes in different learning domains anticipated in the various Unit learning outcomes, research questions and research hypotheses necessitated the incorporation of different forms of question within the questionnaire. Changes in attitudes were monitored with Likert scales, while knowledge change was measured with open response questions, and attribution of blame with a VAS line. In addition, questions were constructed to allow some key demographic data to be recorded.

4.2.2.1 Likert Scales

Attitudes and opinions are “essentially evaluative, reflecting respondents’ value judgements about what is good or bad, effective or ineffective, desirable or undesirable” (McColl *et al* 2001:2). Fife-Schaw (2000b) suggests there is little consensus about how to measure attitudes despite them being of enormous interest. It is preferable to assess attitudes by means of a series of related questions [a summated scale] rather than by means of a single question (Robson 1993), each answer acting as a ‘marker’ and, when totalled, giving some indication of aspects of the attitude.

In many contexts the most popular and easily constructed type of measurement scale (Oppenheim 1992) is a summated rating scale or Likert scale [although this label is only technically accurate if the scale generates normally distributed data (Fife-Schaw (2000b))] which is used to assess the level of agreement or disagreement with certain

statements and commonly used to measure attitudes. Several authors offer procedural details for the creation of a Likert scale which involves the stages of question generation, response categorization and measuring discriminative power (Robson 1993, Kumar 1999). Typically the scale will have 5 or 7 response points, an odd number being used in order to allow a central neutral response, although it is possible for this neutral response to be over-used and lead to more questionable results (Oppenheim 1992).

Purported advantages include allowing degrees of agreement with a statement, rather than absolute 'yes' or 'no' responses to a closed question, and thus increasing the reliability of the scale (Howe 1995). Furthermore, many researchers view Likert scales as interval level measures, with equal intervals between points on the scale and thus suitable for more powerful parametric statistical analysis. Scoring is organised so that higher scores are obtained for agreement with positive statements and disagreement with negative statements. However, the interpretation of total scores on an attitude scale and comparison of the attitudes of different respondents is made more complicated since the same score could have been obtained by very different responses to individual statements.

4.2.2.2 Visual Analogue Scales

Several authors suggest that Visual Analogue Scales (VAS) have become increasingly popular for various assessments in psychology

and medicine since early in the twentieth century (Lingjaerde & Foreland 1998). Polit & Hungler (1995) describe VASs as a means of measuring subjective experiences, such as pain, fatigue, and breathing, while Parahoo (1997) identifies the VAS as an attitude-measuring scale.

Cline *et al* (1992:378) provide a clear, typical description of the VAS, as follows,

“Operationally, the VAS is a vertical or horizontal line, 100mm in length, anchored by terms that represent the extremes of the subjective phenomenon the researcher wishes to measure. Subjects are asked to indicate the intensity of a sensation by placing a line across the VAS at a point that represents the intensity of the sensation at that moment in time. Responses are scored by measuring the distance from the lowest anchor point to the subject’s mark across the line”.

Advantages of the VAS include it being quick and easy to administer (Waltz *et al* 1991), capable of reliable, and relatively sensitive measurement (Cella & Perry 1986), using uncomplicated language which makes the scale simple for the subject to understand (Cline *et al* 1992), and requiring a minimum of motivation (Lingjaerde & Foreland 1998). It can also chart changes over time in the feelings or attitudes being measured.

Some authors suggest the scores obtained are at least of an 'interval level' type (Brantley & Bruce 1986, Cline *et al* 1992), while others (Briggs & Closs 1999) suggest they appear to have the qualities of ratio data, and may be treated as such using parametric statistics, providing that the data are normally distributed. In the area of pain and pain relief measurement, investigators judge VAS to be "a straight-forward, reliable, reproducible, valid and sensitive tool (Huang *et al* 1996).

Possible disadvantages are that people with visual impairment or psychomotor disability may find it difficult to mark the line, and the concept of a line measuring proportions of feelings may be difficult for some respondents to comprehend (Parahoo 1997). Furthermore, while quick to administer, the scale does not produce an instantaneous result and still has to be measured in some way which can be tedious and take rather longer to calculate the value represented by the mark and enter in to a database (Cline *et al* 1992, Choiniere & Amsel 1996). Investigators have used micrometers, clear rulers or transparent template overlays (Huang *et al* 1996).

4.2.2.3 Open Response

McColl (1993) suggests that open-ended questions are particularly useful in pilot studies when the range of a particular issue is being explored, allow respondents scope to highlight particular issues and avoid the possible bias associated with offering a narrow range of forced choice responses.

However, data from these questions can be difficult to interpret, classify and analyse (Oppenheim 1992, Clarke 1999, Fife-Schaw 2000b), the format places an increased burden on the respondent (McColl 1993), and also relies on the participant's ability to articulate their ideas and express them clearly (Parahoo 1993). The extra time and effort required for deliberation and written response may deter the participant from engaging (Clarke 1999).

With regard to subsequent interpretation of responses, Robson (2000) suggests a form of content analysis wherein repeated re-reading of answers will allow the researcher to produce a reduced set of coded categories which cover most responses. In an earlier text Robson (1993) provides more detailed advice about this procedure and advocates taking "a substantial, representative sample of (say fifty cases) selected from the total set of responses" [and not just early responses], copying responses to each question to a large sheet of paper and then attempt to "develop a smallish set of categories (say eight or ten) into which these responses can be sorted" (Robson 1993:253).

The number of categories in part depends on the overall number of cases and required depth of statistical analysis, and, all together, the process of designing coding frames and the actual coding operation are extremely time consuming (Oppenheim 1992).

4.2.2.4 Demographic Data

Fife-Schaw (2000b) highlights that although demographical data is readily available to the respondent themselves many are resistant to revealing this in answers. It is common and apparently easy [but deceptively difficult] to ask respondents for background information on age, biological sex or gender, ethnicity and nationality, social class or socio-economic status and income.

Thought should be given as to the degree of accuracy required in replying to the age prompt since some respondents may not wish to declare their exact age. Similarly, some confusion surrounds the terms sex/gender, ethnicity/nationality and social class and clear prompts are needed here in order to obtain consistent, accurate answers.

Fife-Schaw (2000b) advocates an opt out tick box for some of these categories to allow absence through neglect or accidental omission to be distinguished from deliberate omission and possibly avoid the respondent feeling that they won't respond to anything. Opinions vary about the appropriate position i.e. beginning or end of the questionnaire (McColl *et al* 2001) but several authors (Oppenheim 1992, McColl 1993, Fife-Schaw 2000b) suggest requesting this data towards the end, especially if it might be judged sensitive, when the respondent has settled and become accustomed to answering.

4.3 OTHER ISSUES - BIAS, ERRORS.

In addition to the challenges to internal validity briefly mentioned earlier there are several other issues pertinent to the design and data collection methods chosen for the study. These include possible coercion problems associated with a staff member obtaining data from a captive audience, measuring estimations of differences in self-efficacy over time – the so-called response-shift bias - and the advantages and disadvantages associated with 'insiders' being involved in evaluating their own training courses.

Other issues to do with reliability and validity associated with statistical analysis will be considered in the next chapter when results are presented.

4.3.1 Ethics and Possible Coercion Problems

In many disciplines students have traditionally been used as research subjects and expected to participate in research conducted by their peers or by their lecturers/tutors. In this study, the nature of the lecturer-student relationship and the 'captive audience' circumstances [Section 4.2.1] introduced several concerns into the research process. Under these circumstances there is potential for a number of ethical problems to emerge, including abuse of power, coercion, lack of confidentiality and absence of meaningful informed consent, any one of which could harm the student (Clark & McCann 2005).

Competence to make judgements is desired in respondents but this may be compromised when the research topic is course-related, or when it is performed by a course lecturer (Beauchamp & Childress 1994). Clark & McCann (2005) suggest that consideration must be given to mechanisms for developing informed and voluntary consent, maintaining identity and data anonymously and confidentially, and ensuring fairness, that is to say, including students because the issues relates to them, not because they are convenient 'subjects', and maintaining equitable involvement of different groups within an organisation. Risk of harm has to be minimised

These issues were covered to the degree required at the time by I-WHO and the host institution [See Section 4.6 for further details]. For example, in this case students were informed verbally and using OHP, about the nature of the study and the student commitment in a presentation by the researcher/lecturer that intentionally distanced him from the research. They were reassured that no issues would follow from non-participation of any or all questionnaires. No separate consent form requiring a signature was used and this was considered normal at the time wherein consent was implied by completion of the questionnaire (Clark & McCann 2005).

All forms - completed and non-completed – were returned to the front of the room at the end of the session to protect identity and the fact that non-completed forms were found among completed ones indicated that

some students felt safe to not complete questionnaires -see response rate data in Chapter 5. The use of ID codes on questionnaires and codes stored separately further maintained the anonymity between questionnaire administration points.

Fairness is about ensuring relevance and preventing over-exploitation of willing groups. The host institution Research Access Committee ensured that student groups were not 'over-researched', while the high response rate may be indication of the perceived relevance of the research to the students. Finally, any potential harm associated with thinking about violence issues [extra to that demanded by attending the training Unit] was minimised by reinforcing the supportive mechanisms available to students, namely, personal tutor, Unit lecturer, module and year coordinators and University counselling service.

4.3.2 Response-Shift Bias

Arvey & Cole (1991) discuss the difficulties associated with using self-report methods to measure subjective outcomes, such as, for example, changes in estimation of self-efficacy in managing potentially violent situations. In discussing three different kinds of change they emphasise the so-called 'response-shift bias' wherein an intervention causes a real change in outcome measure but this change is confounded by an internal recalibration of the scale(s) used (Arvey & Cole 1991).

For example if a respondent classified himself as 'average' on a construct before a course and then received input that made them much better but also made them aware of how much there was to know then they may still classify themselves as 'average' at post-test, even though they have dramatically improved from the pre-test. Such a result would wrongly present the training as being ineffective. A possible solution to this is to ask respondents at post-test to rate the variable in a retrospective pre-test 'as it was before the intervention' and as it is after the intervention. Each respondent's scores are then compared individually one at a time in a very long-winded manner.

4.3.3 Internal or External Evaluators?

Under normal circumstances this choice is considered a most important one since any suspected allegiance or suggestion of bias can result in the rejection or invalidation of the results and recommendations (Reid & Barrington 1997). Clarke (1999) discusses the advantages and disadvantages of using internal or external evaluators and the arguments are summarized in Box 4.1.

Robson (2000) identifies the potential problems of using internal or external evaluators, including compromising the objectivity of an evaluation or producing positive or negative reactivity within a team to the evaluation process. Clarke (1999) argues for the combination of the two roles in a single evaluation. In the case of the study being reported, aspects of both internal and external evaluation were incorporated with

an internal evaluator performing the study under guidance of experienced external evaluators [supervisor and team].

Box 4.1: The advantages and disadvantages of using internal or external evaluators [from Clarke (1999)]

Advantages

Internal Evaluators will be:

- familiar with history, background, politics, issues and culture of the organisation
- likely to be more committed to implementing evaluation recommendations, being responsible for producing them
- likely to focus on the central concerns as perceived by management

External evaluators have:

- an independent stance and offer a fresh perspective
- an objective, critical approach
- an overview of numerous organisations to serve as comparisons
- a knowledge and experience of a wide range of evaluation techniques
- a resilience to intimidation by management

Disadvantages

Internal evaluators may:

- have a vested interest in a particular outcome
- often be over-influenced by the history and knowledge of organisational issues
- sometimes be over-influenced by the known views of management
- be unlikely to have had experience of a broad range of evaluation techniques
- be less committed to the need for evaluation
- be inclined to favour programmes developed within their own unit or section
- find it difficult to encourage stakeholders in their organisation to actively participate in the evaluation process

External evaluators may be:

- ignorant of internal matters so that judgements may not reflect the complex reality of the situation
- unaware as to who are the key players in a particular setting and thus more easily misled by interested parties
- more interested in a report than its implementation
- influenced by the need to secure future contracts
- insensitive to organisational norms and internal relationships
- primarily responsible to an external organisation.

This arrangement was obviously essential, since, if the aggression management unit was subjected to only external evaluation then this would have precluded the involvement of this author, and prevented the existence of the PhD training [and this Thesis!]. The only alternative decision would have been about whether to evaluate training at a different educational institution that this author had no personal involvement in. The rationale for the choice of course and sample will be given shortly.

4.4 DESIGN OF STUDY

4.4.1 Choice of Methodology and Data Collection Method

Having reviewed a number of the issues associated with research design along with purported advantages and disadvantages of various data collection methods it is now appropriate to provide a detailed account of the design and methods used along with a rationale for their selection. As previously specified the research study design selected was an interrupted time series design and the data collection method was a purposely-designed questionnaire administered by the main researcher, who was also the person delivering the unit. Like all possible methodological options the approach selected for this study has advantages and disadvantages but it provides an effective way of answering the Research Questions identified at the end of Chapter 1.

This is for a number of reasons. Firstly, the 'juggernaut nature' of a real world pre-registration nurse education programme meant that random allocation was not possible. As explained in Chapter 3, the Unit under study was a pre-existing one slotted into an existing training programme. The combination of the demands of the curriculum and lecturers' competing commitments meant that there was very little administrative opportunity to withhold or delay the training for part of a larger group.

More fundamentally, it could/ would be considered unethical, even negligent, to create control groups by delaying or withholding training from some students, having previously acknowledged that they are on a high risk staff group [See details in Chapter 1]. In any event, an attempt at randomisation or creating control groups would have almost certainly proved futile. Students develop great camaraderie during their training and socialise and talk either in college or on clinical placements or 'off-duty'. Therefore, it is likely that undeterminable 'contamination' of control groups would have occurred with students sharing classroom experiences and clinical placement details, and briefing each other about what happened in different sessions, units / modules and revealing what their friends can expect next.

By applying a rigorous, systematic quasi-experiential research design to an existing course and studying it under normal delivery conditions the study increases the external validity of the findings and any possible

generalisability. Furthermore, the longitudinal nature of the design and the staggered delivery to a number of consecutive cohorts offers the opportunity to detect information or evidence of possible confounding organisational developments that may be occurring [changes in clinical placement admission policy, alteration in number of violent incidents, introduction of separate educational input in other modules or in clinical practice etc.].

The maximum possible number of cohorts was used in the study. This number was effectively determined by the availability of the approved questionnaire, gaining organisational permission to access the groups, and the periodic 'real world' curriculum developments that meant a new curriculum model was introduced in January 2001.

A purpose-designed, self-completion questionnaire was used in preference to direct observation or interviews because of the large student numbers, time restrictions, and subjective nature of some dependent variables. The relative strengths and weaknesses of this approach have already been discussed. The questionnaire was created by pooling the expertise of a number of researchers and tested for validity prior to widespread use.

The researcher was also the unit leader and main lecturer. It could be argued that this arrangement maximised the advantages identified in the section above on the advantages of internal or external evaluators

[see Box 4.1]. Local knowledge was invaluable in terms of interpreting the learning outcomes, maintaining the stability of the unit content and its delivery during the period of evaluation, and creating opportunities for administration of the questionnaire.

Avoidance of any possible allegiance to the teacher/evaluator was attempted. Distance in the eyes of participants was established by presenting the study as an evaluation by another completely separate organisation, emphasised by the questionnaire sporting an independent letterhead and logo. External evaluation expertise from the supervisor was used in guiding the research design, collation of data, and analysis and interpretation of results.

4.4.2 Overall Questionnaire Design Priorities

Chapter 1 concluded with a statement of the Research Questions and Research Hypotheses. The intention was to evaluate the effects of the Unit on the attainment of a number of previously determined and explicitly specified learning outcomes. These learning outcomes related to domains commonly identified in training evaluations (Kirkpatrick 1976, Kraiger *et al* 1993) as identified in Chapter 2, and were similar to those targeted by a limited number of other violence management trainers/ researchers, as identified in Chapter 1.

Hence, there existed a potential integrity or unity between stated learning outcomes, training evaluation model domains/levels, previous

violence management training evaluations, and the research questions identified in this study that needed to be reflected in the sections of the questionnaire. This is in line with the 'outcome-oriented' models of Evaluation Research described in Chapter 2. Box 4.2 provides an outline, and later in the Chapter, a detailed illustration of this arrangement will be provided [Box 4.5].

It will be recalled that the learning outcomes of the Unit related to increases in confidence, self-assessed competence, and knowledge. They also referred to improved attitudes/beliefs about the causation, prediction and prevention of violence, and change to simplistic, stereotypical beliefs about the perpetrators having mental health problems or learning disabilities etc.

It is desirable that all of these areas be represented in the questionnaire. However, this is not say that all questions have to be newly created 'from scratch'. Streiner (1993:140) advocates a "process of gathering potentially useful items from various sources and then winnowing out those which do not meet certain criteria". Similarly, McColl *et al* (2001) caution against unnecessarily 'reinventing the wheel'. Instead, they emphasise the advantages of "drawing on the expertise and experience of others" (McColl *et al* 2001:208) and the time saved in generating and testing the properties of new questions with the usual provisos about applicability to different populations.

Box 4.2: Outline of interrelationship between learning domains, unit learning outcomes and questionnaire sections.		
Learning domains (Kraiger <i>et al</i> 1993)	Unit learning outcomes	Sections of questionnaire
Cognitive outcomes: <ul style="list-style-type: none"> ○ Verbal knowledge ○ Knowledge organisation ○ Cognitive strategies 	Theories and multi-factorial models [including attribution of blame] Risk factors NHS Incidence statistics Role of student nurse Prediction and prevention	Scenarios- open response VAS scales and Likert statements Scenarios-open response Likert statements Likert statements Scenarios/-open response and Likert statements
Affective outcomes: <ul style="list-style-type: none"> ○ Attitudinal ○ Motivational ○ Disposition ○ Self-efficacy ○ Goal setting 	Attitudes Confidence in own ability Staff and patient rights Self respect Motivation to change	Likert statements Likert statements Likert statements Likert statements Likert statements
Skill-based outcomes: <ul style="list-style-type: none"> ○ Compilation ○ Proceduralisation ○ Composition ○ Automaticity 	Therapeutic approach Therapeutic stance Interpersonal skills Verbal and non-verbal de-escalation Breakaways	Likert statements Likert statements Likert statements Likert statements Likert statements

4.4.3 Process of Creating the Questionnaire

The process used to generate the questionnaire followed the earlier suggestions. The author firstly located published examples of questions related to the learning outcomes. These questions were then reviewed by the author, the study supervisor and another experienced violence researcher to determine their suitability for inclusion, with or without modification. Subsequently, additional questions were generated to

permit the evaluation of any learning outcomes inadequately covered by questions obtained from previously published sources.

In addition some relevant demographic details about respondents were sought and a separate section was designed to collect this material.

The questions were then assembled into a number of “modules or sequences” (Oppenheim 1992:109) and scrutinized for utility, content and face validity by peer review.

Several colleagues and experienced researchers offered comments on the original form and, subsequently, a number of amendments were made. Specific amendments included, for example,:

- The Institute of Work, Health and Organisations Letter head and logo were placed prominently at the top of page 1 in order to distance this author from the questionnaire in the eyes of respondents and so remove a potential source of indirect ‘interviewer bias’.
- The typeface was changed to Ariel 12 point to increase clarity and legibility.
- The position of the VAS lines were moved slightly in relation to the scenarios statements to clarify the relationship between them [i.e. which line was associated with which scenario] and the symbols use to separate Scenario 1- Question 2 from Scenario 2 - Question 3, originally a wavy line, was changed to prevent any

possible confusion between the separator and the VAS line above it.

- Minor changes were made to the wording of question 2 to further eliminate 'jargon'.
- The structure and form of wording of the instructions for Scenario 1- Questions 1 and 2, and Scenario 2 – Question 3 and 4 was made identical.
- Where possible, the number of lines allocated to each Likert-type statement on page 3 was made the same [with the exception of statement 12] so that the shorter statements each had the same visual impact as the longer ones. Statement 12 was much longer than the others and occupied another line but was left as such, since it was judged that providing each statement with this amount of space would have taken this section on to two pages, lengthening and unbalancing the entire questionnaire.
- Some additions and modifications were made to the original list of acceptable response categories for Scenario 1 - Question 1 and Scenario 2 - Question 3.

4.5 QUESTIONNAIRE – DETAILS OF TYPES OF QUESTIONS CREATED

4.5.1 Demographic Details

As previously described, demographic details are normally sought via closed, fixed-category questions located either at the beginning or near

the end of the questionnaire. On this occasion, mostly for aesthetic reasons associated with wanting to minimise overall length, and maximise an open, 'white space' appearance (McColl *et al* 2001), these questions were located at the start of the questionnaire.

The variables considered relevant in this case consisted of sex, age, destined nursing speciality, previous management of aggression/ violence training, and type and experience of/involvement in violent incidents during clinical placements. Age, sex and destined speciality were included to enable a subsequent comparative exploration of, for example, possible differences in attitudes regarding vulnerability, confidence in managing aggression, or identification with the age or gender of the characters in the scenarios.

The author was aware that a proportion of students had previously worked in unqualified capacities in statutory, private and voluntary sector health care, and social care as nursing assistants, health care support workers or care workers and a good proportion of these would have received some theoretical or practical training in the prevention or management of aggression. This training would have ranged from brief, theoretical instruction as part of an induction programme to attendance on 5-10 day theory and practice restraint courses. Interest was expressed as to whether there would be any discernable difference in their responses prior to the start of the unit and, if so, whether the difference would remain following completion of the course. This

interest is paralleled with the recent work of Ilkiw-Lavalle *et al* (2002) who found that staff without prior training had the greatest improvement in knowledge scores post-test, especially regarding prediction, management and legal aspects of aggression.

4.5.2 Knowledge

Options in question design for assessing knowledge included multiple-choice questions or short answer, open response designs in response to a stem or prompt. Whittington (1997) effectively utilised two short 'profiles' or scenarios with associated multiple-choice questions as a method of assessing continuing education learning following study of a published journal article. Other researchers report the creation of specific tests or examinations [See Chapter 1] but rarely include details.

On this occasion the primary knowledge concern was in the student nurse's ability to quickly recognise at any early stage risk factors associated with the causation of violence. In the taught sessions, as described in Chapter 3, these factors were educed from the students recent experience supplemented by lecturer input and then located within different categories of a widely advocated total organisational response model (Poyner & Warne 1986).

It was decided to create two knowledge test questions that presented un-dramatic clinical situations as brief scenarios. Maximum student applicability was attempted by avoiding extremes of violence and,

instead, describing everyday, moderate levels of violence or challenging behaviour. The student nurse was required to identify potential risk factors in the scenarios and document their answers in a blank open response answer box. This format was selected, rather than a 'select from options' format, as might accompany a series of multiple-choice answers, since it required the student to answer without any visual prompt, and so, perhaps, demanded a higher level of recall or problem-solving behaviour from the respondent.

The two scenarios were designed to incorporate a number of identified risk factors that were developed in the teaching sessions. These risk factors were first listed and then questions designed, either completely or by modification of suitable previously published questions (Whittington 1997) that included a good number of these risk factors. It was intended that the risk factors presented in the scenarios should represent the different categories of the total organisational response model, namely, assailant characteristics, staff characteristics, environmental characteristics and task/interaction characteristics.

It should be noted that, subsequent to administering the questionnaire but prior to statistically analysing responses, a large sample [N=250 approximately] of completed questionnaires were reviewed. This activity resulted in additional approved responses in some of these categories. The process also highlighted the range of ways in which students responded to the scenarios and allowed the inclusion and exclusion

criteria to be clarified explicitly, and so be consistently interpreted in future analysis.

Scenario 1 related to a young man and contained the following text substantially developed from a published question by Whittington (1997):-

It is late on Friday evening and the A&E Department is already very busy. A 30-year-old man arrives in an intoxicated and dishevelled state. He has a number of cuts to his face and hands and quickly becomes agitated. He shouts obscenities and mentions something about the Patient's Charter. On seeing the man's behaviour, a male nurse approaches and attempts to put his hand on the man's arm in an effort to placate him. The young man pushes him away and becomes even more vociferous.

Scenario 2 was newly developed to complement the risk factors in Scenario 1. It related to an elderly woman and contained the following text:-

Mrs Smith is a 75-years-old woman with a long history of psychotic episodes who lives in a residential home. She now suffers badly with arthritis in her knees and hips and needs staff assistance when using the toilet. On this occasion, having just used the bathroom, she rings the bell to summon help. Most staff are in a hand-over

meeting between shifts and no-one answers her call. After waiting a few minutes she starts shouting for assistance and banging on the cubicle door. A female member of staff, about to finish her shift, arrives to assist the lady back into her chair as quickly as possible. She pulls the old lady to her feet rather quickly, saying "There is no need to shout – I have others to attend to as well you know!". At this point the old lady swears at the member of staff and lashes out at her.

It can be seen that, between the two scenarios, different aspects of patient and staff gender, patient age, care setting, types of aggression/violence, patient symptomatology, insight, and 'blame' are represented. Box 4.3 displays the final list [after analysing 250 completed examples] of acceptable risk factors for Scenario 1, identified and delineated under different categories of the total organisational response model, and the rules about inclusion/exclusion. The list also shows a range of equivalent words that would only be counted as one risk factor even if several similar words were included in one student's responses. Box 4.4 contains the same material for Scenario 2. The written instructions given to students before reading the scenario and before responding to it were in the format of 'Read the following brief scenario and then answer the questions related to it' and 'List below all the factors that you believe contributed to [the young man's aggressive behaviour], or [the old lady lashing out].

A small number of additional aspects of knowledge, for example, of NHS violent incident figures and trends, of the legal use of force, and the role of the student nurse in the management of an incident, were included in some of the Likert questions. These statements will be considered in detail shortly, in a later section.

Intentionally Blank

**Box 4.3: Acceptable risk factors in different categories - scenario 1
[young man A&E] [from 250 questionnaires approx]**

ASSAILANT	<p>Agitation Alcohol - drugs – intoxication Angry /aroused /assault cycle Anxiety –stress - distress Argumentativeness- shouting obscenities Attention-seeking Confusion /Mental state Dishevelled state Embarrassment Fear Feels being controlled/ placated/ patronised Feels threatened/ intimidated /confronted/provoked/ misinterpreted or perceived attack Frustration Having to wait – impatience Knowing his rights – Patient’s Charter – Expectation Late - man tired Pain of injuries – cuts Previous incident/ fight – how sustained injuries Shock Young male</p>
STAFF	<p>Attitude Gender – Sex - Male Nurse Nurse’s stance, Closeness, Invade Space, Proximity Poor observation skills Staff Rushed Tired because late at night Wearing Uniform/ representing authority</p>
ENVIRONMENT	<p>Audience /crowd/ overcrowded Busy A&E Day of week Time of day/night Heat – hot environment Intimidating unfamiliar environment Noisy No security staff</p>
TASK/ INTERACTION	<p>Approach/ intervention/actions Permission [approach/touch] not sought Poor communication Touching arm –hand on arm/restraint/control Touching injuries</p>

Consistency Rules

Approach : only score if isolated statement “being approached” or if states something wrong with the approach. Don’t score if use present participle as part of describing something else “approaching”, “the nurse approaching.....”

Male nurse: only score if explicitly specify that maleness or gender or sex is the issue. Don’t score for mentioning ‘the male nurse.....’

Box 4.4: Acceptable risk factors in different categories - scenario 2
[elderly woman] [from 250 questionnaires approximately]

ASSAILANT	<p>Age -old person Agitation [worked up by shouting /banging] Anger Anxiety –stress- distress Confusion – panic- shock/ incomprehension/ disorientation Dependency / not independent / relying on others Embarrassment/ indignity Fear Frustration Having to wait – impatience- timescale Ignored –no-one answers calls Left alone -neglected - forgotten Low self esteem – low self worth Medication – side effects Nuisance/ feels blamed- sense of injustice Pain/ discomfort of waiting associated with arthritis Psychosis [history of] Response to rebuke – telling off</p>
STAFF	<p>Attitude/ manner / of staff - unsympathetic, Insensitive/ disrespectful Approach, behaviour Gender Impersonal – treat like a number - Patronising - rude - telling off / abrupt Poor communication skills - re delay – no apology Rushed nurse – impatient - wants to go off duty Tired [end of shift] Tone of voice</p>
ENVIRONMENT	<p>Absence of staff – in hand-over - Busy End of shift –confusion who's doing what Isolated toilet area Management of staff hand-over</p>
TASK/ INTERACTION	<p>Assistance [needed / asked for] Intimate task Lack of dignity/ humiliation [state of undress] Perceived attack Rough handling – “pulled up” Task rushed –speed of action</p>

Consistency Rules :
 Count any mention of arthritis or pain associated with arthritis as one factor.
 Count any mention of arthritis or need for assistance because of arthritis as one factor.

4.5.3 Attribution of Blame

Whilst not receiving a specific emphasis in the unit or its learning outcomes the attribution of blame for the causation of violent incidents was of indirect interest. The attribution of blame is associated with the application of multi-factorial total organisational response models. In simplistic models the cause of aggression and violence is always the assailant, that is to say, the patient, relative, visitor etc, and never the staff. In more elaborate models then it is conceded that a number of factors are implicated, included characteristics and features of staff. Hence it was hypothesised that, after presenting a multi-factorial model and encouraging the students to analyse work situations within it, they would be more likely to identify multiple causes of incidents and so reduce the proportion of blame attributed solely to the patient.

Change in the attribution of blame was the focus of Research Hypothesis No 4 listed at the end of Chapter 1. In the questionnaire it was measured using a 10 cm. VAS line. One question relating to each of the two scenarios and requested the student to indicate on the VAS line the proportion of blame for causing the incident that they attributed to the patient. Specifically, they were instructed to 'estimate by making a mark at a point on the line below, to what extent they viewed the [young man] or [old lady] to be to blame for the incident' on a range from 'not at all to blame' to 'totally to blame'.

4.5.4 Attitudes, Competence and Confidence

After reviewing published examples of questions designed to gauge changes in attitudes, competence, confidence etc. it was decided, once again, to utilise appropriate pre-existing questions, following the advice of Streiner (1993) and McColl *et al* (2001). Additional, questions were generated by the researcher, study supervisor and another experienced violence researcher, to meet learning outcomes inadequately covered by previously published ones, as identified in the 'Process' section above [Section 4.4.3].

A Likert-type question design was selected because it is the most straight-forward to create and to answer (Oppenheim 1992) and because of its relative prevalence in the violence training literature. The published work of American trainers/researchers Poster and Ryan (Ryan & Poster 1993, Poster & Ryan 1994, Poster 1996) and British trainer/researcher Collins (1994) furnished twelve statements for the questionnaire.

Poster and Ryan have used several variations of an 18 item [or 31 item] 'Attitudes toward Patient Physical Assault Questionnaire'. The statements within the questionnaires covered "three components of beliefs and concerns about assault: safety concerns, staff performance and legal issues" (Poster 1996:366). Responses were on a five-point scale ranging from strongly agree to strongly disagree with a neutral

position. The study under report adopted the same format for the questionnaire used.

On this occasion, three questions were incorporated unaltered [forming q3, q6 and q24]. A further four questions were included after minor modification to improve clarity or simplicity [forming q5, q7, q8, and q9]. Collins (1994) incorporated 12 attitude statements in his 'Attitudes toward Aggressive Behaviour Questionnaire', including three from the questionnaires compiled by Poster and Ryan. In this case, in compiling the study questionnaire, two questions were used unaltered from Collins (1994) [forming q13 and q18], and the thrust of two questions was taken but altered in the interests of clarity and to remove extremism [forming q1 and q10]. In addition, a question that measured the importance of training being provided was present in the work of both Collins (1994) and Poster and Ryan (1996). This statement was included with a minor modification for application to student nurses in training [forming q12].

A further 12 questions were created to cover other learning outcomes, resulting in a section containing 24 statements. Additional questions covered the areas of stereotypical attitudes about mental illness [q 2], about the predictability and preventability of aggression [q4, q11 and q20], about being non-provocative [q14 and q21], about the student nurse role [q15], about safety and legally protecting oneself [q16, q17, q22 and q23], and about making earlier, proactive lower-level

interventions in preference to later reactive, higher-level ones [q19]. The complete set of 24 questions can be seen in a copy of the final Questionnaire (See Appendix Two). Box 4.5 shows the more elaborate relationship between training evaluation model domains, learning outcomes, research hypotheses and the questions on the questionnaire.

Box 4.5: Relationship between training evaluation model domains, learning outcomes, research hypotheses and the questions on questionnaire			
Type of information	Unit learning outcomes	Hypotheses H	Sections of questionnaire
Demographics		H1	Part 1 'experience of violence'
Cognitive learning domain (Kraiger <i>et al</i> 1993) Outcomes: <ul style="list-style-type: none"> o Verbal knowledge o Knowledge organisation o Cognitive strategies 	-Theories and multi-factorial models -Risk factors	H2	Scenario 1– Q1 & Scenario 2-Q3
	[including attribution of blame]	H4	Scenario 1- Q2 & Scenario 2-Q4
	Assessment, audit, prediction	H2	Likert q22
	-Legal issues -Incidence statistics -Role of student nurse	H2	Likert q15
Affective learning domains (Kraiger <i>et al</i> 1993) Outcomes: <ul style="list-style-type: none"> o Attitudinal o Motivational o Disposition o Self-efficacy o Goal setting 	Attitudes: -Stereotypes of mental illness & learning disability	H5	Likert q 2 Likert q 7 Likert q8 Likert q 9 Likert q 10
	-Prediction and prevention	H5	Likert q1 Likert q 4 Likert q 6 Likert q 11 Likert q 19
	-Confidence in own ability to remain safe	H3	Likert q 17 Likert q 18 Likert q 20
	role of student nurse	H6	Likert q 23

	-Staff and patient rights -Self respect -Motivation to change	H5	Likert q 3 Likert q 5 Likert q 12 Likert q 24
Skill-based learning domains (Kraiger <i>et al</i> 1993) Outcomes: <ul style="list-style-type: none"> o Compilation o Proceduralisation o Composition o Automaticity 	Self assessed competence in -Therapeutic approach -Therapeutic stance -Interpersonal skills -Verbal and non-verbal de-escalation -Breakaways	H7	Likert q 13 Likert q 14 Likert q 16 Likert q 21

4.6 RESEARCH PROCESS

The process of performing the research proceeded through a number of clearly identified stages, for example,

- o Identify the nature of the research, devise hypotheses / research questions/ research methodology
- o Creation of data collection instrument (questionnaire)
- o Determination of sample and pattern of questionnaire administration
- o Obtain access approval
- o Identify opportunities to access student nurses
- o Devise 'Identification Number' coding system
- o Gain access to student nurse groups, present details of study and encourage participation

The first two stages have already been described at length and the remainder will now be covered.

4.6.1 Sample

A number of factors determined the choice of sample. A predominant consideration was to maximise the power of the study by including a large sample of student nurses. As previously discussed, it was not administratively possible to stagger the delivery of training within a cohort, or ethically acceptable to withhold training from some control group of students. The final selection of student cohorts for inclusion was determined by reviewing the projected Department of Nursing and Midwifery pre-registration business plan. The plan revealed that, after the questionnaire was created, three student nurse cohorts on the current curriculum remained available for inclusion in the study.

Furthermore, the plan showed that a new curriculum was to be initiated in January 2001. This development would result in a very different organisation of students' subjective learning experience. There was to be a different sequence of placements, different modules with different themes and theory. Consequently, it would not be possible to continue data collection across this threshold between the two curricula.

The only real decision was whether to complete the study on the three remaining cohorts of the current curriculum or wait for 18 months before performing the similar but different research on a similar but differently constructed aggression prevention and management unit. The former choice was the one selected and the last three cohorts on the 1996 curriculum were included in the study. Each cohort had approximately

80 students and so this arrangement would permit the effects of training on a sample of approximately 240 student nurses to be evaluated.

In relation to the examples of published training evaluations discussed in Chapter 1, this sample size makes the study by far the largest published on student nurses and close to the largest reliable sample of 317 trained staff quoted in a now very dated piece of research (Gertz 1980). It will be remembered from the review in Chapter 1 that over half the studies reported had less than 100 respondents.

4.6.2 Pattern of Questionnaire Administration

With regard to the number and frequency of questionnaire administrations, once again, several issues had to be considered.

In determining the number of separate administrations of the questionnaire within the interrupted time series design an attempt was made to improve on a simple pre-test/post-test design. Administratively and practically it was felt possible to incorporate two additional administration of the questionnaire, one before the unit was delivered and one afterwards.

The agreed data collection points were therefore:

- In the last week of Trimester One – Data Collection Point 1
[PRE 1]
- At the beginning of the first day of the Unit– Data Collection Point 2 [DAY 1]

- At the end of the third day of the Unit– Data Collection Point 3
[DAY 3]
- In the last week of Module Three – Data Collection Point 4
[FOLLOWUP]

These measures produce some of the advantages identified earlier by Robson (1993) without incurring the disadvantages related to fatigue, boredom or testing effects (Fife-Schaw 2001a). In particular the design allowed some measure of test-retest reliability to be determined by comparing the scores obtained on the first two administrations of the questionnaire, before any educational or training intervention has been made.

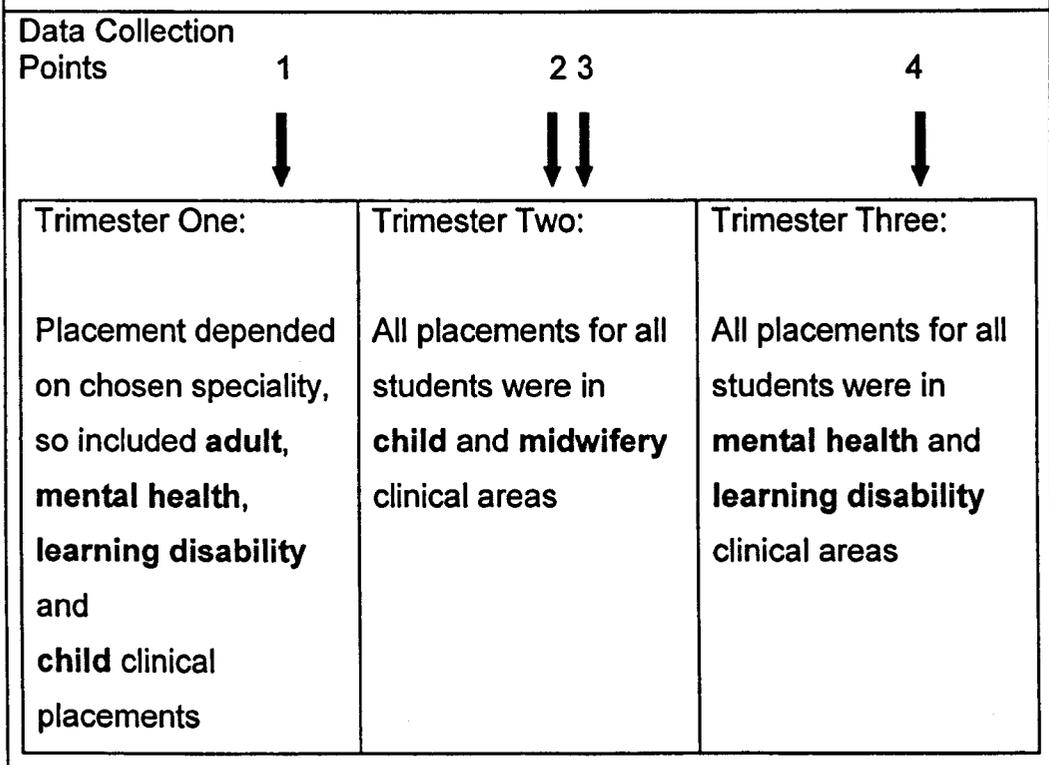
It also permits some monitoring of the enduring effects of the unit as the student nurses interact with the 'real world' of clinical practice and proceed through two three-week clinical placements, one in a mental health setting and one in a learning disability area. Moreover, as previously indicated, the staggered nature of the design provides opportunities for unanticipated and potentially confounding organisational changes to become manifested in the scores of later groups.

The time interval between the first and second administration of the questionnaire was approximately 16 weeks, similar to the interval between the third and fourth administration -12 -13 weeks. The time

interval between the second and third administration was the length of the unit for each student nurse – between 3-4 days. This arrangement meant that the stability of test scores over the period prior to completing the unit could be compared with those over a very similar period after its completion. Also, the immediate effects on students of completing the three-day unit can be measured.

The design of the study also effectively spanned three trimesters for each student cohort and permitted data on number and type of violent incidents involving students in three different types of setting to be gathered at Data Collection Points 1, 2 and 4. Specifically, Data Collection Point 1 referred to the placement in Trimester One, Data Collection Point 2 asked about the placements in Trimester Two, and Data Collection Point 4 asked about the placements in Trimester Three. Figure 4.1 shows diagrammatically this arrangement as it applies to one student cohort and includes an indication of the types of placements that different students would be completing in each Trimester [the placement in Trimester One depended on each student's destined speciality].

Figure 4.1: The study design applied to a single student cohort showing trimester structure, placement types and data collection points



At the time it was anticipated that the study design – 3 cohorts times 80 students times 4 administrations - would result in approximately 1000 questionnaires for subsequent analysis. Figure 4.2 provides a clear plan of the 17 months timescale over which the questionnaire was administered to the three different cohorts, and also indicates diagrammatically the various points within their training programmes when data were collected.

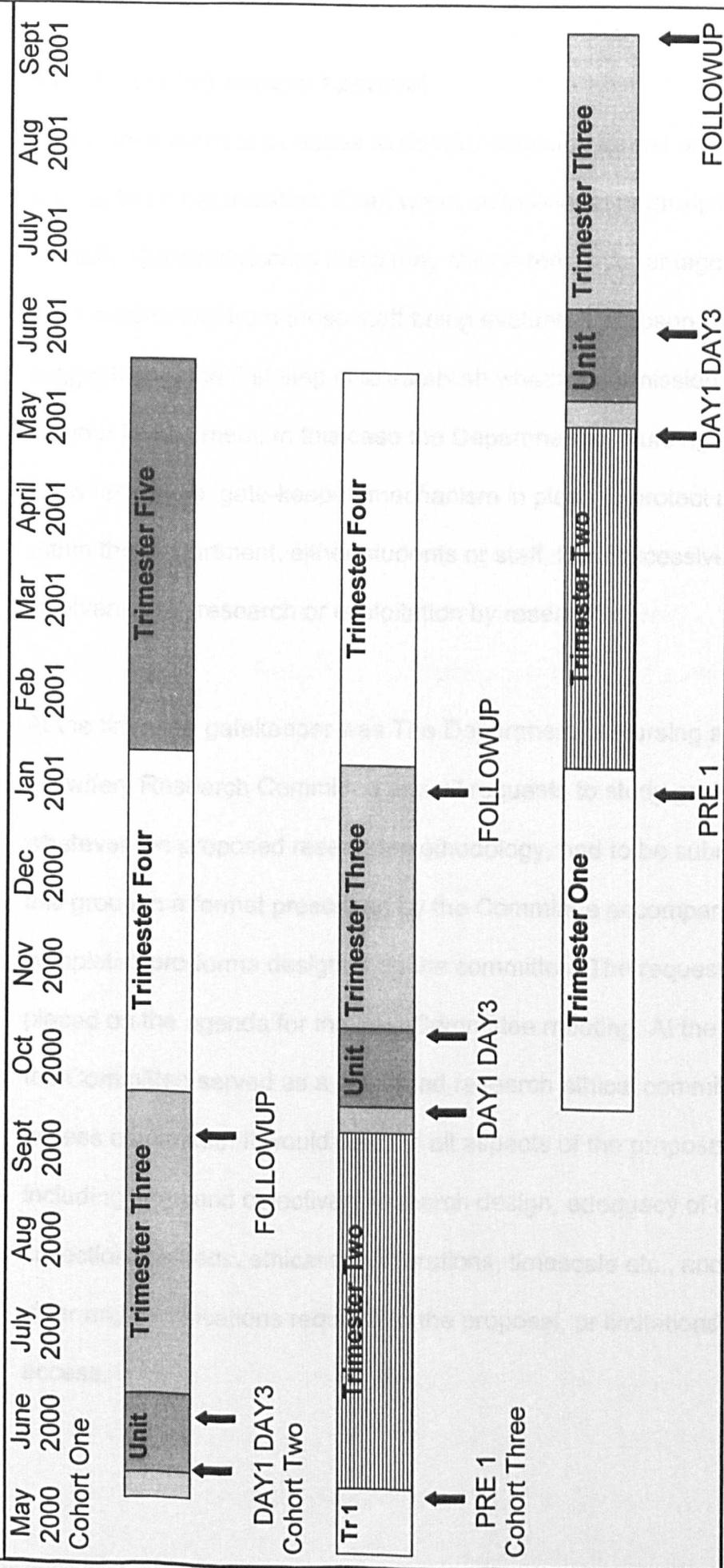
The staggered arrangement is clearly evident in Figure 4.2 and shows, for example, that Cohort One was in Trimester Three undertaking the module at a similar time to when Cohort Two was being tested for the first time at the end of Trimester One. Similarly, at approximately the

same time as Cohort Two was in Trimester Three completing the module, Cohort One was undergoing follow-up testing at the end of Trimester Three. Hence any changes in the wider organisation manifested in clinical settings are likely to be revealed in test scores taken at these points, given that a broad range of clinical placements in several specialities was being used in Trimester One and even more mental health and learning disability placements were being accessed in Trimester Three.

Figure 4.2 also shows that, in the event, Cohort One was tested with the questionnaire on only three occasions, not four as originally planned. The first planned administration of the question at the end of Trimester One [PRE 1] was not performed for this group. This came about because of the relative timings of creating and developing the questionnaire, the rolling programme of 'real world' student nurse training, and the wish to include as many student cohorts as possible before the curriculum changed.

It should also be noted that, although a new pre-registration nurse training curriculum was introduced in January 2001 as planned, the cohorts of students already in the system continued on their own approved curriculum pathway. Cohort Three was still on the old curriculum, the same as the previous two studied cohorts, and no results were confounded by this new curriculum initiative.

Figure 4.2: Time scale of data collection for the three cohorts of student nurses included in the study



4.6.3 Obtaining Access Approval

There are a number of issues to do with obtaining formal and informal access to an organisation. Even when an individual or group has formally approved access there may still be reticence, antagonism or overt resentment from those staff being evaluated. Robson (2000) suggests that the first step is to establish whether permission is needed for your involvement. In this case the Department of Nursing and Midwifery had a 'gate-keeper' mechanism in place to protect any group within the Department, either students or staff, from excessive involvement in research or exploitation by researchers.

At the time, the gatekeeper was The Department of Nursing and Midwifery Research Committee and all requests to study groups, whatever the proposed research methodology, had to be submitted to this group in a format prescribed by the Committee accompanied by completed pro-forma designed by the committee. The request was then placed on the agenda for the next Committee meeting. At the meeting, the Committee served as a combined research ethical committee and access committee. It would discuss all aspects of the proposal, including aims and objectives, research design, adequacy of data collection methods, ethical considerations, timescale etc., and make clear any modifications required to the proposal, or limitations placed on access.

In this case a proposal and completed pro-forma were submitted to the Committee [See Appendix Three for Completed Request for Access Form]. The proposal succinctly outlined the background and importance of the research before identifying the three cohorts for which access was requested. Permission to proceed with the research was granted and approval was given for access to all three cohorts with one proviso - that at some later stage the Committee was informed of the findings.

4.6.4 Identify Opportunities to Access Student Nurses

Once permission to perform the research had been obtained then the researcher made arrangements to access the student nurses in order to inform them of the situation and, hopefully gain their willing participation. Access to student nurses was judged to be easiest when they were gathered in large groups for taught sessions within the Department of Nursing and Midwifery, rather than attempting to see students in smaller groups or individually on clinical placements. On all occasions, at each data collection point, several 'appointments' had to be negotiated since a single cohort was invariably subdivided and taught in at least two smaller groups.

For the initial administration of the questionnaire a session was identified in the timetable of a taught block near the end of Trimester One and a thirty-minute period negotiated with the Lecturer due to deliver the session. Similarly for the final 'follow-up' administration of the questionnaire a taught session was identified in the timetable of the last

week of Trimester Three and a thirty-minute period of access negotiated with the Lecturer. The second and third administrations of the questionnaire were planned to occur within the Unit and so easily under the control of the researcher.

4.6.5 Devising Identification Number Coding System

The longitudinal, repetitive nature of the research design meant that a system was required by which a respondent's scores at different times could be compared. At the same time, in order to achieve this objective, it was not necessary to include the respondent's name on the form. Hence, the plan was to provide the student nurses with confidentiality but not anonymity.

An Identification Number scheme was devised prior to presenting details of the study to each cohort of student nurses. It was known that attendance at each session was recorded on a group attendance register. A copy of the attendance register for each group involved in the study was adapted to include a different number against the name of each student. The adapted form was then transferred to an acetate sheet, which could be projected from an overhead projector on to a screen within a session, thus allowing the student nurse completing a questionnaire to insert their correct Identification Number at the designated position, at the top right-hand corner of Page One of the questionnaire.

4.6.6 Gain Access to Student Nurse Groups, Present Details of Study and Encourage Participation

There were several tasks to be completed in the limited time period allocated for student access. This was particularly the case on the first occasion [PRE 1] when the rationale and background for the research had to be given, participation encouraged, queries answered, and instructions given about completing the questionnaire, which was being seen for the first time.

In addition, the researcher was meeting the group of 40-50 student nurses for the first time and attempting to manage any nervousness associated with this. Hence, it was felt necessary to create a 'script' in order to standardise the presentation, deliver material in the most logical sequence, maximise the clarity of the presentation, prevent any omissions, present all details to permit informed choice, and encourage participation. A script was devised beforehand [See Box 4.6] and used consistently throughout each initial meeting.

This first meeting was also used as an opportunity to emphasise the repetitive nature in a time series design, the desirability of continued participation, and highlight clearly that the *same* questionnaire would be presented to the students on a number of occasions in the next few months. As previously indicated, an attempt was made within the presentation to remove any possible bias caused by association with researcher by stressing that the research was being conducted and

coordinated by staff at the Institute of Work , Health and Organisations, Nottingham University, as clearly evidenced by the logo and letterhead at the top of Page One of the questionnaire.

A number of very important ethical issues were also incorporated in to this first meeting. For instance, any possible coercive effects were minimised by clearly indicating within the script that they were under no pressure to participate on this occasion, and that, if they did, then they were free to opt out on any future occasion, without fear of prejudice etc. They were also made aware of the lack of anonymity inherent in the research design, the reasons for it and the measures taken to ensure confidentiality, including the separate storage of Identification Numbers and completed questionnaires.

Furthermore the possibly emotive nature of the subject was acknowledged. Some students would have been involved in aggression management as part of their previous job. More would have been exposed to violence or aggression since they commenced the training. Consequently, some student nurses may have already sustained physical or psychological injuries. Therefore every effort was made to treat the subject sensitively. Reassurance was offered at the time, as was the opportunity to seek support from either the researcher or the student nurse's personal tutor should mention of the topic cause any distress at any point in the near future.

Box 4.6: Script for briefing students about the study and gaining their consent for participation

Hello, I'm not [X X] the lecturer you are expecting but Bernard Beech. I am a lecturer on the Mental Health branch with an interest in violence in health settings. I am a registered trainer in this area. I will be with you a little later in the course in Trimester 3 for 3 days when we will complete a unit on "preventing and managing violence in health care settings". You might already be aware that this is an area of interest for the Department of Health and currently subject to a lot of study. In this regard, the Departmental Research Committee and Head of Pre-Registration Training [X X] have given permission for this unit to be studied by a Group from the Institute of Work, Health and Organisations at the University of Nottingham, of which I am also a part.

This study involves using a questionnaire with certain student groups at various points in their training – a while before completing the unit, at the very start and end of the Unit, and at some more distant point in their training.

Your group is one of those selected. Today I have negotiated 30 minutes at the start of this session to administer the first questionnaire and I wanted to just point out a few things before asking you to complete the form:-

- The same form will be presented on each occasion.
- I should emphasise that you are not forced to complete it - participation is completely optional - and you are free to opt out of this or future rounds - without any detriment or repercussions to your progress.
- All data collected will be treated and stored confidentially but the design of the study - repeated measures - means that it cannot be anonymous.
- It firstly asks for some details of yourselves and your recent experiences but this will be stored anonymously (using ID No without names). I assure you that the list of names and corresponding ID numbers will be stored separately and securely.
- The questionnaire then asks three different sorts of questions about violence – open and analogue (mark a line) questions on pages 1 and 2 and a set of Likert-style statements on page 3 and there are more instructions with each type of question about how to respond to it. [For the analogue and Likert-style statement questions it is probably better not to dwell too long on each statement before making a response].

Are there any questions or anything further that you wanted to know?
[Answer any questions or queries]

I will now display on the screen a form with your name on it. You will see that each name has an identity number against it and this same number will be used on all occasions when you complete the form. If you are willing to participate in the study then you should right this number at the top right hand corner of the form now and then go on to answer each of the questions.

4.7 SUMMARY STATEMENT

This Chapter has attempted to move from the generalities of research texts on design and data collection methods through to a detailed description and justification of the precise research design, data collection methods and processes used in this study. In so doing it has highlighted the rationale for decisions made, acknowledged the restrictions imposed by attempting to perform research in a 'real world' organisation, and indicated the measures chosen to overcome or minimise these restrictions. The remaining two Chapters of this work will go on to present the results [Chapter 5] of the study outlined in this Chapter, and then discuss these results in relation to the originally posed Research Questions and Research Hypotheses, and relate them to the findings of other researchers [Chapter 6].

CHAPTER 5 – RESULTS

5.1 INTRODUCTION

This chapter will report the analysis of data and present findings. A number of recognised preliminary stages are suggested for this process prior to any statistical analysis of the data. Robson (1993) suggests that, firstly, a coding system has to be devised for each variable. Some mention of this was made in the previous chapter with regard to the open-response questions on the questionnaire. A similar scheme is required to record responses to the other types of question. For example, a method of recording responses to the questions on gender, age, speciality branch etc. is required, as is a method of recording responses to the 5-point Likert statements. A coding method for indicating missing data has also to be devised. Once this was agreed then a 'data set' has to be designed and created, with rows corresponding to respondents and columns corresponding to variables. In this case all the data were entered in to a database created using The Statistical Package for the Social Sciences (SPSS) Version 10 software. Data is then entered in to the database using the numerical codes or direct scores.

Prior to any statistical exploration, the data set must be cleaned and Robson (1993) suggests a number of straightforward techniques for achieving this. Simple frequency plots could be performed for each variable to highlight the presence of invalid scores. Oppenheim (1992)

and Bowling (1997) advocate the same procedure and additionally suggests the use of range checks, so, for example, when using 5-point Likert scales, only values in the range 1-5 plus a value for missing items, 9 say, should be present. In addition, some checks of internal consistency can be made via cross-tabulations of variables to detect impossible correlations.

Oppenheim (1992) also highlights the decisions to be made about missing data, and distinguishes between the treatment of the odd missing response, many omissions and the complete non-respondent. The major concern is the risk of possible bias more so than reduced sample size. Oppenheim (1992:280) cautions that the researcher needs to be satisfied " the reasons or causes for non-response or missing data are unconnected with the topic in our questionnaire, so that there are no '*correlated biases*'. Programmes offer the facility of deleting from statistical tests on a listwise or pairwise basis, the former removing all of a respondent's data from analysis, while the latter only temporarily removes a respondent from those tests for which they have incomplete entries.

In addition, some further preparatory work might be required before certain statistical tests can be performed. In particular, the researcher needs to ensure that variables satisfy certain assumptions about being normally distributed within a population before being exposed to more powerful parametric statistical analysis. Fife-Schaw (2000c:366)

concedes that “in practice you are unlikely to have access to information about the distribution of scores in the population” forcing the researcher to look only at the distribution of scores in the sample. Fife-Schaw (2000c:366) adds that “minor deviations from normality will not unduly undermine the value of many common parametric tests” and a measure of the approximation to normality can be determined by prior calculation of skewness and kurtosis.

All the stages identified above were completed prior to the analysis of the data collected for this study. That is to say:-

1. A coding system was devised for each variable and an agreed list of risk factor responses was determined for the open-response questions related to the two scenarios.
2. A data set was created within SPSS consisting of, initially, 142 variables and questionnaire data was entered in to the database.
3. The data set was cleaned. One source of error identified in this case was caused by the use of group registers to allocate Identification Numbers to student nurses. It was assumed that the registers were up-to-date but this was discovered to be incorrect. A small number of students had left the course prior to the first questionnaire being administered and this resulted in a small correction to the number of student nurses included in the sample [a reduction of 11, from 254 to 243] [See Table 5.1].

5.1.1. Measures of Reliability

McDowell & Newall (1987) suggest that reliability is concerned with the estimation of random errors in the administration or completion of scales. Burns (2000) identifies a number of synonyms for reliability that illustrate its importance as a characteristic of a measurement instrument, including, dependability, consistency, predictability and accuracy.

Bryman & Cramer (1997) identify two separate form of reliability, namely, internal and external reliability. Internal reliability [more often called *internal consistency*] "is particularly important in connection with multiple-item scales. It raises the question of whether each scale is measuring a single idea and hence whether the items that make up the scale are internally consistent" (Bryman & Cramer 1997:63). External reliability refers to the degree of stability or consistency of a measure used by different individuals over time, so-called *test –retest* and *inter-rater* reliabilities.

All three measures are calculated by measuring the correlations between, say, different combinations of items within a scale, different raters' scores, or scores obtained by the same rater at different times. Hence, a correlation coefficient is produced which ranges between 0 and 1. Since a variety of conditions affect the computation of scores "no assessment or techniques has a single reliability coefficient" (Burns

2000:345) and it is recommended that they be calculated on each administration (McDowell & Newall 1987, Gibbon 1995). Internal consistency can be calculated using a number of tests, including split half [dividing the items into two groups and comparing the correlation between each half and the popular Cronbach alpha that calculates the average of all possible split half divisions. Prior to calculation items must be coded in the same direction (Bryman & Cramer 1997) in terms of implied positive or negative attitude.

With regard to determining acceptable levels of reliability many authors (Hammond 2000) cite Nunnally (1978) who suggested that reliability coefficients should be > 0.7 before a scale can be used as a research tool. Gibbon (1995) suggests >0.6 is acceptable for most purposes. Streiner & Norman (1989:89) suggest that sample size affects the determination of acceptable reliability "since a sample of 1000 can tolerate a much less reliable instrument than a sample of 10".

The Cronbach alpha formula involves the number of items and, thus, the number of items in the scale affects the value of alpha obtained [few items equating with lower scores]. Hence a balance has to be made between the length of the scales, possible fatigue or non-completion, and reliability coefficients, although Hammond (2000:187) counsels that "a small number of items does not excuse poor reliability estimates". However, Leather *et al* (1998b) suggest "the mean inter-item correlation is a more appropriate estimates of the reliability of small scales". Values

of this mean inter-item correlation in the range of 0.1 – 0.5 are considered acceptable (Cox & Ferguson 1994).

With regard to test-retest reliability “there is no standard duration of time which should separate two administrations, however a minimum of one day and, perhaps a maximum of one year” are generally considered acceptable (Burns 2000). Obviously, the characteristic being measured must remain stable during the period. Bowling (1997) suggests using Cohen’s Kappa, weighted Kappa or Pearson product moment, depending on whether the data is of nominal, ordinal or higher level, to determine test-retest and inter-rater reliability.

McDowell & Newall (1987) make the important point that different sorts of reliability become most important under different applications. Sometimes stability over time is crucial [for example when making predictions], whilst on other occasions, internal consistency is most important.

In this study a number of measures of reliability are appropriate. Measures of test-retest reliability can be applied between scores obtained at t1 and t2 on several variables, including demographic variables, number of risk factors and attribution of blame in the two scenarios, and factors scores pertaining to the 24 Likert-type statements. In addition, when determining the open responses to the

number of risk factors identified in the two scenarios it is also desirable to obtain a measure of inter-rater reliability.

Internal consistency can be measured for the same variables on each administration. Obviously, measuring internal consistency between all 24 Likert statements prior to factor analysis makes no sense since these were designed to measure different aspects of change and some will be unrelated [uncorrelated] to others. Oppenheim (1992:201) asserts that as a first resort, factor analysis is “clearly a better way of ensuring uni-dimensionality” in a scale than internal consistency methods. However, once factors have been identified and factor scores calculated, then measurement of internal consistency and test –retest reliability for each factor has merit. Watson (1995:56) advised “carrying out a factor analysis of the data derived and then calculating the alpha coefficients of each factor”. This procedure was performed in this study.

In order to maximise clarity and appreciation of the findings it is thought appropriate to present a summary of the schedule of questionnaire administration at various data collection points and indicate the significance of each of these points, prior to any results being presented. Figure 5.1 summarizes this information.

Figure 5.1: Summary of data collection schedule				
	4 months prior to Unit: t1	Day 1 of the Unit: t2	Day 3 of the Unit: t3	3 months after the Unit: t4
Cohort 1		✓	✓	✓
Cohort 2	✓	✓	✓	✓
Cohort 3	✓	✓	✓	✓
Significance and relationship to clinical placement	At end of Trimester One after completion of first placement in destined speciality area. Could relate to Adult, Child, Mental Health or Learning Disability settings	At start of Trimester Three. Clinical placement questions relate to Trimester Two when all students completed Child and Midwifery placements	3-4 days after t2. Students did not complete questions related to their clinical placement	At end of Trimester3. Clinical placement questions referred to Trimester Three when all students completed placements in Mental Health and Learning Disability settings

Once all this preparatory work has been completed, Bowling (1997) presents a sequence to data analysis. Similarly, Oppenheim (1992:281) suggests that the "analysis of a typical survey will usually have to go through several predictable stages:

- a. Uni-variates;
- b. Bi-variates;
- c. Multi-variates;

d. Special sub-group studies”.

This is the order that will be used predominantly to structure the presentation of findings within this Chapter.

5.2 SAMPLE

5.2.1 Size And Response Rate

Table 5.1 illustrates the sample size and response for each group/cohort of student nurses on each occasion that the questionnaire was administered. It also shows the corrected number of students in each cohort, determined by retrospectively checking the records of those students who failed to complete any questionnaires in order to determine their status i.e. left the course, sick, transferred etc.

Cohort	Sample Size		Response Rate N, %			
	Original number on register	Corrected number	Pre- unit t1	Day 1 t2	Day 3 t3	Follow-up t4
1a	40	38	**	36, 94.7	36, 94.7	33, 86.8
1b	45	41	**	38, 92.7	35, 85.4	35, 85.4
Total	85	79		74, 93.7	71, 89.9	68, 86.1
2a	39	38	34, 89.5	34, 89.5	33, 86.8	34, 89.5
2b	38	37	37, 100	35, 94.6	34, 91.9	26, 70.3
Total	78	75	71, 9.7	69, 92.0	67, 89.3	60, 80.0
3a	47	44	38, 86.4	41, 93.2	41, 93.2	38, 86.4
3b	45	45	42, 93.3	45, 100	45, 100	42, 93.3
Total	92	89	80, 89.9	86, 96.6	86, 96.6	80, 89.9
Total	254	243				
** No data collected						

A sample of 243 students was eligible to participate in the study. Consecutive cohorts contained 79, 75 and 89 students respectively. The response rate for any group of students [approximately half of a cohort] on any occasion ranged from 85.4% to 100%.

5.2.2 Homogeneity Of The Three Cohorts In The Sample

An early task was to ascertain that the student nurses in the three cohorts that constituted the sample were broadly similar. This was necessary since the later intention was to combine the results for each cohort in order to summarize the effects of training on student nurses. Therefore, before data from the three cohorts could be combined and processed it was essential to show that each cohort was similar on important demographic variables. Table 5.2 shows summaries for separate cohorts with regard to the main demographic variables, namely, gender, age, destined branch, previous experience, and experience of violence during placements.

Obvious differences in this table are the absence of any learning disability students from Cohort Two, the increase in adult branch students in Cohort Three, and the absence of material on involvement in violent incidents for Cohort One in Trimester One. Furthermore, Cohort Two reported experiencing more verbal violence in Trimester One than Trimester Three while Cohort Three reported the opposite effect. Statistical analysis of different types can be applied to this data. Tests to demonstrate normality via measures of skewness and kurtosis

were performed and normal distribution was assumed if skewness was <2.0 and kurtosis <5.0.

Table 5.2: Summaries for each cohort on demographic variables N, %			
Variable	Cohort 1	Cohort 2	Cohort 3
Age yrs mean[s.d.]	25.84 [7.78]	26.95 [7.73]	26.14 [7.90]
Gender:			
Male	8, 10.1	13, 17.3	14, 15.7
Female	69, 87.3	62, 82.6	75, 84.3
Missing	2, 2.5		
Destined Branch:			
Adult	38, 48.1	46, 61.3	50, 56.2
Child	9, 11.4	9, 12.0	10, 21.3
Learning Disability	11, 13.9	-	12, 13.5
Mental Health	19, 24.1	20, 26.7	17, 19.1
Missing	2, 2.5		
Previous Training:			
Theory	13, 16.5	16, 21.3	15, 16.9
Breakaways	12, 15.2	13, 17.3	12, 13.5
Restraint	12, 15.2	8, 10.7	9, 10.1
Number of students involved in incidents:			
Trimester One			
Verbal	-	38, 50.7	28, 31.5
Physical	-	23, 30.7	19, 21.3
Trimester Two			
Verbal	5, 6.3	10, 13.3	5, 5.6
Physical	3, 3.8	2, 2.7	2, 2.2
Trimester Three			
Verbal	40, 50.6	30, 40.0	35, 39.3
Physical	31, 39.2	25, 33.3	34, 38.2

One factor independent groups analysis of variance [ANOVA] can be used if the data are shown to be of a higher level of measurement [interval, ratio] and normally distributed [for example, age]. Alternatively,

a Chi –square test can be applied to data measured at lower categorical levels [for example, gender, destined branch, previous training and involvement in incidents].

ANOVA results showed that the key variable of age was normally distributed within cohorts. The Chi-square statistical analysis demonstrates that, with one exception, there are no statistically significant differences between the students in each cohort with regard to the key demographic variables, although the numbers in each destined branch was almost significant. The exception is verbal violence during Trimester One and it is unfortunate that this variable could not be measured for Cohort One. Otherwise, the composition and characteristics of the three cohorts are considered very similar and, so, able to be used in a combined form. [See Table 5.3].

Table 5.3: Statistical analysis of homogeneity of the three cohorts on significant demographic variables

ANOVA Test	Sum of Squares	df	Mean Square	F	Sig
Age					
Between Groups	49.091	2	24.546	.403	.669
Within Groups	14262.234	234	60.950		
Total	14311.325	236			
Pearson Chi-Square Test					
		Value	df	Asymp. Sig. (2- sided)	
Gender		1.641	2	.440	
Destined Branch		12.533	6	.051	
Previous Training :Theory		1.097	2	.578	
Previous Training :Breakaways		0.684	2	.710	
Previous Training :Restraint		1.047	2	.592	
Trimester One experience					
Violence - verbal		4.875	1	.027 *	
- physical		1.273	1	.259	
Trimester Two experience					
Violence - verbal		4.255	2	.119	
- physical		0.333	2	.847	
Trimester Three experience					
Violence - verbal		3.349	2	.187	
- physical		0.324	2	.851	
* significant at the .05 level (2-sided)					

5.3 RESULTS – DEMOGRAPHIC DATA

5.3.1 Age, Gender, Destined Branch, Previous Training

Table 5.4 shows the demographics for the combined sample. The mean age of respondents was 26.3 years [s.d. 7.79 years], the youngest students being 18 years and the oldest 52 years. Two-thirds of the student nurses were under 30 years old, and over 90% were under 40 years old. Just under 85% were female. A little over half were destined for adult nursing with approximately one quarter intending to be mental health nurses and approximately 10% each going in to child and learning disabilities branch. No conclusion can be drawn regarding whether these figures are typical of the constitution of student nurse cohorts throughout the country since no national figures about the constitution of student nurses in training, with regard to gender, age speciality, are kept by the Department of Health or NMC.

Approximately 20% had previously received some theoretical instruction related to aggression prior to starting the training programme, with progressively smaller numbers receiving breakaway or restraint training.

Table 5.4: Student nurse demographics

Category	n	%
Gender:		
Male	35	14.4
Female	206	84.8
Undisclosed	2	0.8
Age:		
18-19	43	17.7
20-29	119	49.0
30-39	60	24.7
40-49	12	4.9
50 plus	3	1.2
Undisclosed	6	2.5
Destined Branch:		
Adult	134	55.1
Learning Disability	23	9.5
Child	28	11.5
Mental Health	56	23.0
Undisclosed	2	0.8
Previous Training:		
Theory - Yes	44	18.1
- No	187	77.0
- Undisclosed	12	4.9
Breakaways - Yes	37	15.2
- No	194	79.8
- Undisclosed	12	4.9
Restraint - Yes	29	11.9
- No	203	83.5
- Undisclosed	11	4.5

5.3.2 Experience of Violence

The questions related to experience of violence and 'involvement' in either verbal or physical, during the placement showed that on all occasions and in all clinical placement areas verbal incidents were more common than physical ones. Potentially, any interpretation of the term 'involvement' could range from being present in the vicinity, having a minor part, through to being the focus of the incident. During the initial briefing the term 'involvement' was clarified as being towards the more severe end of this dimension and 'defined' as the student nurse being the focus of a verbal or physical incident or playing a major part in the incident and its management.

The highest rates were noted in the mental health and learning disability placements completed in Trimester Three where a little over half of the student nurses experienced incidents involving verbal abuse or threats, while slightly fewer were involved in incidents of physical violence. The lowest rates occurred in Trimester Two [Child and Midwifery placements] where less than 10% of students encountered verbal threats and seven students [3%] experienced physical violence.

Trimester One involved the student completing a clinical placement in their destined speciality area and consequently included mental health, learning disability, adult, and child placements. The numbers of students involved in incidents during this Trimester [based on only two cohorts] are between the two extremes of Trimester Two and Three

with just over one quarter experiencing physical violence and approximately 43% verbal threats or abuse. Table 5.5 shows the relative numbers of student nurses reporting involvement in verbal or physical incidents in each of the three Trimesters from data collected at points t1, t2, and t4.

Table 5.5: Student nurses involvement in verbal or physical violent incidents during three trimesters				
	Verbal Incidents		Physical Incidents	
	n,	% of total	n,	% of total
Trimester One [All clinical areas]	66	43.4	42	27.6
Trimester Two [Child and Midwifery placements]	20	8.7	7	3
Trimester Three [Mental Health and Learning Disability placements]	105	50.5	90	43.5

As previously mentioned during Trimester One the student nurses completed a placement in their destined branch speciality. It was possible to further analyse the data relating to this first placement using bivariate statistical analysis. A crosstab chart showed the different numbers of students answering 'yes' and 'no' to the statements about being involved in verbal or physical violent incidents during this placement - see Table 5.6 [percentages exclude missing responses].

During this Trimester mental health student nurses had the highest percentage of verbal aggression, while mental health and learning disability students reported the highest percentages of physical aggression. Student nurses working in child settings reported the lowest incidence of both verbal and physical violence. Application of the Pearson Chi-square test showed these figures to be statistically significant, both for verbal incidents [12.536, d.f. 3, p= .006], and for physical violence [9.348, d.f. 3, p= .025].

Table 5.6: Crosstabs breakdown of student nurse involvement in verbal and physical violence in trimester one by placement type, N,%

	Trimester One - Placement type				
	Adult	Learning Disability	Child	Mental Health	Total
Experience verbal violence during placement –					
Yes	41, 45.1	4, 44.4	1, 5.9	20,57.1	66, 43.4
No	50, 54.9	5, 55.6	16,94.1	15,42.9	86, 56.6
Experience physical violence during placement –					
Yes	25, 27.5	4, 44.4	0, 0	13,37.1	42, 27.6
No	66, 72.5	5, 55.5	17, 100	22,62.9	110,72.4

5.3.3 Number of Incidents

Additionally, when involvement was indicated, the questionnaire asked the students to quantify the number of incidents of different types that they were involved in. A small number of respondents omitted to provide a numerical estimate of the number of incidents despite indicating that they were involved in incidents. A handful of others indicated the number of incidents using a vague statement, for example, 'daily', 'numerous', or 'hundreds', rather than specific numbers.

These responses presented some difficulties to analysis. Whilst obviously indicating a serious problem it was decided to ignore these responses when determining the frequencies [hence the slight difference in N for Trimesters One and Three between Tables 5.5 and 5.7]. For convenience the data was reduced to a limited number of categories, '1-2 incidents' and '5 or less incidents'. Summarised results are presented in Table 5.7.

As seen in the previous section, the smallest number of students reported involvement in any number of violent incidents of any type during Trimester Two when the students attended child and midwifery placements. This Trimester also showed the smallest range for number of incidents for each student who reported incidents, with regard to both verbal and physical aggression. Students were involved in the highest number of both verbal and physical incidents during the clinical

placements on mental health and learning disability wards/unit during Trimester Three. Again Trimester One, with its broader range of adult, child, mental health and learning disability placements is between the two extremes of Trimester Two and Trimester Three.

Table 5.7: Numbers of incidents [verbal and physical] which students were involved in during placements				
	No. of Students reporting incidents	1-2 incidents n, %	5 or less incidents n, %	Range
Trimester One placements	Verbal [n=61]	22, 36	56, 91	1-15
	Physical [n=39]	30, 76.9	38, 97.4	1-7
Trimester Two placements	Verbal [n=20]	13, 65	18, 90	1-14
	Physical [n=7]	6, 85.7	7, 100	1-5
Trimester Three placements	Verbal [n=97]	41, 42.3	79, 81.4	1-20
	Physical [n=87]	60, 69	81, 93.1	1-10

The varied placements that occurred in Trimester One allowed an analysis of the different levels of violence occurring in different placements types at any time to be investigated in relation to the students destined for different specialities. Unfortunately most of the scores for the variables relating to number of incidents verbal or physical during placements were not normally distributed and it was necessary to transform this data into square root form before it could be analysed using parametric statistical tests. When one-way ANOVA for Independent Samples is performed on the square root data for number

of incidents against destined branch then no score for Trimester One were statistically significant.

5.4 RESULTS - SCENARIOS

Data collected for each of the two scenarios consisted of the number of risk factors identified in the scenario and the attribution of blame for causation of the incident. Analysis of these variables will be considered separately.

5.4.1 Number of Risk Factors

The number of risk factors identified in the scenario was determined via interpretation of the student nurse's answer to the open response question. As described in the previous chapter, an initial list of 'acceptable' or 'approved' responses for risk factors in the two scenarios was generated at the time that they were written. Suitable items were listed under the sub-headings of Staff, Assailant, Environment, and Task.

This list was subsequently increased by reading the responses in the completed answers of a sample [n=250 approximately] of returned questionnaires – See Box 4.3 and Box 4.4. Once again it was ensured that all risk factors were delineated so as to only appear under one of the four sub-headings. Consistency rules were also devised to manage the counting of similar or equivalent words or phrases. When calculating

a participant's score only phrases or words that appeared in the accepted/ approved lists were counted.

In addition the stability of these approved risk factors was determined via a measure of inter-rater reliability. The responses for risk factors identified in Scenarios One and Two of a sample [one cohort at one time point] were chosen and subjected to separate analysis by a colleague who had been briefed in the response frameworks and exclusion criteria. Statistical analysis was performed on the resulting data. Percentage agreement and Pearson Product Moment correlation tests were calculated and the results of this inter-rater analysis are shown in Table 5.8. For each scenario thirty-six of the 38 student responses were the same [almost 95 % agreement]. In both cases the correlation coefficient was at a very acceptable level of equivalent interpretation of student responses by both scorers using the two frameworks.

Table 5.8 : Inter-rater reliability results for number of risk factors identified in scenarios				
	Percentage Agreement %	Correlation Pearson Product Moment		
		N	r	Significance
Scenario One	94.7	38	0.987	0.000**
Scenario Two	94.7	38	0.992	0.000**
** Correlation significant at 0.01 level [2 tailed]				

The measurement level of these two variables made them potentially amenable to parametric statistical analysis. Measures showed the variables to be normally distributed with no evidence of excessive skewness or kurtosis. Paired sample t –tests were performed on the data. The results are presented in Table 5.9 and show that the mean number of risk factors increased by almost three-quarters of a risk factor for both Scenarios [0.71 for Scenario One and 0.73 for Scenario Two] between t2 and t4, over the course of completing the module and associated clinical placements.

Furthermore, the range of number of risk factors identified by students increased from 8 before the unit to 10 after the unit for Scenario One and the maximum number of factors identified increased from 9 prior to the unit, to 11 immediately after, and 12 at follow-up for Scenario One. Similar increases also occurred for Scenario Two, where the range increased from 7 prior to the Unit to 9 immediately afterwards to 12 at follow-up, and the maximum number of factors identified increased from 8/9 before to 10 immediately afterwards to 13 at follow-up.

With regard to the statistical analysis, Table 5.9 also identified there was no significant change between t1 and t2 for either Scenario One and for Scenario Two. This result can be seen as an initial indication of test-retest reliability for this part of the questionnaire, more of which will be said later [section 5.4.3]. A statistically significant change between t2 and t3 was noted for both Scenarios on immediately completion of

Table 5.9: Number of risk factors identified at four time points for scenarios – mean scores, min, max, range, and paired sample t-test				
Number of risk factors identified [Mean scores, range, min, max]				
Scenario One	pre unit [t1]	day 1 [t2]	day 3 [t3]	follow-up [t4]
Mean	4.19	4.18	4.66	4.89
Range	8	8	10	10
Min	1	1	1	2
Max	9	9	11	12
Scenario Two				
Mean	4.52	4.37	4.88	5.10
Range	7	7	9	12
Min	2	1	1	1
Max	9	8	10	13
Number of risk factors identified -Paired sample t-test				
Scenario One :		t	df	Sig. (2-tailed)
Pair 1:	pre unit – day 1	0.580	140	0.954
Pair 2:	day 1 – day 3	-4.098	206	0.000**
Pair 3:	day 1 – follow-up	-6.958	196	0.000**
Scenario Two :				
Pair 1:	pre unit – day 1	1.338	140	0.183
Pair 2:	day 1 – day 3	-4.884	206	0.000**
Pair 3:	day 1 – follow-up	-5.826	196	0.000**
**Significant at the less than 0.01 level				

the Violence Management Unit and this was maintained at time t4 following the two clinical placements, with a statistically significant difference between t2 and t4 scores being evident. These findings suggest that attendance on the Unit had an immediate desirable effect on the number of risk factors identified by student nurses in both

scenarios, and that this change endured throughout the two subsequent clinical placements.

5.4.2 Attribution of Blame

The attribution of blame for causation of the incident was calculated via measurement of the point at which the 10 cm VAS line was intersected.

The two variables were considered to be of interval or ratio level and potentially amenable to parametric statistical analysis since they were normally distributed with no evidence of excessive skewness or kurtosis. Paired sample t-tests were performed on the data.

Table 5.10 shows the results for the questions relation to attribution of blame and, in this case, the changes are more complex. A reduction in score equates to a reduction in the amount of blame attributed to the patient in each scenario, and a consequent increase in the proportion of blame attributed to other factors, for example, the staff, environment or task. Hence, as the Unit promulgated a multi-factorial model of violence rather than a simplistic 'patient at fault' one, then a reduced score would be desired and seen as the Unit material exerting a positive effect.

For Scenario One, the mean score reduced by 0.24 between t1 and t2, although this was not statistically significant. It decreased significantly more [0.78] between t2 and t3, gaining statistical significance on completing the Unit, and rose slightly but remained statistically significant at t4 compared with t2. In addition the full range of scores [0.00-10.00] were recorded at t3 and the maximum score [10.00] was

used at all data collection points. For Scenario Two, mean scores were considerably lower at all points than in Scenario One, ranging between 2.72 at t1 and 3.03 at t4. The range of scores at all time points was similar to Scenario One and the lowest possible minimum score [0.00] was used on all occasions. Paired sample t-tests showed that there were no statistically significant changes between scores at any of the following pairs of time points; t1-t2, t2-t3 or t2-t4.

Table 5.10: Attribution of blame [length of VAS line] identified at four time points for scenarios – mean scores and paired sample t-test results				
Length of VAS line [centimetres] [Mean scores]				
Scenario One	pre unit [t1]	day 1 [t2]	day 3 [t3]	follow-up [t4]
Mean	7.14	6.90	6.12	6.43
Range	9.80	8.90	10.00	7.90
Min	0.20	1.10	0.00	2.10
Max	10.00	10.00	10.00	10.00
Scenario Two				
Mean	2.72	2.90	2.78	3.03
Range	8.55	9.50	9.45	9.50
Min	0.00	0.00	0.00	0.00
Max	8.55	9.50	9.45	9.50
Length of VAS line [centimetres] Paired sample t-test				
Scenario One :		t	df	Sig. (2-tailed)
Pair 1:	pre unit – day 1	0.9712	136	0.333
Pair 2:	day 1 – day 3	6.828	204	0.000**
Pair 3:	day 1 – follow-up	4.607	194	0.000**
Scenario Two :				
Pair 1:	pre unit – day 1	-1.193	140	0.235
Pair 2:	day 1 – day 3	0.880	207	0.380
Pair 3:	day 1 – follow-up	0.148	195	0.882
**Significant at the less than 0.01 level				

It appears that considerably less blame was attributed to elderly woman in Scenario Two than the young man in Scenario One, at all time points. The scores for Scenario One followed the predicted pattern, the mean score reducing after explication of a multi-factorial model. No discernable pattern could be noted for Scenario Two, the mean score on completing the Unit being between those recorded at t1 and t2. Indeed it appears that at t4, after completing the two clinical placements, the elderly woman is seen as more to blame than at the commencement of the Trimester or Unit.

5.4.3 Test-Retest Reliability

The stability of the scores in the four-month time period prior to the Unit being presented [between t1 and t2] is an indication of the test-retest reliability of the measure. This test-retest reliability can be quantified using the Pearson product moment correlation test. The results of this test for Scenario One and Two are shown in Table 5.11 and indicate positive correlations for risk factors in Scenarios One and Two ranging between $r = +0.457$ and $+0.472$. Correlations in the range 0.3-0.5 are described as low and between 0.5-0.7 as moderate (Hinkle et al 2003). Similarly the positive correlations between t1 and t2 for the attribution of blame for Scenarios One and Two are within the same low-moderate range, being $+0.513$ and $+0.436$ respectively.

Table 5.11: Test-retest reliability - correlation tests for scenario variables [number of risk factors identified and attribution of blame] at t1 and t2

	N	Pearson Correlation	Significance (2 tailed)
Scenario One			
Number of risk factors identified at t1 v t2	141	0.457	0.000**
Attribution of Blame at t1 v t2	137	0.513	0.000**
Scenario Two			
Number of risk factors identified at t1 v t2	141	0.472	0.000**
Attribution of Blame at t1 v t2	141	0.436	0.000**
** Correlation is significant at the 0.01 level (2-tailed)			

5.4.4 Bi-variate Tests

A number of tests were performed to ascertain whether demographic differences were associated with differences in performance, in terms of number of risk factors identified or the attribution of blame for causing the incidents described in the two scenarios. The variable 'age' was transformed in to two groups, based on the mean age of 26 years, that is, student nurses less than 26 years old and equal to or more than 26 years. Independent samples t-tests were performed on variables with two possible groups, for example, previous theoretical training or gender. Independent samples ANOVA was conducted when three or

more groups were identified, for example, when analysing the variable 'destined branch'.

Selected results are shown in Tables 5.12 and 5.13. Table 5.12 shows the relevant means for a number of pairings whilst Table 5.13 shows the results of statistical analysis . With regard to age, nothing significant or consistent was noted for the number of risk factors identified by students in either group [<26 or ≥ 26 years]. Only Scenario One at t1 produced statistically significant different scores, wherein older students identifying a mean of 0.45 more risk factors. On all occasions, for both scenarios, the mean score for blame attributed to the patient [young man in Scenario One and elderly woman in Scenario Two] by older student nurses was lower than that attributed by younger students. This difference was statistically significant [$p < .01$] for Scenario Two on all four occasions and on two occasions [t2 and t3] for Scenario One. It appears that older student nurses blamed the patient less than younger nurses did, and, presumably, incorporated other factors into their explanation of causation.

There were no differences noted between male and female students regarding the attribution of blame. There were gender differences regarding the number of risk factors identified in the scenarios. On all four occasions, for both scenarios, the mean number of risk factors identified by female student nurses was higher than that identified by male student nurses. The mean difference in scores ranged from 0.23

factors for Scenario One at t2 to an extra 1.10 factors for Scenario Two at t3. These differences were statistically significant at t1, t3, and t4 for Scenario One and at all times for Scenario Two.

Student nurses who had received theoretical training prior to starting nurse training identified more risk factors than those without previous training. On all occasions students with previous theoretical training identified slightly higher mean numbers of risk factors than those without prior training. However, these differences were only statistically significant on one occasion [Scenario One at t4] when an additional 0.68 mean risk factors were identified by those with prior theoretical training [t2.06, df 51.04, p= .045].

On all occasions student nurses with previous breakaway training identified higher mean numbers of risk factors than those without previous training although this difference was never statistically significant at any time. No obvious pattern or significant difference was noted for previous restraint training with regard to the number of risk factors identified or the attribution of blame in either scenario.

Table 5.12: Selected results for bi-variate analysis of demographic and scenario variables - means

Age v No of Factors	t1	t2	t3	t4
Scenario 1:				
>=26 years	4.42	4.13	4.74	4.95
<26 years	3.95	4.14	4.53	4.82
Scenario 2:				
>=26 years	4.40	4.27	4.63	4.96
<26 years	4.43	4.23	4.87	5.02
Age v Blame attributed				
Scenario 1:				
>=26 years	7.01	6.48	5.66	6.16
<26 years	7.18	7.20	6.40	6.62
Scenario 2:				
>=26 years	2.20	2.37	2.11	2.56
<26 years	3.03	3.29	3.26	3.31
Gender v No of Factors				
Scenario 1				
Male	3.61	3.94	3.87	4.30
Female	4.28	4.17	4.78	5.00
Scenario 2				
Male	4.00	3.91	3.88	4.20
Female	4.48	4.31	4.97	5.13
Previous theory training v No of Factors				
Scenario 1				
Yes	4.33	4.44	4.78	5.44
No	4.14	4.07	4.65	4.76
Scenario 2				
Yes	4.48	4.60	4.87	5.49
No	4.37	4.18	4.82	4.90
Previous breakaway training v No of Factors				
Scenario 1				
Yes	4.48	4.51	4.91	5.38
No	4.11	4.07	4.63	4.80

Scenario 2		t1	t2	t3	t4
	Yes	4.48	4.31	5.06	5.50
	No	4.38	4.25	4.78	4.92
Experienced violence-verbal v No of Factors					
Scenario 1					
	Yes	4.58	4.22	4.53	5.09
	No	3.87	4.05	4.17	4.60
Scenario 2					
	Yes	4.75	4.38	4.72	4.98
	No	4.15	4.05	4.60	5.07
Experienced violence- physical v No of Factors					
Scenario 1					
	Yes	4.57	4.30	4.41	5.24
	No	4.03	4.06	4.30	4.65
Scenario 2					
	Yes	4.76	4.22	4.68	5.57
	No	4.28	4.18	4.64	4.82
Experienced violence-verbal v Blame attributed					
Scenario 1					
	Yes	6.90	6.84	6.14	6.59
	No	7.30	7.03	6.52	6.53
Scenario 2					
	Yes	2.79	3.05	2.97	3.24
	No	2.67	2.74	2.86	3.05
Experienced violence- physical v Blame attributed					
Scenario 1					
	Yes	6.89	7.02	6.26	6.49
	No	7.23	6.93	6.39	6.58
Scenario 2					
	Yes	3.33	3.31	3.49	3.38
	No	2.49	2.72	2.70	3.04

With regard to experiencing verbal violence during Trimester One, on all but one occasion [Scenario Two at t4] for both Scenarios students who reported experiencing verbal violence during Trimester One reported a higher mean number of risk factors than those who did not report verbal violence, although this observation was only statistically significant at time t1 for both scenarios. On all occasions students who had experienced physical violence during Trimester One identified a higher mean number of risk factors for both scenarios than students who did not report physical violence during Trimester One, however the difference was only statistically significant at t1 for both Scenarios. These findings may suggest that exposure to verbal or physical violence in clinical settings and associated reflection and experiential learning allowed the students with prior experience of violence to more ably relate to the situations described in the scenarios.

In addition, student nurses who had experienced verbal violence in Trimester One attributed a greater mean amount of blame to the elderly woman in Scenario Two on all four occasions, and a lower mean amount of blame to the young man in Scenario One at t1, t2, and t3 than students who did not report experiencing verbal violence during Trimester One but these differences were not statistically significant. Moreover, student nurses who had experienced physical violence in Trimester One attributed a greater mean amount of blame to the elderly woman in Scenario Two on all four occasions, and a lower mean amount of blame to the young man in Scenario One at t1, t3, and t4

Table 5.13: Selected results for bi-variate analysis of demographic and scenario variables 'number of risk factors identified' and 'attribution of blame' for scenarios

	t1			t2			t3			t4		
	t	df	sig (2-tailed)	t	df	sig (2-tailed)	t	df	sig (2-tailed)	t	df	sig (2-tailed)
Age v No of factors												
Scenario One	2.06	117.75	.041									
Scenario Two		ns										
Age v Blame attributed												
Scenario One	-.573	129.7	.568	-3.19	194.8	.002**	-2.7	1723.0	.007**	-1.82	157.9	.071
Scenario Two	-2.61	137.7	.010**	-3.79	218.3	.000**	-4.44	207.7	.000**	-2.83	176.4	.005**
Gender v No. of factors												
Scenario One	-3.05	44.07	.004**		ns		-3.23	45.96	.002**	-2.03	40.47	.043*
Scenario Two	-2.25	38.57	.030*	-2.05	44.25	.046*	-4.16	48.29	.000**	-2.84	39.78	.007**
Experience violence –verbal												
v No of factors												
Scenario One	3.24	125.29	.002**		ns							
Scenario Two	3.17	135.36	.002**		ns							
Experience violence- physical												
v No of factors												
Scenario One	2.33	80.65	.022*		ns							
Scenario Two	2.16	66.44	.034*		ns							

Equal variances not assumed in above tests *significant at the 0.05 level (2-tailed) **significant at the 0.01 level

than students who did not report experiencing physical violence during Trimester One but these differences were only statistically significant at t2 for Scenario 2 [t 2.316, df 74.786, p=.023]. Finally, destined branch was not shown to be significant with regard to any variable in this section.

**Later in this Chapter the data relating to number of risk factors will be analysed further at a more detailed level in order to explore the more subtle changes in the number of risk factors identified as a result of attending the Unit.

5.5 24 LIKERT STATEMENTS

It will be recalled that there were 24 Likert-type statements on the final page of the Questionnaire, each with a five point scale, scored as follows - 1 strongly agree, 2 agree, 3 don't know, 4 disagree, 5 strongly disagree. The statements were written to allow a determination of changes in attitudes, confidence, self-assessed competence, personal value etc. as originally intended in the Unit's learning objectives and outcomes. In most instances, several statements were written for each learning outcome and Box 4.5 identified the statements associated with specific learning outcomes.

It is much more preferable to analyse the combined scores for a group of statements rather than analyse statements individually since they are

then more stable and less susceptible to individual variance. Box 4.5 in Chapter 4 provides the theoretical groupings for statements as they were originally written but the performance of the scale had to be determined in practice. The intention in this case was to perform a factor analysis on the Likert statements in order to determine (a) whether the 24 statements could be statistically reduced to a satisfactory smaller number of representative factors and (b) whether these factors bore any relation to those originally envisaged, i.e. the learning outcomes. Once factors were identified then factor scores at the four time points could be calculated and changes in scores noted and analysed for statistical significance.

5.5.1 Initial Review

A preliminary task, however, was to examine the ability of statements to discriminate between respondents, that is, the adequacy of the spread of responses for each statement at each of the four time points. Streiner (1993) suggests that attention be given to the 'endorsement frequency' and 'restriction in range' of scale items, adding that "if some items are answered in one direction or another more than 90% or 95% of the time they may be worse than useless" (Streiner 1993:142) and exclusion should be considered. Appendix Four shows a summary of the percentage of responses for each statement at each of the four time points presented as a 100% stacked chart. This simple rule is complicated by a number of other considerations. For example, the Streiner article refers to applying the rule to a scale measuring a single

dimension rather than an inventory that measures a number of different variables [as in this case].

The study is an evaluation of training effects over time and there will [hopefully] be dramatic changes in scores between certain time points, therefore the distortion referred to above would be more problematic if found in the responses obtained prior to completing the Unit. Visual analysis of the 100% stacked charts and simple descriptives [frequencies] tables showed that a small number of statements transgressed these rules in the period prior to completing the Unit. Specifically,

- Statement 3 [It is unacceptable for nurses to protect themselves when being physically assaulted by a patient] was scored as 'strongly disagree' by 52% at t2 and, in total, 91% of respondents disagreed or strongly disagreed with this statement at that time.
- Statement 5 [When staff members are assaulted and have no injuries, there is no need to report the incident] was scored 'strongly disagree' by over 71% of respondents on t1 and t2 prior to the Unit and, indeed, over 98% of respondents either disagreed or strongly disagreed with this item at t1 and t2.
- Statement 7 [The members of staff who are physically assaulted are generally those who are least competent in their job] was scored as 'strongly disagree' by 60% at t1 and, in total, over 93% of respondents disagreed or strongly disagreed with this statement at time points t1 and t2.

- Finally, Statement 12 [Staff should be educated about the prevention and management of aggressive behaviour as part of their pre-registration preparation] was scored as 'strongly agree' by over 76% of respondents at t1 and t2 and, indeed, 100% of respondents agreed or strongly agreed with this statement on commencing the unit at t2.

These statements had fairly high endorsement frequency for one response and a restricted range of responses used [i.e. over 90% of respondents answering in one direction] and would, on this basis, be candidates for removal. However each of the questions was the sole question relating to that particular area, staff right to defend themselves, importance of reporting incidents, incompetent staff being victims, and importance of staff preparation and training. The later process of determining factors could very well result in the removal of variables and, for this reason, the statements identified above were left in the analysis at this stage.

5.5.2 Factor Analysis

Several authors (Royce 1958, Child 1990) report a long developmental history for the techniques of factor analysis, citing the contributions of Sir Francis Galton, Karl Pearson and Charles Spearman. Comrey (1978) asserted the methods have increased in popularity with the availability of computerised statistical packages, and, in the interests of

competent reporting, influentially identified potential pitfalls and ways of overcoming them.

Streiner (1994:136) suggests that "the purpose of factor analysis is to determine if a small number of underlying factors can explain the pattern of scores obtained on a battery of tests". Factors are hypothetical constructs and Bryman & Cramer (1997) suggest that the techniques of factor analysis have three main purposes, these being to determine the degree to which variables in an assessment instrument are tapping the same construct, to reduce a large number of variables to a smaller set for statistical advantage, and to aid understanding of social behaviour by reducing its complexity.

There are several factor analysis techniques, and one of the most popular is called Principal Component Analysis, the one chosen for use in this study. It is described as "the solution of choice for the researcher who is interested in reducing a large number of variables down to a smaller number of components" (Tabachnick & Fidell 1989:626).

Principal Component Analysis (PCA) can produce independent [orthogonal] components that are then able to be used in parametric statistical tests.

The techniques hinge on statistically analysing the levels of correlation between variables [hopefully, all variables on a scale correlate well with some other variables but not with others] and "describing the variation

or variance within the scores of respondents on three or more variables (Bryman & Cramer 1997). Initially in an analysis there are as many factors as there are variables (Streiner 1994) and the process needs rules and procedures to be applied at different stages in order to reduce the number of factors by removing ineffectual ones, standardise analysis, and increase statistical rigour. Unfortunately, some of these rules and protocols still lack general widespread agreement.

Gorsuch (1978) offered an algorithm to formalise an analysis. Ferguson & Cox (1993) describe three stages to the process of exploratory factor analysis: pre-analysis checks, extraction and rotation and offer heuristics for each stage that updated earlier advice offered by Comrey (1978). Tabachnick & Fidell (1989:598) identify key steps in factor analyses as “selecting and measuring a set of variables, preparing the correlation matrix, extracting a set of factors from the correlation matrix, determining the number of factors, (probably) rotating the factors to increase interpretability, and , finally interpreting the results”.

Interpretability and scientific utility seem to be key since “after extraction [of the factors] there are an infinite number of rotations available, all accounting for the same degree of variance in the original data, but with factors defined slightly differently” (Tabachnick & Fidell (1989:598). Contentious issues include those regarding the minimum sample size and number of respondents, the ratio of variables to respondents, level of measurement used [5-point ordinal is acceptable (Ferguson & Cox

1993), demonstration of univariate normality, the determination of the number of factors to be retained and rules for extraction, rules for accepting variables within factors, the management of cross-loading variables and factorial complexity [Streiner (1994) suggests three options - these are including variables on all factors , inclusion only on highest loading factor or remove from analysis], rules for the type of rotation used to maximise the loading of variables on only one factor, and rules for calculating factor scores.

Reassuringly, selection from the enormously large number of possible extraction and rotation techniques is probably not critical since Tabachnick & Fidell (1989) suggest that in practice any differences among them are minor. These issues and statistical terms will not be explored further but it is proposed to provide a detailed protocol of the particular rules used in the analysis of data obtained in this study.

5.5.3 Protocol Used for Principal Component Analysis

- I. Data from different cohorts of students and different occasions was reframed. Responses were stacked to create a separate database containing 24 variables and 1016 respondents – the Factor Database – (Leather *et al* 1994, McKeachie *et al* 1997).
- II. Variables were checked for univariate normality and acceptable levels of skewness and kurtosis.
- III. Checks were performed for adequate numbers and levels of correlations between some variables via Bartlett's Test of

Sphericity and Kaiser-Meyer-Olkin Measure of Sampling Adequacy

- IV. Correlation, Partial Correlation and Anti-Image Correlation matrices were examined. All variables should correlate with at least one other at $r > 0.2$ and Anti-image Correlations should be > 0.5 (Field 2000).
- V. Principal Components Analysis with Orthogonal Varimax Rotation was performed to extract factors.
- VI. Number of factors to be retained was determined by studying Eigenvalues and discarding factors with Eigenvalues < 1.0 [The Kaiser 1 or K1 criterion which is the default on most computer programmes].
- VII. Factor Matrix was studied and variables with communality less than 0.3 and variables that cross-load on more than one factor at > 0.4 were removed.
- VIII. Principal Components Analysis was repeated [back to point 3] until no Communality < 0.3 or cross-loading variables.
- IX. Results were interpreted and factors named.
- X. Variables with negative factor loadings were re-scored.
- XI. Factors scores were calculated.
- XII. Factor Scores were subjected to parametric statistical analysis [matched pair t-tests].

Tabachnick & Fidell (1989:602) urge wariness against combining data from different samples or from the same sample on different occasions

unless the samples produce the same factors, in which case, "pooling them is desirable because of increase in sample size". In this study, given cohorts of approximately 70-80, some combination was necessary in order to satisfy minimum sample size and ratios of variables to sample size guidelines. Ferguson & Cox (1993)] suggest at least 100 in a sample, while Field (2000) cites several authors who quote the need for 300 respondents. Moreover, cited desirable ratios of respondents to variables [between 1-4 and 1-10 are suggested by Streiner (1994)] means that at least 125-240 respondents are required for the 24 statements/variables in this part of the questionnaire. In addition, earlier in this chapter, different cohorts have been shown to be similar on key demographic variables.

On several occasions during the data collection period, after a batch of questionnaires had been collated and entered in to the factor database, a factor analysis was performed. However, these time-points did not necessarily neatly coincide with the summary material for a particular cohort. For example, a factor analysis was first performed on 12th January 2001 on assorted forms from Cohorts One and Two and then repeated on 26th March 2001 when the final forms related to Cohort Two and initial forms from Cohort Three were collated and computed. The factor analysis was repeated again on 19th September 2001, 15th November 2001 and 3rd December 2001 as further data from Cohort Three were obtained and inputted into the database. The factor structures obtained were recognisably similar on each occasion.

Indeed, several factors were present and identically constituted on several occasions. Hence, it was thought permissible to combine all responses to the 24 Likert statements in a new database.

5.5.4 Principal Components Analysis Solution For 24 Likert Statements Combined For All Cohorts On All Occasions.

The reformed factor database consisted of 1016 respondents.

Measures of skewness and kurtosis for the 24 statements revealed all statements to be within limits except statement 12 [skewness 2.483, kurtosis 11.382]. Ferguson & Cox (1993) offer a complicated heuristic for determining whether it is possible to retain items affected by skew and kurtosis depending on the number and proportion of initial correlations. However, in this case, because of the large kurtosis score, and the fact that it had also performed badly on assessment of endorsement frequency for one response and a restricted range of responses [see Section 5.5.1] this statement was removed from the analysis. [It will be recalled that this statement asserted " Staff should be educated about the prevention and management of aggressive behaviour as part of their pre-registration preparation" and received 100% agreement on Day One of the module].

The optimum factor solution was obtained after the removal of four variables and repetition of the analysis three times. Statements 2 was removed because of a communality of <0.3 , while statements 18, 20 and 7 were removed because they cross-loaded on several Factors at

>0.4 . The analysis was performed with Varimax rotation and Kaiser Normalisation and converged after 5 iterations. The final score for KMO Measure of Sampling Adequacy was 0.820 and for Bartlett's Test of Sphericity Approximate Chi-square 3700.300, d.f. 171, significance 0.000. These scores indicate that there are sufficient correlations between variables to perform a factor analysis and that the sample is adequate [KMO >0.8 is 'very good' Field (2000)]. Table 5.14 shows details of Communalities and reveals that extracted communalities for Statements 10 & 19 are close to the cut off [<0.300].

	Initial	Extraction
Question 1	1.000	0.514
Question 3	1.000	0.456
Question 4	1.000	0.553
Question 5	1.000	0.374
Question 6	1.000	0.525
Question 8	1.000	0.603
Question 9	1.000	0.632
Question 10	1.000	0.300
Question 11	1.000	0.483
Question 13	1.000	0.599
Question 14	1.000	0.471
Question 15	1.000	0.362
Question 16	1.000	0.764
Question 17	1.000	0.623
Question 19	1.000	0.312
Question 21	1.000	0.500
Question 22	1.000	0.802
Question 23	1.000	0.826
Question 24	1.000	0.403

Table 5.15 shows the Eigenvalues for the Factors and cumulative percentage of variance accounted for by the rotated factors that remain in the analysis. It reveals, after Variamax Rotation, a five-factor solution that explained 53.166% of the variance.

Factor	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative%	Total	% of Variance	Cumulative %
1	4.166	21.928	21.928	3.446	18.138	18.138
2	2.082	10.959	32.887	1.906	10.034	28.171
3	1.385	7.289	40.177	1.746	9.190	37.171
4	1.336	7.031	47.207	1.518	7.989	45.351
5	1.132	5.960	53.166	1.485	7.816	53.166
6	0.930	4.894	58.060			
7	0.914	4.809	62.869			
8	0.866	4.556	67.425			
9	0.838	4.412	71.837			
10	0.751	3.951	75.789			
11	0.734	3.861	79.650			
12	0.690	3.633	83.282			
13	0.645	3.393	86.676			
14	0.606	3.191	89.867			
15	0.541	2.850	92.717			
16	0.510	2.683	95.399			
17	0.445	2.341	97.740			
18	0.265	1.394	99.134			
19	0.165	0.866	100.00			

Figure 5.2 shows the Scree Plot that lead the researcher to select the same five factor solution.

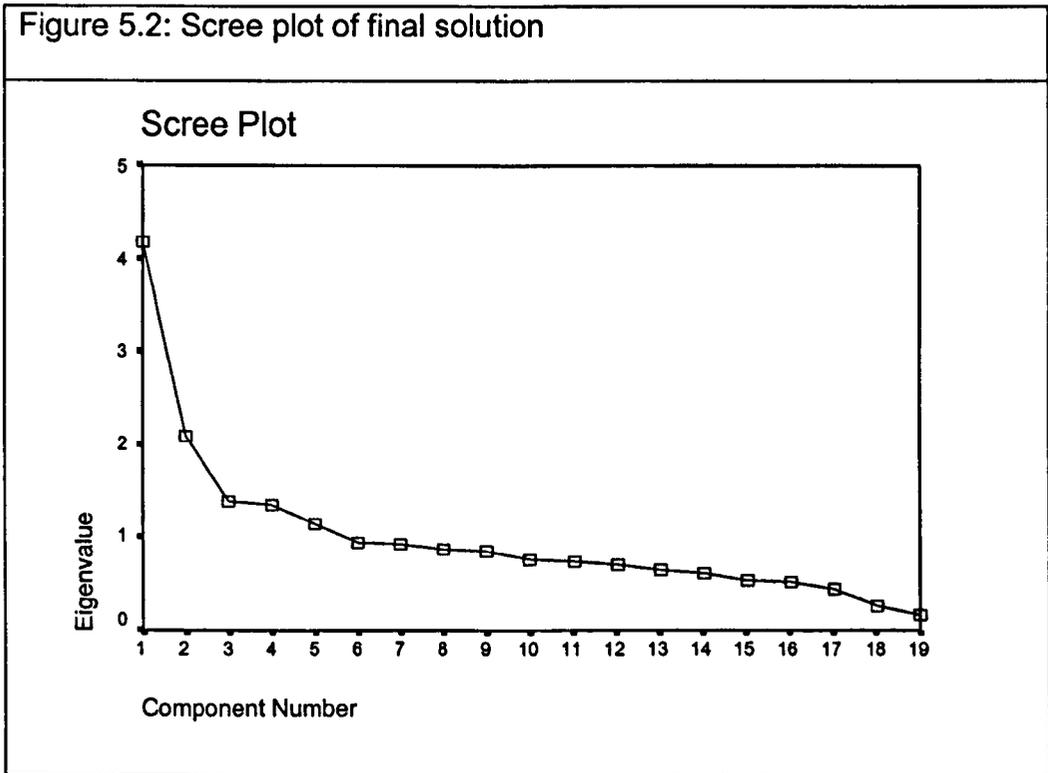


Table 5.16 shows the Rotated Component Matrix that reveals the loading of each statement associated with a particular Factor [factor loading coefficients]. It reveals that Factors 1-4 each have four variables associated with them while Factor 5 has two variables. Statement 19 was not associated with any factor at >0.4 and so was excluded from further analysis.

Factor 3 has two variables negatively scored and Factor 4 has one variable negatively scored. Student scores associated with these three statements were reversed before any calculation of Factor Scores was made.

Table 5.16: Rotated component matrix					
	Factor/ Component				
	1	2	3	4	5
Statement 23	0.894				
Statement 22	0.889				
Statement 16	0.862				
Statement 17	0.700				
Statement 19					
Statement 6		0.719			
Statement 4		0.716			
Statement 11		0.660			
Statement 10		0.524			
Statement 13			-0.759		
Statement 21			0.586		
Statement 14			0.518		
Statement 15			-0.505		
Statement 3				0.647	
Statement 5				0.566	
Statement 1				0.557	
Statement 24				-0.544	
Statement 9					0.779
Statement 8					0.747

Next, each Factor was reviewed and named according to the sense that the statements associated with it portrayed. To aid clarity, Table 5.17 shows the full text of each statement grouped under its factor name. A brief overview of the essence of statements grouped under each factor and allotted factor names follows:-

- Factor 1 consisted of four statements that included phrases to do with being confident in one's ability to remain safe at work and use breakaway skills and legal levels of self-defence. Hence it was entitled '**Confidence in Maintaining Personal Safety**'.
- Factor 2 again consisted of four statements. Two items talked about the predictability of aggression, one suggested that most aggression could be prevented and one addressed the issue of estimating patient responsibility for aggression. This factor was entitled '**Prediction and Prevention**'.
- Factor 3 also had four statements. All four statements referred to clinical experience and the respondent's own ability to interact and remain calm in the presence of agitated or aggressive patients. One question, in addition, also suggested that the student nurse's role was to remain in the background and be available to offer assistance. This statement received a negative loading, as did a negatively phrased one about become 'nervous' in the presence of aggressive patients. Factor 3 was entitled '**Personal Practical Ability**'.
- Factor 4 contained four statements. Three of these statements illustrated examples of low self-respect or disregard for staff rights and consequently suggested poor practice responses in response to staff assault. The fourth statement was scored in the opposite direction and suggested that staff had rights, including the right to take legal action against a patient who assaults them. This factor was named '**Self-Respect and Staff Rights**'.

- Factor 5 consisted of two statements. These statements suggested extremes of behaviour –authoritarian or submissive approaches - which were likely to be perceived by a patient as provocative and, hence, more likely to lead to assault. The factor was named '**Provocative Approach**'.

Table 5.17: Questionnaire statements associated with each factor name	
Factor 1 – Confidence in maintaining personal safety	Q23. I am confident of my ability to protect myself using legally permissible force if attacked by a patient or visitor Q22. I am fully aware of the legal framework within which self defence is legitimated Q16. I can describe the main principles of breakaway techniques Q17. I am confident of my ability to remain safe at work
Factor 2 – Prediction and prevention	Q6. Prediction of patient assault is within the competence or ability of practicing psychiatric nursing staff Q4. Most aggression and violence by patients is predictable Q11. Most aggression and violence by patients is preventable Q10. We under-estimate how much people with mental illness or learning disabilities are responsible for their behaviour
Factor 3 – Personal practical ability	Q13. When a patient becomes increasingly aggressive I get so nervous that I can hardly think straight Q21. I am able to talk in a calming and reassuring way to a verbally aggressive patient/relative and manage the situation Q14. I am able to demonstrate a non-provocative approach towards, and stance in front of, a patient Q15. The student nurse's role in a violent incident is to remain in the background and let qualified staff deal with the situation
Factor 4 – Self-respect and staff rights	Q3. It is unacceptable for nurses to protect themselves when being physically assaulted by a patient Q5. When staff members are assaulted and have no injuries, there is no need to report the assault Q1. Health professionals should accept that being assaulted is an inevitable part of the job Q24. Staff have a right to take legal action against patients who have assaulted them
Factor 5 – Provocative approach	Q9. Staff with a yielding and submissive manner are more likely to be assaulted Q8. Staff with an authoritarian manner are more likely to be assaulted

5.5.5 Factor Scores

Firstly, the scores of the three statements which received negative loadings [statements 13, 15, and 24] were reversed so that 'strongly agree' scoring 1 became 'strongly disagree' scoring 5, and 'agree' scoring 2 became 'disagree' scoring 4. Factor scores were then calculated by simply totalling the individual statement response scores [the 5 point Likert scale scores] for the statements related to each factor. No sign of skewness or kurtosis within the factor scores meant they were considered normally distributed and suitable for parametric analysis.

Matched pair t-tests were then performed on these results to determine change in factor scores between different time points. Specifically differences between t1 and t2, t2 and t3, and t2 and t4 were compared. These pairs of time points permit evaluation of the stability of the questionnaire prior to the Unit [test-retest reliability], the immediate effects of the Unit and the on-going effects of the Unit and subsequent clinical placements to be made. Table 5.18 shows the results of this analysis and provides mean score details [with grey arrows showing anticipated direction of change], along with t-test scores.

5.5.5.1 Factor 1 [Confidence In Maintaining Personal Safety]

It was anticipated that, following attendance of the Unit, student nurses would increasingly agree with these statements, resulting in a reduction in factor scores [strongly agree=1, agree=2 etc.] The scores for Factor

1 reveal no significant change prior to the Unit but a statistically significant reduction immediately following completion of the Unit that was maintained at the 3-month follow-up. The mean score reduced to a great extent from 13.67 on Day One of the Unit to 8.02 on Day Three and remained below 8.8 at the follow-up.

5.5.5.2 Factor 2 [Prediction and Prevention]

A similar pattern was anticipated for scores related to Factor 2. Following material on risk factors, models, and risk assessment, increasing agreement with statements about violence being predictable and preventable would result in lower factor scores. In this case the reduction in mean score was more modest, from 11.77 on Day One to 10.47 on Day Three. The matched pair t-test scores showed no significant change prior to the Unit but a statistically significant change immediately following it that was maintained at the 3-month follow-up.

5.5.5.3 Factor 3 [Personal Practical Ability]

The interpretation of scores is much more complicated in the case of Factor 3. Following re-scoring of statements 13 and 15, it was anticipated that a *lower* score for statements 13, 14, and 21 would be associated with a positive Unit effect as students perceived themselves as more competent in their interactions and ability to remain calm.

However, following re-scoring, a *higher* score for statement 15 was anticipated post-Unit, indicating that the student had taken on board the

Unit's message about the student staying in the background, offering assistance, and allowing regular, qualified staff to deal with incidents. Hence there was a bi-directional tension within the statements and this affected the factor score. The factor scores revealed a statistically significant increase prior to the Unit, a statistically significant decrease immediately on completion and at follow-up compared with Day One of the Unit, and a final score that was lower than the initial score.

If Statement 15 was removed from the calculation, because of its opposing effect, and the factor scores re-calculated for statements 13 [reversed], 21 and 14 then a slightly better picture emerges, although still far from ideal. In this case, there is still instability evident in scores prior to attending the Unit with a statistical significant increase in factor scores, [but less so than previously, $t = -2.287$, $df = 132$, $p = .024$].

Importantly however, following the Unit factor scores fall below *either* pre-Unit score and remain lower at the three-month follow-up at statistically significant levels, for t_2-t_3 [$t = 6.614$, $df = 206$, $p = .000$] and for t_2-t_4 [$t = 10.610$, $df = 193$, $p = .000$]. Mean factor scores for these three statements were 7.13, 7.61, 6.80, 6.36 at times t_1 , t_2 , t_3 , and t_4 respectively. Hence there is again a decrease in modified factor scores following attendance on the Unit, as anticipated.

A separate analysis of Statement 15 also revealed interesting results. If the original student responses to Statement 15 are considered [prior to re-scoring] then *lower* scores are anticipated following attendance on

the Unit if the student nurse agrees with the original statement. It is found that the mean score for statement 15 shows a statistically significant reduction between t1-t2 prior to the unit, from 3.11 to 2.84 [t 3.354, df 137, p= .001]. As anticipated, the mean score shows a further statistically significant reduction t2-t3 immediately following the Unit, from 2.84 to 2.67 [t 2.075, df 210, p= .039].

However there is an unanticipated increase in score to 2.83 at t4, following the clinical placements, [t2-t4 t -1.831, df 192, p= .069] that returns the mean score almost back to its position prior to students completing the Unit. It appears that the scores to Statement 15 are adversely affected by the realities of clinical practice, wherein it may not be possible for students to 'take a place in the relative background'. Hence, singularly both Statement 15 and the other three statements that make up a modified Factor 3 appear to respond as anticipated appears to respond as anticipated to the Unit but in a combination that forms Factor 3 they appear to be opposed following the requisite re-scoring indicated by the factor analysis process.

5.5.5.4 Factor 4 [Self Respect and Staff Rights]

Following the re-scoring of Statement 24 scores perusal of Factor 4 suggested that the factor score should *increase* following attendance on the Unit as students should be more inclined to disagree with the statements. Factor scores were not stable prior to the Unit showing a statistically significant change, although in the anticipated direction, with

mean scores increasing t1 to t2, from 16.89 to 17.30. There was a further statistically significant increase on immediate completion of the Unit, maintained at the three-month follow-up, although mean scores reduced slightly between t3 and t4 following the two clinical placements.

5.5.5.5 Factor 5 [Provocative Behaviour]

It was anticipated that factor scores would *decrease* following attendance on the Unit, as students should be more likely to agree with these statements. Factor scores showed no statistically significant change in the period prior to the Unit but, as expected, a statistically significant decrease in score occurred on completion of the Unit and was maintained at the three-month follow-up [compared to Day One]. Mean scores at all times after completing the Unit were lower than those prior to attendance.

Table 5.18: Factor scores identified at four time points – mean scores and paired sample t-test results

[Mean scores]					
		pre unit [t1]	day 1 [t2]	day 3 [t3]	follow-up [t4]
Factor 1	↓	13.74	13.67	8.02	8.78
Factor 2	↓	11.82	11.77	10.47	10.56
Factor 3	↓	9.99	10.77	10.12	9.51
Factor 4	↑	16.89	17.30	17.79	17.72
Factor 5	↓	7.01	7.27	6.61	6.84
Factor Scores - Paired sample t-test					
Factor 1: Confidence in maintaining personal safety			t	d.f.	Significance (2-tailed)
Pair 1:	pre unit – day 1		0.810	132	0.419
Pair 2:	day 1 – day 3		25.458	208	0.000
Pair 3:	day 1 – follow-up		22.066	194	0.000
Factor 2 : Prediction and prevention			t	d.f.	Significance (2-tailed)
Pair 1:	pre unit – day 1		-0.464	129	0.643
Pair 2:	day 1 – day 3		7.691	201	0.000
Pair 3:	day 1 – follow-up		7.313	187	0.000
Factor 3: Practical ability			t	d.f.	Significance (2-tailed)
Pair 1:	pre unit – day 1		-3.219	130	0.002
Pair 2:	day 1 – day 3		4.651	204	0.000
Pair 3:	day 1 – follow-up		9.794	191	0.000
Factor 4: Self-respect and staff rights			t	d.f.	Significance (2-tailed)
Pair 1:	pre unit – day 1		-2.293	138	0.023
Pair 2:	day 1 – day 3		-3.711	211	0.000
Pair 3:	day 1 – follow-up		-3.593	196	0.000
Factor 5: Provocative approach			t	d.f.	Significance (2-tailed)
Pair 1:	pre unit – day 1		-1.583	140	0.116
Pair 2:	day 1 – day 3		5.730	212	0.000
Pair 3:	day 1 – follow-up		4.192	197	0.000

5.5.6 Graphic Representation

For many, charts or graphs are more meaningful and easily discernible.

It was felt that the mean factor scores changes in Table 5.18 might more clearly illustrate training effects if presented as a set of bar charts.

The idealised shapes for such a series are illustrated in Figure 5.3 and consist of a horizontal line on the occasions prior to training, followed by

a plateau step up [or down] immediately following training representing the effect of training, and a second horizontal line from training to

follow-up representing the maintenance of change and absence of undermining clinical effects. Figure 5.4 illustrates the mean data

relating to Factor 1-5 scores at times t1, t2, t3, and t4, and possibly shows more clearly the effects described in the preceding section

Figure 5.3: Idealised bar charts for effects of training

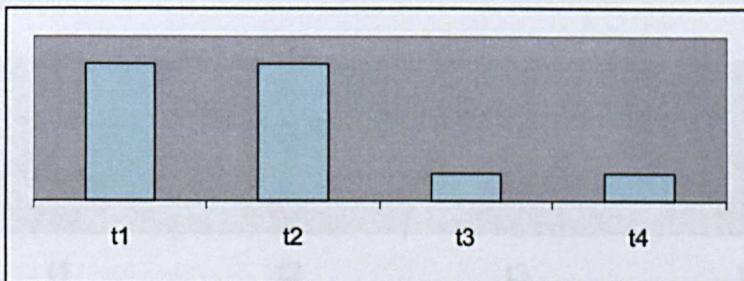
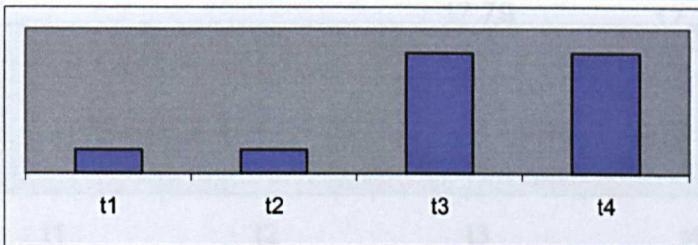
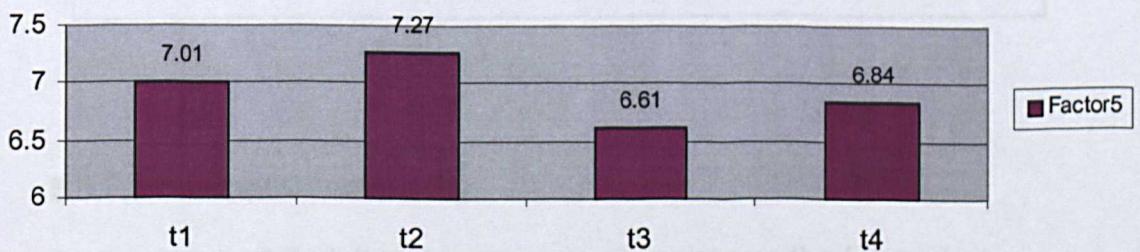
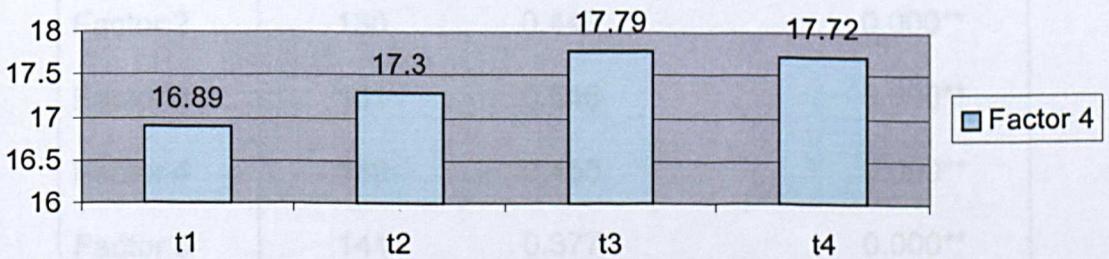
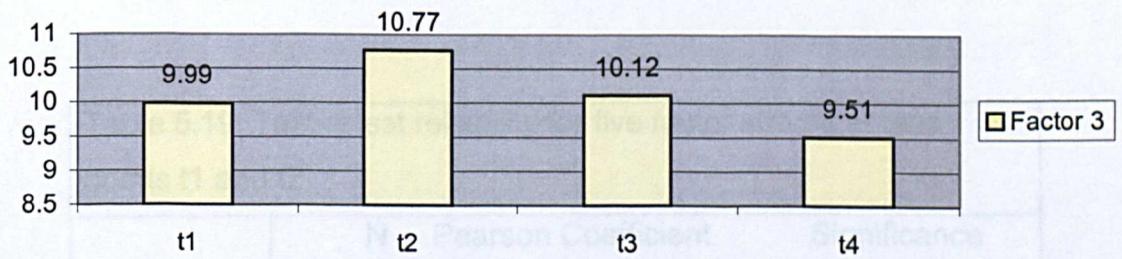
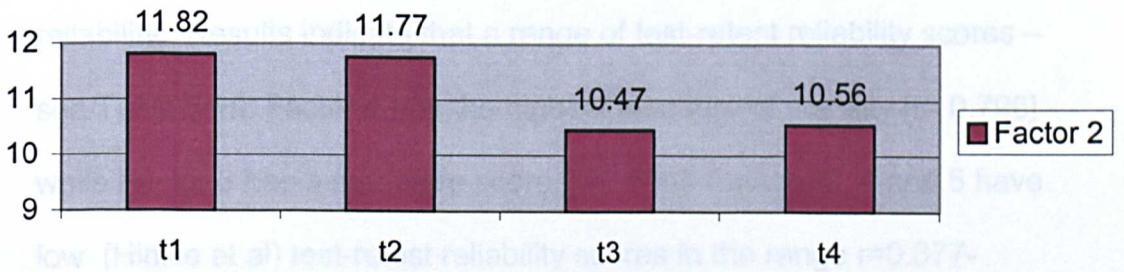
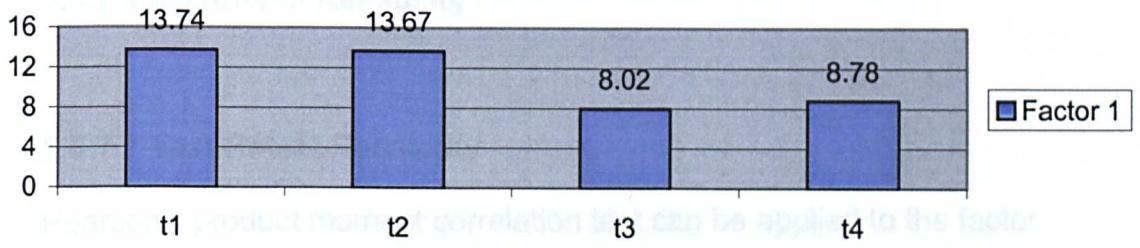


Figure 5.4: Bar charts of mean scores for factors 1-5 at time points t1, t2, t3, & t4



5.5.7 Measures of Reliability

5.5.7.1 Test-Retest Reliability

Pearson's product moment correlation test can be applied to the factor scores obtained for time points t1 and t2 as a measure of test-retest reliability. Results indicate that a range of test-retest reliability scores – see Table 5.19. Factor 1 has the highest measure of stability [$r= 0.796$] while Factor 3 has a moderate score [$r=0.546$]. Factors 2, 4 and 5 have low (Hinkle et al) test-retest reliability scores in the range $r=0.377-0.455$.

	N	Pearson Coefficient	Significance
Factor 1	133	0.796	0.000**
Factor 2	130	0.440	0.000**
Factor 3	131	0.546	0.000**
Factor 4	139	0.455	0.000**
Factor 5	141	0.377	0.000**

** Correlation significant at 0.01 level (2 tailed)

5.5.7.2 Internal Consistency

As previously stated, it makes no sense to measure the internal consistency of the 24 Likert statements prior to factor analysis since that part of the questionnaire was designed to measure different

aspects of change. Some statements are intentionally unrelated [uncorrelated] to others. However, the test is advocated following a factor analysis procedure and calculation of factor scores (Watson 1995). It is possible to calculate a Cronbach Alpha α coefficient for each occasion that the questionnaire was administered. This has to be done retrospectively once the factor analysis procedure has indicated the items [or statements] that constitute each Factor.

In this case such a procedure would yield over 20 such coefficients and prove difficult to present. However, it is feasible to calculate Cronbach Alpha α on the combined cohort data for each time point and the results are presented in Table 5.20, along with a mean Cronbach Alpha score for the four occasions. It will be seen that only Factor 1 satisfies Nunnally's (1978) standard of reliability coefficients being > 0.7 before a scale can be used as a research tool. Factors 2 and 5 approach Gibbon's (1995) suggests value of >0.6 .

No Factor contains more than four items, and, given that the number of items in the scale affects the value of Cronbach alpha obtained [few items equating with lower scores] it would be possible to increase the Internal Consistency of the questionnaire by increasing the number of statements in that part of the data collection tool, should one wish to persist with it.

Table 5.20: Cronbach alpha α for factor scores calculated at time points t1-t4 and mean alpha score					
	t1	t2	t3	t4	Mean alpha [t1-t4]
Factor One Items 23, 22,16, 17	N=145 0.8389	N=225 0.8345	N=222 0.7645	N=207 0.7337	0.7929
Factor Two Items 6, 4, 11, 10	N=142 0.4365	N=223 0.6282	N=215 0.6060	N=203 0.5647	0.5589
Factor Three Items13 [rescored], 21, 14, 15 [rescored]	N=146 0.4620	N=222 0.5653	N=220 0.4869	N=203 0.5376	0.5130
Factor Four Items 3, 5, 1, 24 [rescored]	N=151 0.3504	N=226 0.4316	N=223 0.3435	N=208 0.5024	0.4070
Factor Five Items 9, 8	N=151 0.4862	N=228 0.5833	N=222 0.5265	N=208 0.5778	0.5435

However, as mentioned earlier, Leather *et al* (1998) suggest calculation of the mean inter-item correlation is a more fitting estimation of the reliability of small scales. In this case, given the small number of items associated with each Factor, it is appropriate to also perform this calculation. Acceptable values of this mean inter-item correlation lie in the range of 0.1 –0.5 (Cox & Ferguson 1994). Table 5.21 shows the calculation of this value for each of the five Factors for the combined cohorts at time points t1-t4.

It can be seen that all Factors, at all time points, now have Mean Inter-Item Correlation scores that fall with the range suggested by Cox & Ferguson (1994). Factor 1 continues to have the highest score, with its mean average virtually at the upper limit of the indicated range. Hence on this measure all Factors are considered to have acceptable internal consistency scores.

Table 5.21: Mean inter-item correlation for factor scores calculated at time points t1-t4					
	t1	t2	t3	t4	Mean Inter-Item Correlation averaged for t1-t4
Factor 1 Items 23, 22, 16, 17	N=145 0.5612	N=225 0.5497	N=222 0.4460	N=207 0.4205	0.4944
Factor 2 Items 6, 4, 11, 10	N=142 0.1660	N=223 0.2952	N=215 0.2775	N=203 0.2339	0.2432
Factor 3 Items 13 [rescored], 21, 14, 15 [rescored]	N=146 0.1870	N=222 0.2634	N=220 0.2205	N=203 0.3656	0.2591
Factor 4 Items 3, 5, 1, 24 [rescored]	N=151 0.1301	N=226 0.1653	N=223 0.1176	N=208 0.2053	0.1546
Factor 5 Items 9, 8	N=151 0.3214	N=228 0.4179	N=222 0.3628	N=208 0.4103	0.3781

Given the earlier interest in constitution of Factor 3 [see Section 5.5.5.3] and the apparent opposing effect of a re-scored Statement 15, it was considered appropriate to examine further the internal consistency of this Factor. When performing the Cronbach Alpha test within SPSS it is possible to request an estimation of the Cronbach Alpha if any item was removed. Table 5.22 shows the results of this test for combined cohorts at time points t1-t4. It indicates that for Factor 3, at all time points, the Cronbach Alpha coefficient α is reduced if Item [Statement] 13, 21, or 14 is removed. It also indicates that, on every occasion, the Cronbach Alpha coefficient α for Factor 3 is *increased* if Item 15 [rescored] is removed.

Table 5.22: Calculation of cronbach α for factor 3 if item removed at time points t1-t4				
Factor 3	Cronbach α if item removed			
	t1	t2	t3	t4
Original α score	0.4620	0.5653	0.4869	0.5376
Removed Item				
Item 13 rescored	0.2822	0.4010	0.2793	0.4197
Item 21	0.3965	0.4443	0.3700	0.4020
Item 14	0.3682	0.4894	0.4422	0.4424
Item 15 rescored	0.5134	0.6298	0.5740	0.6276

Removal of Item 15 and re-calculation of Cronbach Alpha for Factor 3 [now consisting of Items 13 [rescored], 21 and 14] at time points t1-t4 produced coefficients for α which were slightly different [higher or lower] than those estimated above. Nevertheless, on each occasion they were higher than the original calculated value of α shown in Table 5.20. In addition the newly calculated Mean Inter-Item Correlation was higher at time points t1-t3 [and the same at t4] than those originally calculated and presented in Table 5.21. These results are presented in Table 5.23.

Table 5.23: Effects on cronbach alpha and mean inter-item correlation of removing item 15 from factor 3 at time points t1-t4				
Factor 3: Items 13 [rescored], 21, 14	Re-calculated Cronbach Alpha α if Item 15 removed at Time Points t1-t4			
	t1	t2	t3	t4
	N=148	N=222	N=222	N=205
	$\alpha = 0.4878$	$\alpha=0.6312$	$\alpha =0.5686$	$\alpha =0.6164$
	Mean Inter-Item Correlation if Item 15 removed at Time Points t1-t4			
t1	t2	t3	t4	
0.2406	0.3742	0.3124	0.3656	

These statistical tests present a dilemma for the researcher. It appears that the Internal Consistency of the statement part of the questionnaire would be improved by removing statement 15 from the constitution of Factor 3. At the same time, statement 15 is a key statement and the

only one that refers specifically to Hypothesis 6 [about the role of the student nurse in the management of violent incidents].

Watson (1995:59) considers this issue and suggests, "in addition to analysing these statistical properties, the substantive properties of the instrument should also be considered". In so doing, it may be determined that items that do not "contribute to the construct validity of the instrument may be retained because they are considered to be important for other reasons. Such reasons may be practical rather than statistical" (Watson 1995:59).

In this case it was determined that there was a good deal to be lost by removing statement 15 from the analysis, in terms of its direct and exclusive relationship to Hypothesis 6. In addition, there was essentially nothing to be gained from its elimination since the questionnaire is not being developed for more widespread utilisation. Indeed, it should be remembered that the questionnaire was designed to reflect the particular learning outcomes pertaining to the pre-existing aggression management Unit.

5.5.8 Bi-variate Analysis

Possible differences between factor scores related to demographic variables were analysed using independent samples t-tests. Variables tested were gender, age [less than or equal to/greater than 26 years], and previous training. Destined branch was compared using one-way

ANOVA . Table 5.24 shows instances where significant results were returned. There were no relationships with regard to destined branch at any time point. Age makes a difference to scores for Factor 2 [Prediction and prevention] with students aged 26 years or over having lower mean scores at all time points, a result statistically significant at t1, t2, and t4. These findings suggest that older student nurses consider violence to be more predictable and preventable than younger student nurses.

Gender created a difference on scores for Factor 1[Confidence in maintaining personal safety] with males having lower mean scores than females at all time points although this was only statistically significant at t3. Males also had a lower mean score at all time points on Factor 3 [Personal practical ability] with differences being statistically significant at all time points. These findings suggest that male student nurses feel more confident in their own ability to remain safe and protect themselves at work than their female counterparts. The findings also suggest that male student nurses feel more competent in their ability to interact with and calm agitated or aggressive patients and manage the situation.

All the results for previous training showed similar effects, irrespective of the type of training that had been received –theory, breakaway training or physical restraint. Students with previous training of any sort had a lower mean score on Factors 1 and 3, at all time points They also

recorded a higher mean score on Factor 4 at t1, t2 and t4 [also at t3 for student with restraint training]. These differences in factor scores were statistically significant at t1, t2 and t4 but not at t3.

These findings suggest that previous training of any sort increases the student nurses confidence in their personal ability to remain safe. It also increases their self-assessed competence in managing aggressive behaviour calmly, and also gave them greater self-respect and belief in their rights to work safely than colleagues who had not received previous training.

However, put another way, the finding that previous training of any type made no statistical significant difference in factor scores for any factor on immediate completion of the module [t3] suggests that, on completion of the Unit, students without prior training performed similarly [in terms of confidence in ability to remain safe, self assessed competences etc.] to those who had undertaken additional prior training, possibly amounting to five, even ten days in duration.

Interestingly, findings did not reveal that previous training made any difference to student nurses' belief that violence was predictable and preventable [Factor 2].

Table 5.24: Selected results for bivariate analysis of demographic variables and factor scores

	t1		t2		t3		t4		
	t	df	sig (2-tailed)	t	df	sig (2-tailed)	t	df	sig (2-tailed)
Age:									
Factor 2	-3.321	137	0.001**	-2.318	216	0.021*	-1.294	208	0.197
Gender:									
Factor 3	-2.054	144	0.042*	-4.451	218	0.000**	-3.619	217	0.000**
Previous training: theory:									
Factor 1	-5.569	134	0.000**	-8.052	222	0.000**	-1.562	214	0.120
Factor 3	-2.751	135	0.007**	-5.554	218	0.000**	-1.796	212	0.074
Factor 4	2.305	140	0.023*	1.408	223	0.160	-0.958	215	0.339
Previous training: breakaways:									
Factor 1	-7.326	134	0.000**	-8.843	222	0.000**	-1.683	214	0.094
Factor 3	-3.464	135	0.001**	-4.628	218	0.000**	-1.215	212	0.226
Factor 4	2.941	140	0.004**	2.286	223	0.023*	-0.132	215	0.895
Previous Training: physical ::									
Factor 1	-6.585	134	0.000**	-7.278	223	0.000**	-1.529	214	0.128
Factor 3	-3.756	135	0.000**	-5.158	219	0.000**	-1.563	212	0.119
Factor 4	2.765	140	0.006**	1.574	224	0.117	0.718	215	0.473

Equal variances not assumed in above tests

*significant at the 0.05 level (2-tailed) **significant at the 0.01 level (2-tailed)

5.6 SECONDARY ANALYSIS OF SCENARIO RISK FACTOR SCORES

As indicated at the end of Section 5.4.4, an interest in the subtler changes within the open response answers to the two scenarios lead to a secondary analysis of risk factor scores. This analysis would take the interpretation to a higher level of detail and indicate changes in the proportion of identified risk factors within the different components of an integrated organisational model of violence, namely assailant factors, staff factors, environment factors, and task/interaction factors.

As previously stated the main interest was in whether the student incorporated the Unit material into their analysis of the scenarios and so moved from a 'simple' interpretation that predominantly consisted of assailant/patient factors to a more sophisticated one that incorporated, for example, more staff factors or patient risk factors following training. Hence the main difference being investigated was mention [or non-mention] of risk factors that did not involve the assailant – abbreviated to 'non-assailant risk factors'.

This intention necessitated the creation of a second database that located each approved student response under one of the four exclusive headings detailed in Boxes 4.3 and 4.4. Checks were then made that the totals of these new variables equated to the totals of the original analysis detailed in Section 5.4.4. Hence the new database

initially consisted of the four component risk factor scores for each Scenario at four time points.

Section 5.4.4. reported that the mean increase in number of risk factors identified was just under three-quarters of a risk factor for each scenario between t2 and t4, from 4.18 to 4.89 for Scenario One, and 4.37 to 5.1 for Scenario Two. A mean increase of less than one risk factor made the task much more difficult since the mean three-quarters risk factor would be sub-divided within the four component of the more sophisticated model and potentially lost to analysis. Hence the set of variables in the database was increased by creating additional variables that represented whether any non-assailant risk factors were reported by a student at each time point [yes-no], and the total number of non-assailant risk factors mentioned by a student at each time point [total number of risk factors minus the assailant risk factors].

Firstly, the proportion of student nurses mentioning a non-assailant risk factor prior to training was encouraging. Over three-quarters of students mentioned at least one non-assailant risk factor at t2 [Day One of the Unit, immediately prior to any educational input], 77% for Scenario One and 86% for Scenario Two. This suggests that, by the middle of the first year of training, the majority of student nurses have already moved on from the simplistic 'blame the patient not ourselves' explanation of workplace violence.

Statistical analysis was then completed. Strictly speaking the data represented frequency counts of categorical-level variables. Hence non-parametric chi-square analysis of these data was appropriate. However, the variables representing number of non-assailant factors mentioned were normally distributed and showed no signs of skew or kurtosis so parametric analysis using paired sample t-tests analysis was conducted. Table 5.25 shows the results of the Chi square analysis of number of students mentioning a non-assailant risk factor. It reveals that there was no significant difference in the number of students mentioning a non-assailant risk factor at any time point [it will be remembered that the number at time point t1 refers to only two cohorts].

Table 5.25: Number of student nurses mentioning non-assailant risk factors for scenarios at four time points				
Scenario One:	t1	t2	t3	t4
Mention non-assailant factors:				
Missing /No	37	55	41	55
Yes	120	188	202	188
Pearson Chi Square	Value 4.187	df 3	Significance [2 sided] 0.242	
Scenario Two:				
Mention non-assailant factors:				
Missing/ No	18	34	42	46
Yes	139	209	201	197
Pearson Chi Square	Value 2.675	df 3	Significance [2 sided] 0.445	

Table 5.26 shows the mean number of non-assailant risk factors identified at different time points, and the paired sample t-test analysis. Mean number of non-assailant risk factors identified in both scenarios increased by 12-15% following attendance of the training Unit. In the case of Scenario One there was a further increase in mean score at t4, following completion of the two clinical placements, while the mean Scenario Two score for non-assailant risk factors showed a slight decrease at t4 compared with the score on immediate completion [t3]. For both Scenarios, there were no statistically significant changes in the number of non-assailant risk factors between t1 and t2 but statistically significant change did occur following the Unit [t2-t3] that were maintained at the three-month follow-up point [t2-t4].

Table 5.26: Number of non-assailant risk factors identified at time point t1-t4 - means and paired sample t-test				
	Mean Number of Non-assailant Risk Factors			
	t1	t2	t3	t4
Scenario One	1.74	1.71	1.93	2.01
Scenario Two	1.80	1.72	2.10	1.95
Scenario One:	Paired Sample t-test			
	t	df	Significance [2 tailed]	
t1-t2	0.267	95	0.790	
t2-t3	-3.369	262	0.001	
t2-t4	-4.403	152	0.000	
Scenario Two:				
t1-t2	0.769	122	0.444	
t2-t3	-5.329	181	0.000	
t2-t4	-3.326	177	0.001	

Suggested risk factors produced by student nurses are still distorted towards assailant factors. Simple arithmetical analysis of the mean numbers shows that, following training, approximately 60% of risk factors are assailant-related [mean number of non-assailant risk factors / mean number of risk factors approximately for both scenarios = $2 / 5 = 40\%$]. Hence the interpretation of this data is hampered by the small increases in total risk factors. There is evidence that the majority of student nurses considered non-assailant risk factors and mentioned at least one prior to attending the unit; the number of student nurses including any non-assailant factors appears to be unaffected by training; the mean number of non-assailant risk factors identified by student nurses is positively influenced by training, that is, it remains stable prior to training but increases immediately afterwards and remains elevated for a period of time after training.

5.7 SUMMARY OF RESULTS

- 243 students were included in the study, response rates were between 85.4-100%, the three cohorts were considered homogeneous and representative of student nurse cohorts in training at that time, in terms of age range, gender, and destined branch
- rates of involvement in incidents of verbal and physical violence in clinical placements varied by placement type and ranged from 8.7 - 50.5% for verbal incidents and 3.0 - 43.5% for physical

incidents, highest rates being noted in mental health and learning disability settings

- where involvement in incidents occurred, the numbers of incidents for individual students ranged from 1 – 20 for verbal and 1 –10 for physical violence
- the mean number of risk factors identified in the two scenarios increased following Unit attendance by almost three-quarters of a risk factor [to approximately five risk factors in both cases]
- female students consistently identified more risk factors in the scenarios than their male counterparts
- the level of blame attributed to the patient in Scenario One reduced following Unit attendance
- older student nurses appear to blame patients less than younger colleagues for the causation of violent incidents
- a five factor solution to factor analysis confirmed the desired construction of the 24 Likert statements
- following Unit attendance students expressed more confidence in their ability to maintain their personal safety
- there was a tendency for male students to be more confident at all times than females
- following Unit attendance students expressed increased belief in the predictability and preventability of violence
- older students considered violence to be more predictable and preventable than their younger colleagues

- following Unit attendance there was evidence that students expressed a more positive self-appraisal of their own practical skills in interacting with aggressive patients or relatives
- at all times male students made more positive appraisals of the practical ability to manage aggressive patients than females
- following Unit attendance students expressed increased self-respect and belief in their rights to be safe at work, and have these rights protected by legal recourse
- following Unit attendance students were more able to recognise provocative approaches exhibited by staff
- the effects of previous training could be detected in the factor scores relating to confidence in personal ability to remain safe, self-assessed competence and self-respect but not on immediate completion of the Unit
- at the point of completing the Unit students without prior training were not statistically different to their colleagues who had received previous training in their estimations of confidence and competence
- over three-quarters of students mentioned at least one non-assailant risk factor for the two scenarios *prior* to attending the Unit, and attendance made no significant change to the number of students including at least one non-assailant risk factor
- in the case of those students who mentioned a non-assailant risk factor, Unit attendance significantly increased the subsequent mean number of non-assailant risk factors identified.

CHAPTER 6 – DISCUSSION

This chapter will provide an explanation of the results presented in Chapter 5. An attempt will be made to explain the individual effects of discrete questionnaire sections, and also, by inter-connecting these individual findings, give an overall impression of the effects of attending the training Unit. In so doing, it will provide answers to the Research Questions, and support or refute the Research Hypotheses listed at the end of Chapter 1.

The findings of this study will also be compared and contrasted with those reported by other researchers discussed in Chapter 1, and related to the theory and practice of training evaluation discussed in Chapter 2. Furthermore, a critical evaluation of the study will be made that will include comment on the effectiveness and utility of the different styles of data collection incorporated in to the questionnaire, and the strengths, weaknesses, successes and failures of the research process adopted in this particular case.

Finally it will consider the practical implications of a study of this nature in relation to the current 'political climate' which, quite rightly, moves relentlessly towards demonstration of clinical and cost effectiveness. Prior to any explanation of findings it will prove useful to present a summary of the Research Questions and Research Hypotheses that were stated at the end of Chapter 1.

6.1 RE-PRESENTATION OF STUDY RESEARCH QUESTIONS AND RESEARCH HYPOTHESES

Research Questions

The literature review raised a number of important questions which this study has endeavoured to answer:-

- Question 1:** What is the extent of the exposure of student nurses to aggression and violence during the first year of their training course?
- Question 2:** Does a relatively short, three-day unit on the prevention and management of aggression in health care settings have a positive effect on its student nurse course participants?
- Question 3:** Can desired changes in learning domains, including knowledge, attitudes, confidence, and self-assessed competence be unambiguously detected following attendance on the course?
- Question 4:** If changes in different learning domains are detected can these be monitored and subtly explored in detail and depth?
- Question 5:** Do any immediately detected changes [identified in question 4] remain, increase or deteriorate over time when the student nurses return to clinical practice?

Question 6: Do different student nurse sub-groups respond differentially to the training course?

Question 7: How do a range of techniques for measuring training effects differ in sensitivity, usefulness and ease of use?

From these Research Questions a number of Research Hypotheses were generated for investigation.

Research Hypotheses:

- 1. Student nurses are involved in substantial numbers of violent incidents during their training**
- 2. After attending a three-day training unit as part of their first year training programme student nurses will demonstrate increased knowledge about causation and prevention of violent incidents in health care settings**
- 3. After attending a three-day training unit as part of their first year training programme student nurses will demonstrate increased confidence in their ability to remain safe while interacting with aggressive clients**
- 4. After attending a three-day training unit as part of their first year training programme student nurses will more broadly attribute the blame for causation of violent incidents and not always blame the assailant**
- 5. After attending a three-day training unit as part of their first year training programme student nurses will demonstrate more**

adaptive and realistic attitudes about workplace violence and the possibility of reducing its incidence.

6. After attending a three-day training unit as part of their first year training programme student nurses will demonstrate more adaptive and realistic attitudes about the role of the student nurse in the management of violent incidents.
7. After attending a three-day training unit as part of their first year training programme student nurses will assess themselves as being more competent in interacting with aggressive clients.
8. Desired changes observed on completion of the unit will still be detected three-months later following two short clinical placements

6.2 EXPLANATION OF FINDINGS

6.2.1 The Sample: Size, Response Rate, Homogeneity, Generalisability

The analysis of published studies in Chapter 1 suggested that the highest reliable sample size was 317 in the early study by Gertz(1980) and that over half the studies included in the analysis were on samples of less than 100. The sample size of 243 makes this study one of the largest international studies, possibly the second largest performed and reported to date, and certainly the largest reported study on student nurses.

The response rate was very good on all occasions, ranging from 85.5-100%, and may be accounted for by a combination of factors, including the Department of Nursing and Midwifery's insistence on student attendance at all sessions, a 'captive audience' effect, the reassurances about confidentiality, support etc. offered in the introduction, and perhaps, the perceived relevance and importance of the topic. The tests of homogeneity of the three cohorts revealed them to be very similar, as would be expected since there had been no sudden change in recruitment policy at the time.

It is not possible to determine whether the sample demographics represented [and still represent] a typical example, in terms of age, gender and destined speciality, of many cohorts of student nurses with the U.K. since these figures are not currently kept by the Department of Health or the governing body, the NMC. Workforce Development Confederations keep details of locally commissioned numbers but accept that these may differ widely from those actually recruited, and also that attrition will affect numbers qualifying. From personal anecdotes, cohorts in some departments of nursing may now vary in terms of the number of available specialities since not all departments offer Learning Disability Nursing or Children's Nursing options. In addition, the level of training – diploma or degree - may now be different [and is different in the Department where the study occurred], wherein the proportion of training places offered at Level 3 first-degree level is

increasing. Indeed, some departments of Nursing now only offer pre-registration nurse training at first-degree level.

This difference in course level should not be over-emphasised. In the Department under study, the aggression prevention and management training is still currently provided in the first year of training and all student nurses, on either degree or diploma pathways are provided with identical content material, time, and input. Locally, cohorts still appear to have similar proportions of males, different age categories biased towards younger students and similar proportions of students enrolled on the different branches. However, it is acknowledged that the results cannot be generalised to pre-registration nurse education in the United Kingdom or further afield.

Care must also be taken before attempting to generalise these results more widely to qualified or unqualified staff working in the health services since they may well differ on a number of pertinent points. Their educational qualifications may be significantly different, their clinical experience is likely to be different [longer periods of time on fewer wards rather than the frequent shorter placements that student nurses have], the training these staff receive is also likely to include a large proportion of time dedicated to physical restraint skills. However, from personal experience of courses in the U.K. and Australia, it should be said that much of the content of the Unit under study [for example, risk factors, risk assessment, de-escalation, non-provocative approaches,

breakaways skills] is included in longer courses [5-10 days] that also include restraint training. Hence, it may well be possible to generalise some of the findings about some sections or the utility of different data collection approaches to these courses.

It is important to remember, in addition, that recent reports have concluded that the balance of training for NHS staff has been wrong (Nursing and Midwifery Council [NMC] 2004a, NAO 2003), with an over-emphasis on 'hard' restraint skills over 'soft' inter-personal, non-provocative, customer care skills. For example, in its recent Newsletter (2004a) the NMC commented on the findings of the inquiry in to the death of David 'Rocky' Bennett, a psychiatric patient who died while being restrained. The NMC professional advisor on mental health Rick Tucker, who was also an expert witness at the public inquiry said "there has been an over-emphasis on how to physically manage incidents of violence at the expense of recognition and prevention. This has led to a defensive and reactive culture that has de-skilled many practitioners. We believe the current strategy is failing both patients and carers" (NMC 2004a:7). Hence, it is felt that, in the future, the findings of this study will be increasingly relevant and potentially useful to aggression management trainers working with NHS and social services staff.

6.2.2 Extent of Student Nurse Experience of Violence

Hypothesis 1 suggested that student nurses are involved in substantial numbers of violent incidents during their training. The findings of this

study referred to the first three clinical placements that the student nurses undertook in the first year of their training, support this hypothesis and make depressing reading. The trends in the figures obtained reflect those from published surveys discussed in Chapter 1, particularly the Health Services Advisory Committee study reported by MacKay (1994).

Different definition and classification of incident type prevent exact comparison but the study results show that verbal incidents are more common than physical incidents [as did MacKay 1994] and that most incidents occurred in mental health and learning disability settings [as did MacKay 1994, and NAO 2003], where slightly over half the sample were involved in verbal incidents and slightly under half in physical violence. Averaged for the year, these rates are slightly lower than those reported for students by MacKay (1994) in that he found rates of approximately 38% [380/1000 staff/ year]for physical incidents [major + minor incidents] compared with an average of 21% [16.5 +3+43.5/3] for the year covered by the three Trimesters in this study.

For verbal threat MacKay (1994) cited a figure of 40% [402 incidents/1000 staff/year] and this study found an average of 28.4% [26+8.7+50.5/3] for the year covered by the three Trimesters in this study. The NAO (2003) study found 2.5 times more incidents in mental health and learning disability settings than other Trusts and this study found similar proportionate rates of 16.5% for Trimester One [all

placements] and 43.5% for trimester Three [mental health and learning disability placement]. Least incidents occurred in child and midwifery placements where substantially smaller percentages were reported [8.7% verbal, 3% physical].

Face validity is increased since the reported figures for the first placement [when students went to their destined speciality area] are between the extremes of the second placement [when *all* students went to child/midwifery areas] and the third placement [when *all* students went to mental health and learning disability]. Face validity is further enhanced by the analysis of Trimester One figures that showed this hierarchical trend within the placements, the differences being statistically significant.

As previously mentioned 'involvement' was defined in the initial presentation as the student nurse being the 'focus' for the incident or playing a major part in the incident and its management. Hence these findings show a serious problem and it remains to be determined as to whether this prevalence is reflected in the official statistics for aggressive/violent incident figures submitted by clinical areas.

Equally concerning is the quantity of incidents that involve student nurses. As previously mentioned, a handful of students indicated 'daily', 'numerous', or 'hundreds' in response to the question about the number of incidents. Of those respondents that were more precise, the

maximum number of incidents for a single student was 20 for verbal incidents and 10 for physical incidents [both in mental health and learning disabilities areas]. Findings that half of the total number of student nurses in the first year of their training have been verbally threatened or abused by clients/patients, some of them up to twenty times in one placement, is disconcerting. It will affect student/ staff retention and is a clear justification for the inclusion of this type of training early in nurse training programmes.

6.2.3 Number of Risk Factors Identified in the Scenarios

Hypothesis 2 suggested that "after attending a three-day training unit as part of their first year training programme student nurses will demonstrate increased knowledge about causation and prevention of violent incidents in health care settings". Hence for this to be supported the data would need to show no change in the number of risk factors identified on occasions prior to Unit attendance and an increased number listed on occasions after training.

Adequate preparation of the second rater, resulted in reassuringly high inter-rater reliability scores. The data showed low-moderate levels of test-retest reliability in the range 0.463-0.521 between t1 and t2.

Moreover, t tests showed no change in scores between t1 and t2. The scores were stable prior to training and increased at all time-points after training, and the scores differences between t2-t3, and t2-t4, were statistically significant. Hence Hypothesis 2 is supported.

However, personally, there were a couple of surprising features in the scores reported. Firstly, the range of scores recorded was much larger than expected. Given that all the students were at the same point in training, it was unanticipated that some students listed up to 13 risk factors whilst others identified only one. Secondly, the mean number of risk factors recorded was lower than expected. The final lists of acceptable risk factors contained almost forty possible risk factors for each scenario, and yet the mean number identified was less than five. The immediate 'training effect' between day 1 and day 3 of the course resulted in just about half an extra factor being identified [increase in mean number of identified risk factors t2-t3 was 0.48 for Scenario One and 0.51 for Scenario Two]. At three-month follow-up, after completing the two associated placements this increase between t2-t4 was 0.71 for Scenario One and 0.73 for Scenario Two.

Given the emphasis within the Unit on risk assessment and audit [several hours on Day 1 were devoted to it], this increase was a little disappointing. Apparently, a statistically significant difference is not the same as a gratifying educational one. On the one hand, it could be argued that the additional three-quarters of a factor amounted to a 17-18% increase over the score at t2, and that the additional factor identified could make all the difference between the student nurse recognising a situation as dangerous [and leaving] or safe [and staying].

On the other hand, it was optimistically expected that the students would identify two or three additional factors after attending the Unit.

The effects of respondent fatigue may offer a partial explanation. Open response questions place a greater cognitive demand on respondents than, say choosing between presented options. At t3 the student nurse had just completed the final day of the Unit, a very full day, half of which consisting of breakaway training. This type of training is both cognitively [memory, visualisation], psychologically [very close physical contact, higher level of involvement and visibility in a smaller group], and physically [some physical exertion in standing up, practicing breakaways] demanding.

Hence the student, more accustomed to sitting and making notes, may well have been exhausted and in a poor state to set about giving their optimum performance when completing open response questions. This situation was compounded by knowledge that they were free to leave as soon as the questionnaire was completed [obviously, they were able to leave without completing the questionnaire but this option was psychologically more difficult]. Under these circumstances they may well have decided to settle for sub-optimum response levels to the most demanding questions. It is anticipated that such a phenomenon would be less likely to occur in, say, the Likert questions since they have only to indicate level of agreement and not generate new material.

Other studies have tended to measure change in level of knowledge via multiple-choice or short answer tests and have generally reported increased knowledge at post-test (McDonnell 1997, Calabro *et al* 2002). It could be argued that multiple choice questions are less demanding than open-response since they involve elements of recognition rather than recall.

The gender effect was worthy of note. On all occasions for both scenarios the mean number of risk factors identified by female student nurses was greater than that identified by males. For Scenario Two at t3 this difference was 1.17 factors, higher than the mean training effect t2-t4. It is not possible to say whether this is a distinct, perceptual gender difference or a greater willingness on the part of females to persist with answering open questions but this observation is surely worthy of further investigation. Further mention of this gender difference will be made shortly when discussing 'confidence'.

There was evidence that student nurses who had received prior theoretical aggression training or breakaway training identified a slightly higher mean number of risk factors than those without prior training but this difference was rarely significant. More intriguing is the absence of such a consistent effect for those student nurses who had received restraint training. Experience of violence in Trimester One fell in to the same category with students who reported involvement in verbal or physical incidents identifying a higher mean number of risk factors for

both scenarios but again this effect was rarely statistically significant. Here, there is possibly some limited evidence of students learning experientially, that is reflecting on meaningful clinical personal experience and incorporating this into their knowledge base.

6.2.4 Attribution of Blame

Hypothesis 4 suggested that attending the training Unit would cause student nurses to “more broadly attribute the blame for causation of violent incidents and not always blame the assailant”. An associated secondary ‘prediction’ was that the elderly woman in Scenario Two would be perceived as less to blame than the young man in Scenario One.

The two Scenarios produced different effects. Once again the test-retest reliability coefficients between t1-t2 were in the low-moderate range 0.436-0.513, although the t-tests scores confirmed that there was no change in VAS scores for either Scenario between t1-t2.

Mean VAS scores for Scenario One corresponded to the anticipated pattern [lower score equating to less patient blame] with a small reduction t1-t2 prior to training being followed by a larger reduction t2-t3 on immediate completion. The reduction lessened slightly t2-t4 at follow-up, but remained statistically significant.

Scenario Two scores were lower at all time points indicating support for the secondary prediction about relative blame. Otherwise the trends in

Scenario Two were much less clear and no changes between any time points provided statistically significant. In fact the highest levels of blame were reported at t4 after completing the two clinical placements, many of which would have including caring for older clients.

It has been reported that staff [and therefore student nurses] may identify with clients /patients and then minimise, rationalise or 'excuse' their abhorrent behaviour. Bi-variate tests showed no gender effects but did reveal an age effect in that, on all occasions for both Scenarios, student nurses older than 26 years blamed the patient less than their younger colleagues. It could be that the life and work experience of older student nurses allowed them to view the situations more holistically but, as reported above, they failed to demonstrate this when identifying risk factors for the two scenarios. Given that the patients in both Scenarios were in the over 26 years category, it is possible that older student nurses identified with them and possibly made some allowance for the behaviour.

An intriguing bi-variate observation without an obvious explanation concerns the attribution of blame of those student nurses who had been involved in physical violence in Trimester One. Their mean attribution of blame score was *lower* than that of their colleagues [who had not been involved in physical violence in Trimester One] for Scenario One at time t1, t2, and t3 and *higher* for Scenario Two on all occasions although these differences were not statistically significant. It is not obvious why

they should tend to blame the younger man less than their colleagues, and the elderly woman more than their colleagues, unless their involvement in violence had been directly linked with aggressive old ladies and they had generalised a stereotype from this experience. Intriguingly, many, probably the majority, of the Trimester One clinical placements were in 'Care of the Elderly' areas.

The observation that virtually all possible scores for both scenarios were recorded at all time points provides an insight into the diverse factors that must contribute to attribution of blame and the complexity of studying this area further. It should be noted that the attribution of blame was not considered specifically or separately to any extent within the Unit and so the change that was sought via the VAS score was an indirect one. It was anticipated that incorporation of a multi-factorial model for causation of violence would distribute the attribution of blame more widely between a number of factors related to the assailant but also to staff, the environment and to the interaction or task. No other studies were found that measured attribution of blame.

6.2.5 Likert Statements

The Likert statements were designed to evaluate attainment of the Unit learning outcomes related to, knowledge and awareness, beliefs and attitudes, self-esteem, confidence, and skill [competence]. The statements also relate to Hypotheses numbers 2, 3, 4, 5, 6, 7, and 8. Box 5 in Chapter 4 illustrated the assumed links between evaluation

model domains, learning outcomes, research hypotheses and statements on the questionnaire. The initial screening of statements for endorsement frequency and restriction of range, and subsequent performance of the principal components factor analysis resulted in a loss of six statements from the original list of 24. This equates to a 25% reduction, and such a loss needs consideration when generating an initial pool of statements.

The factor analysis was performed to ensure that the topics in each Unit learning outcome [and Hypothesis identified above] were indeed present. In this regard it is felt that the procedure worked very well. Five factors were identified and further study revealed that all of these relate directly to learning outcomes and research hypotheses. Simplistically one could state that the factor analysis procedure is intended to identify separate orthogonal [un-related] factors and one could say, crudely, that Factor 1 [confidence] is about attitudes, Factor 2 [prediction] is about knowledge, Factor 3 [ability] is about behaviour, Factor 4 [respect] is about attitudes, and Factor 5 [provocative approach] is about knowledge [or behaviour?].

However, in practice knowledge, attitudes, emotions and behaviours are inter-connected. Attitudes are recognised as containing knowledge, emotional and behavioural components, while competent skill performance is seen to incorporate knowledge, attitude and behaviour.

Indeed, as stated in Chapter 3, the fundamental ethos of the training curriculum was to produce the 'knowledgeable doer', a nurse with high level practical skills backed by up-to-date knowledge about when to perform the skills, how to modify the skill in response to different criteria, when not to perform the skill, etc., and also be aware of the inter-personal, and attitudinal components that need to accompany the skill performance.

Whilst the 24 statements were written to allow a determination of the achievement of learning outcomes it is felt that the relationship between a particular group of statements [a Factor] and a Research Hypothesis was generally not a simple, unitary linear one. It can be seen that some of the Factors, particularly Factors 2, 3, 4, and 5, can be viewed as relating to more than one Research Hypothesis. More precisely

- Factor One [Confidence in Maintaining Personal Safety] relates to Research Hypothesis 3 concerned with increased confidence post –Unit
- Factor Two [Prediction and Prevention] relates primarily to Research Hypothesis 5 regarding 'more adaptive and realistic attitudes about workplace violence and the possibility of reducing its incidence'. However it can also relate to Research Hypothesis 2, concerned with increased knowledge of causative and preventive factors, and to Research Hypothesis 4 concerned with broader attribution of the blame for causation of violent incidents.

- Factor Three [Personal Practical Ability] relates to Research Hypothesis 7 concerned with increased self-assessed competence 'in interacting with aggressive clients increased confidence' post –Unit, and also contains Statement 15 regarding the role of the student nurse which specifically relates to Research Hypothesis 6 on 'the role of the student nurse in the management of violent incidents'.
- Factor Four [Self-Respect and Staff Rights] was more difficult to locate but it seems to relate to Research Hypothesis 2 about knowledge of risk factors and causation [particularly with respect to Statement 1 about the inevitability of being assaulted], and to Research Hypothesis 5 regarding attitudes to workplace violence [via Statements 3, 5, 24 dealing with self-protection, recording and documentation, and prosecution].
- Factor Five [Provocative Approach] contained only two statements and is concerned with extreme approaches – submission and authoritarian- which relate to Research Hypothesis 2 on knowledge of risk factors and causation, and Hypothesis 4 regarding a broader attribution of the blame for causation of violent incidents and not always blame the assailant.

6.2.6 Factor Scores

6.2.6.1 Factor One [Confidence in maintaining personal safety]

This Factor displayed high levels of stability prior to the Unit and also high internal consistency. The changes in the scores of Factor One over time offer considerable support for Research Hypothesis 3 in that student nurses report increased confidence in maintaining their own personal safety on completion of the training Unit. Furthermore, this effect was maintained at the three-month follow-up offering support for Research Hypothesis 8, although the major analysis related to this Hypothesis will be considered separately later.

The mean scores for this Factor reduced by over 40% between t2-t3 and represent a gratifying increase in student confidence. This effect has been noted by other researchers, for example, recent Australian studies by McGowan *et al* (1999) and Ilkiw-Lavalle *et al* (2002) both reported increased confidence on immediate completion of training. The three-month follow-up finding is more difficult to compare since generally studies failed to monitor changes in immediate post-training effects over time. Ilkiw-Lavalle *et al* (2002) failed to monitor follow-up while McGowan *et al* (1999) conducted a partial follow-up analysis at 6 months. Thackrey (1987) reported that increased confidence was still detected at 18-month follow-up.

The gender difference noted in this study, wherein male students felt more confident at all time points, although only at a statistically

significant level at t3 was also reported in a small number of studies reviewed by Allen (2001). It seems males are more confident prior to training and remain so, even though, as identified earlier [Section 6.2.3], they identified few risk factors in both scenarios on all occasions.

Possible explanations for this apparent inconsistency include perceptual distortions [males wrongly seeing themselves as more skilled, bigger, stronger, able to cope etc.], genuine obliviousness [not seeing the risks so not being concerned about them] or personality differences [possibly laziness - males perceiving risk factors in the Scenarios but failing to write them down].

With regard to the differential effects of previous training on Factor One scores the results showed that at all time points the mean 'Confidence in maintaining personal safety' score for students with any form of previous training was lower. That is to say, the students with previous training were *more* confident than students without previous training. Interestingly, this effect was statistically significant at time t1, t2 and t4 but not at t3. It would be hoped/expected that previous training increased confidence at t1 and t2 prior to this Unit. However these results suggest that on immediate completion of the Unit the confidence of students without previous training has risen more than those with previous training [which could have amounted to 5-10 days duration] and any remaining difference has lost statistical significance.

Unfortunately, at three-month follow-up after exposure to the two clinical placements the statistical difference in scores had returned. It is possible that a transferability or [in]congruity issue exists between the Unit themes and messages and the 'real world' clinical placements. Observing incompetent, poorly skilled staff, being told exaggerated horror stories about super-human strength, or given inaccurate information about physical touch or rights to defend oneself may have undermined the message of staff rights and usefulness of self-defence skills proffered in the Unit. This transferability theme will be further developed in Section 6.2.6.3.

This 'differential improvement of previously untrained staff' effect has also been noted in other studies. For example, the works of Allen & Tynan (2000) and Ilkiw-Lavalle *et al* (2002) both demonstrated relatively greater increases in performance by staff without prior training, reaching levels similar to [but still slightly less than] staff with previous training on course completion. Neither of these studies conducted a follow-up to determine possible enduring effects.

6.2.6.2 Factor Two [Prediction and prevention]

A similar desirable pattern was obtained in this case. The scores had lower test-retest and internal consistency coefficients but were stable prior to the Unit, changed in the desired direction at t3 and maintained the change at t4, both statistically significant. As such these changes offer support for Research Hypotheses 2, 4, and, in particular, to

Research Hypothesis 5. The age effect is interesting in this case. Older students [>26 years] had lower mean scores at all time points, statistically significant at t1, t2, and t4, suggesting that 'older' student nurses thought violence to be *more* predictable and preventable than their younger colleagues.

Possibly older students were more likely to have worked in health or social care prior to commencing training and so received previous aggression management training of some sort that could account for their belief in the predictability and preventability of aggression. However, the bi-variate analysis failed to reveal any significant correlation between previous training and Factor Two [although it identified one for Factors 1, 3, and 4] so any explanation has to be more intriguing.

Possibly, it is to do with relevant previous personal or clinical experience other than previous training. Put simply, working in clinical areas with clients would have allowed the respondent to observe the pre-cursors of incidents. They may have been privy to staff discussions or seen, say, the characteristic behavioural changes that many clients exhibit prior to an incident. No other studies were identified that particularly studied effects in this area.

6.2.6.3 Factor Three [Personal Practical Ability]

This factor relates directly to Research Hypothesis 7 concerned with increased self-assessed competent 'in interacting with aggressive clients'. However, it also contains Statement 15 regarding the role of the student nurse, which specifically relates to Research Hypothesis 6 on 'the role of the student nurse in the management of violent incidents'.

The inclusion of all four statements identified in the factor analysis produced inconsistent results. Indeed, the factor analysis suggested that Statement 15 dealing with the role of the student nurse in a violent incident should be scored in the opposite way to that which was anticipated when it was combined with the other three statements to create Factor Three. Test-retest reliability was moderate at 0.546, mean internal consistency was 0.513 and mean inter-item correlation was adequate at 0.259. The factor scores were not stable in the time period prior to the module, moving in the opposite direction to that which was anticipated.

Statistically significant changes between t2-t3, and t2-t4 were in the expected direction [reduction] for statements 13[rescored], 14, and 21 thus indicating increased self-assessed competence. This finding is more pronounced if Statement 15 is removed from the calculation but both results offer support for Research Hypothesis 7. Separate analysis of Statement 15 revealed interesting changes. The message presented in the Unit was that student nurses are temporarily placed in

clinical areas and inexperienced both in the general management of aggression and the particular individualised management of clients. Hence they should leave the management to regular staff but be available in the 'background' to give assistance.

Changes in mean scores of the original responses were in the anticipated direction between t1-t2, and t2-t3 [reduction indicating agreement with the statement]. However, there was a large unanticipated increase in scores between t2-t4, following the two clinical placements, which returned the score to its pre-Unit level. It is probable that a combination of the large unanticipated reduction t1-t2 prior to the Unit, and the large unanticipated increase t2-t4 confounded the factor analysis and resulted in the illogical bi-directional tension within the factor and inconsistent factor score.

Hence it appears that Research Hypothesis 6 is supported at t3 but not at t4 [more related to Research Hypothesis 8]. The large increase in score for Statement 15 at t4, following the two clinical placements may indicate the previously mentioned [Section 6.2.6.1] gulf between the idealistic nurse educator and the reality of nurse training wherein student nurses are often allocated to poorly or inadequately staffed clinical placements. In reality the student nurse may not be considered a super-numerary, additional observer/participant but, in effect, a necessary part of the clinical team. Whilst this greater involvement may prove beneficial to many student nurses, possibly allowing them greater

exposure to clinical experience and a sense of belonging to a team, it must also be recognised that it could result in too great an expectation being placed on other students.

Once again there are intriguing bi-variate effects between Factor Three scores and gender or previous training. Males viewed themselves as being more competent at all time points than females while students with previous training of any type [theory, breakaway, restraint] considered themselves more competent to manage aggression than colleagues without previous training. These effects were present at all time points for both variables, and also statistically significant at all time points for gender but only t1, t2, and t4 for previous training. Turning these results around, once again it is personally gratifying that on immediate completion of the Unit there was no statistically significant difference between the competence ratings of those students who had received previous aggression management training prior to this Unit and those who had not.

An obvious criticism of this Factor is that the scores relate to self-assessed competence and are therefore entirely subjective. No attempt was made to objectively measure the competence displayed by students in the interactions with aggressive clients, or in role-plays with actors. With 80 students in a cohort this would have been a serious additional logistical endeavour and very time-consuming. As previously

explained in Chapter 3 the Unit was added to an existing curriculum timetable and time was tight.

Some researchers have taken this self-assessed approach (Collins 1994, McGowan *et al* 1999) whilst a few others have used independent assessment of performance in role-played situations (Paterson *et al* 1992, Philips & Rudestram 1995, McDonnell 1997) but on relatively small samples [<25] it must be said. Paterson *et al* (1992) demonstrated increased competence in de-escalation, disengagement and restraint skills post-test as judged by independent raters. Philips & Rudestram (1995) demonstrated increased competence post-test compared to pre-test and control groups, as measured by independent judges. McDowell (1997) only assessed skill performance in restraint against a set of criteria at post-test but showed that all attendees passed.

An additional potential problem is the phenomenon called 'response shift bias' referred to in Chapter 4, wherein, because of an internal recalibration of the construct, a respondent judges no change in their performance over time, when in fact it has changed. If self-assessment were to be persisted with then a solution to this possible problem would be to ask the respondent to perform a 'retrospective pre-test' when completing the post-test [how it was before the intervention and how it is after the intervention] (Arvey & Cole 1991) or estimate performance relative to any earlier time, say, pre-Unit [better than before, same as before, worse than before etc.].

6.2.6.4 Factor Four [Self-Respect and Staff Rights]

The formulation of this Factor was rather unexpected and more difficult to associate to any one particular Research Hypothesis. However it did relate to Research Question 2 with regard to the phrase “have a positive effect on its student nurse course participants” and it was also associated with Research Hypothesis 2 about knowledge of risk factors and causation [particularly with respect to Statement 1 about the inevitability of being assaulted], and to Research Hypothesis 5 regarding attitudes to workplace violence [via Statements 3, 5, 24 dealing with self-protection, recording and documentation, and prosecution].

This factor had both the lowest mean internal consistency 0.407 and mean inter-item correlation 0.1546 of the five factors, and a low test-retest reliability of 0.455. The desirable responses to the four statements were correctly detected by the factor analysis process, that is to say, anticipated disagreement with Statements 1, 3, and 5 and agreement with Statement 24. Following re-scoring of Statement 24 then the scores for Factor Four were expected to *increase* post –Unit.

As in the case of Factor Three, the factor scores prior to the Unit were unstable although in this case the statistically significant change in scores t1-t2 was in the anticipated direction. It is possible that the first completion of these particular questions prompted a re-consideration of the issues of staff rights and self-valuation that was reflected in the

responses immediately pre-Unit at t2. This point would illustrate the different considerations of the researcher and educator. As an educator it was gratifying that a desirable change occurred in response to an educational intervention [questionnaire completion]. However, as a researcher, this effect is seen as an example of undesirable 'reactivity', wherein completion of a supposedly neutral data collection tool in itself produces a change in behaviour, attitude etc.

Notwithstanding this there was a larger, statistically significant increase in factor scores on immediate completion of the Unit, which can be attributed to Unit attendance that was almost completely maintained following the clinical placements. It may be that the slight reduction resulted from the previously mentioned, [Section 6.2.6.1 and 6.2.6.3] apparent barriers to transfer and realities of clinical placements. For example, seeing incidents going unreported, hearing stories about staff being unsupported in their attempts to proceed with prosecutions against clients, or listening to negative, jaundiced pronouncements about the inevitability of being assaulted, may eke away at the gains in self-respect produced by attending the Unit.

Once again the effects of previous training of any sort on Factor Four scores are interesting. It appears that previous training increased self respect and belief in the right to be able to work safely, hence there were differences in scores of students who had received previous training compared to those who had not at t1, t2 and t4 [statistically

significant for all sorts of training at t1 and t4]. The fact that no differences were discernable at t3, on immediate completion of the Unit, once again, illustrates the ability of training to, at least temporarily, reduce the differences between the previous educational experience of student nurses.

Three of these statements originated from the published work of Poster and Ryan on their 'Attitudes towards Patient Physical Assault Questionnaire' (Ryan & Poster 1993, Poster & Ryan 1994, Poster 1996). Statements 3 and 24 were incorporated unaltered whilst Statement 5 received minor modification. Statement 1 was a modified version of an attitude statement included by Collins (1994). Poster and Ryan limited themselves to descriptive surveys of staff attitude at one time-point so no comparison can be made with regard to intervention effects. Collins (1994) performed a pre-post unit comparison and found that the percentage of staff agreeing with a statement similar to Statement 1 decreased following a five-day course.

6.2.6.5 Factor Five [Provocative Approach]

This factor contained the minimum possible number of statements and is concerned with extreme approaches –submission and authoritarian– that are associated with increased rates of assault. The factor concerns staff-centred risk factors and the assailant's response to them and, as such, is associated with Research Hypothesis 2 on knowledge of risk factors and causation, and Hypothesis 4 regarding a broader attribution

of the blame for causation of violent incidents and not always blame the assailant.

Factor five had the lowest test-retest reliability of the five factors [0.377], with more respectable mean internal consistency 0.5435 and acceptable mean inter-item correlation of 0.378. Following Unit attendance greater agreement with the statements was anticipated producing a *decrease* in factor scores. Results corresponded to the anticipated format. There was stability prior to the Unit, statistically significant change in the desired direction immediately on completion of the Unit that was still evident following the two clinical placements.

These two statements were modified versions of two attitude statements used by Poster and Ryan in their work mentioned above which sought to link increased risk of assault with staff personality traits and competence level. Unfortunately, these studies reported surveys of staff attitudes and failed to monitor change over time following an educational intervention so no comparison is possible.

6.2.7 Reliability of Factor Scores

As previously mentioned, there is still no agreement about one acceptable level of reliability. In addition it may be that some of the variables and factors created in the questionnaire may be 'fluid' in nature. Only Factor One attained a good mean test-retest reliability score and also satisfied Nunnally's (1978) standard of internal

consistency. Others achieved either a moderate level of test-retest reliability [Factor Three] or internal consistency [Factors Two, Three and Five]. All factors achieved acceptable levels of mean inter-item correlation.

Given the nature of the Cronbach alpha formula, increasing the number of statements in each factor would increase the reliability coefficients of the factors. Hence simply increasing the number of Likert statements in the questionnaire to produce a longer questionnaire would increase the number available to the factor analysis procedure. However, it should be remembered that the resulting increased reliability coefficients are an artefact of the formula and the process will not necessarily result in increased correlations between individual statements or greater validity of the instrument (Nunnally 1978).

Similarly, Chapter 5 highlighted the improvement in Cronbach alpha coefficient of internal consistency for Factor Three if Statement 15 was removed. The dilemma is that a small increase in internal consistency coefficient is obtained at the cost of removing statements from the analysis and creating factors with few items. In this case the effect would be starker because it would result in the elimination from the analysis of Statement 15 that related to Research Hypothesis 6. Indeed, it was the only statement to relate directly to this Hypothesis. Hence, it was felt that there would be a net loss to the analysis if Statement 15 was removed. This decision reflected the opinion of Watson (1995) who

suggested that the substantive properties of an instrument should be considered in addition to its statistical ones.

6.2.8 Secondary Analysis of Identified Risk Factors

This stage of the analysis proved to be more difficult and less informative than originally intended due, in large part, to the lower than hoped for number of risk factors identified and lower than expected increase in number of risk factors following training. It had been optimistically envisaged that students might typically identify 5-6 risk factors in scenarios prior to training, and 8-10 after training. In this case it may have been possible to 'track' the location of the additional risk factors and, hence, say something more detailed about the subtler effects and emphases of training.

In the event, as has been discussed in Section 6.2.3, the training effect increased mean scores by three-quarters of one risk factor to slightly under [Scenario One] or slightly over [Scenario Two] five risk factors after training. Thus, finding the meaningful location of the additional three-quarters of a risk factor became much more statistically unlikely. The analysis was modified from the original intention to track every risk factor category [assailant, staff, environment, task] to a simplified one, namely, mentioned assailant factors and mentioned non-assailant factors [sum of staff, environment and task factors].

Firstly, the finding that over three-quarters of students mentioned a non-assailant risk factor before completing the Unit was encouraging and clearly suggests that the majority of student nurses have already begun to develop a more sophisticated mental model of causation of workplace violence. The results of this statistical analysis showed that there was no significant change in the *number of student nurses* mentioning a non –assailant risk factor after completing the training. There was a statistically significant 12-15% increase in the *mean number of non-assailant risks* identified for both Scenarios [mean number approximately 2] following training that persisted at three-month follow-up.

It still remains that the suggested risk factors produced by student nurses are distorted towards assailant factors. Arithmetical analysis of the mean numbers shows that, following training, approximately 40% of risk factors fell within the three non-assailant related categories, while 60% related to the single category of assailant factors. However, there is also definite evidence that the presentation and development of a multi-factorial model of workplace violence had a positive impact on the mental conceptualisation of the causation of the violent incidents presented in the two Scenarios.

6.3 SUMMARY OF FINDINGS IN RELATION TO RESEARCH

HYPOTHESES

There was **clear evidence to support** Research Hypothesis 1 that “student nurses are involved in substantial numbers of violent incidents during their training”:-

- Over 50% of student nurses has been involved in verbal abuse on learning disability and mental health clinical placements in their first year of training
- Slightly less than 50% of student nurses has been involved in physical violence on learning disability and mental health clinical placements in their first year of training
- Figures for child and midwifery placements were much lower [but not zero].

There was **clear evidence to support** Research Hypothesis 2 that “attending a three-day training unit ... will increase knowledge about causation and prevention of violent incidents in health care settings”:-

- Student nurses identified more risk factors in two scenarios after attending the Unit
- Females student nurses reported more risk factors in the two scenarios than male student nurses on all occasions, significantly more on most occasions

- Students with prior theoretical training [but not breakaway or restraint training] identified a higher mean number of risk factors in both scenarios on all occasions but this difference was rarely significant
- Students who had experienced verbal or physical violence in their first placement identified a higher mean number of risk factors than other students but this difference was rarely significant
- Changes in Factor Two [Prediction and prevention] scores showed an increased belief in work-place violence being both predictable and preventable
- Changes in Factor Four [Self-respect and staff rights] scores, particularly in relation to Statement 1 about the apparent inevitability of assault showed increased knowledge of the preventability of violence
- Changes in Factor Five [Provocative behaviour] scores showed increased knowledge of undesirable, provocative staff behaviours that are associated with increased rates of involvement in incidents.

There was **clear evidence to support** Research Hypothesis 3 that Unit attendance will provide students with “increased confidence in their ability to remain safe while interacting with aggressive clients”:-

- Changes in Factor One [Confidence in maintaining personal safety] scores showed a pronounced increase in confidence in one's ability to remain safe at work.
- Males were more confident in their ability to remain safe than females at all time-points
- At all time points student nurses with previous aggression management training were more confident than those with prior training. On immediate completion of the Unit this difference lost any statistical significance suggesting that attending the Unit had a disproportionately positive effect on those students without prior training.

There was **limited evidence to support** Research Hypothesis 4 that Unit attendance and presentation of a multi-factorial model of causation of work-place violence would result in student nurses more broadly attributing “the blame for causation of violent incidents” rather than always blaming the patient;-

- VAS scores for Scenario One showed that students reduced the proportion of blame attributed to the patient following Unit attendance
- However, VAS scores for Scenario Two showed no such effect. In fact, the proportion of blame attributed to the patient increased over time and was highest following the two clinical placements

- Students were able to differentiate the proportion of blame attributed to the patient in each scenario, consistently blaming the elderly woman much less than the younger man
- Older student nurses [>26 years] attributed less blame to the patient in both scenarios than their younger colleagues to a significant degree on most occasions
- Students who had experienced physical violence in their first placement tended to blame the young man less, and the elderly woman more, than their colleagues on most occasions but not to any significant level
- The 'qualitative' analysis showed that over three-quarters of student nurses identified at least one non-assailant risk factor and, furthermore, the mean number of non-assailant risk factors identified in both scenarios increased statistically significantly following training
- Changes in Factor Two [Prediction and prevention] scores would indicate that more accurate prediction and prevention would take into account a larger number of risk factors including those related to staff, clinical environment, task etc.
- Changes in Factor Five [Provocative behaviour] scores would indicate a growing awareness of the influence of provocative staff behaviours

There was **clear evidence to support** Research Hypothesis 5 that attending the Unit will cause student nurses to “demonstrate more adaptive and realistic attitudes about workplace violence and the possibility of reducing its incidence”:-

- Changes in Factor Two [Prediction and prevention] scores showed an improved positive attitude towards the predictability and preventability of work-place violence
- Changes in Factor Four [Self-respect and staff rights] scores demonstrated that post-Unit student nurses had increased self-respect and more adaptive and realistic views of their rights, for example, to protect themselves or to pursue a legal case against an assailant
- Changes to Factor Five [provocative behaviour] scores also showed an increasingly positive attitude towards the role of staff behaviour in the causation of work-place violence

There was **limited transitory evidence to support** Research Hypothesis 6 that Unit attendance would result in student nurses demonstrating “more adaptive and realistic attitudes about the role of the student nurse in the management of violent incidents”:-

- Separate analysis of Likert Statement 15 showed that student nurses demonstrated increased agreement about their 'background' role on immediate completion of the unit
- However, this effect was completely under-mined by the realities of clinical placements

There was **clear evidence to support** Research Hypothesis 7 that Unit attendance would result in student nurses assessing themselves as "being more competent in interacting with aggressive clients":-

- Changes in Factor Three [Personal practical ability] scores, particularly when only three relevant items are included, indicate that students rate themselves as more competent in their clinical interactions after attending the Unit
- Males student nurses rated themselves as more competent at all times than their female colleagues to a statistically significant degree
- Student with previous training of any type rated themselves as more competent in interacting with agitated or aggressive clients at all time points than those without prior training, statistically significant at t1, t2, and t4.
- Previous training made no statistical difference on immediate completion of the Unit indicating that the Unit had a

disproportionately greater positive effect on the self-assessed competence of those students without previous training.

There was **clear evidence to support** Research Hypothesis 8 that changes produced by attending the Unit would endure for three-months following two short clinical placements:-

- The number of risk factors identified in the two scenarios at the follow-up time-point remained significantly higher than pre-Unit total
- Factor One [Confidence in maintaining personal safety] scores at the follow-up time-point remained significantly lower than pre-Unit scores indicating that confidence was maintained during and after the placements
- Factor Two [Prediction and prevention] scores at the follow-up time-point remained significantly lower than pre-Unit scores indicating that the student's belief in the predictability and preventability of work-place violence endured the two clinical placements
- Factor Three [Personal practical ability] scores at the follow-up time-point remained significantly lower than pre-Unit scores indicating that the students' self assessment of their own competence was not eroded by the two clinical placements,

indeed mean scores were lower than at t3, indicating higher self-rated competence

- Factor Four [Self-respect and staff rights] scores at the follow-up time-point remained significantly higher than pre-Unit scores, although they were slightly lower than at t3, indicating that self-respect and appreciation of staff rights was not unduly affected by the clinical placements
- Factor Five [Provocative behaviour] scores at the follow-up point remained significantly lower than pre-Unit scores, although slightly higher than at t3, indicating that the students' belief that staff behaviour influenced the rate of violence was unaffected by the time interval and clinical placements
- However, VAS scores relating to attribution of blame in the two scenarios at the follow-up point were higher than at t3 in both cases, indicating a greater amount of blame being attributed to the patient, and reversing the movement achieved by the Unit. The scenario concerning the elderly woman produced a larger amount of blame for the elderly woman at follow-up than at any other time-point
- However, scores for Statement 15 with regard to the role of the student nurse in the management of violent incidents at follow-up showed a reversal of the trends achieved by Unit attendance at t3.

In addition, two Research Questions involved areas not directly covered in the Research Hypotheses. Research Question 4

asked about subtler monitoring of differences in questions, while Research Question 6 queried possible differences in responses between different student types. In this regard,

- there was some evidence to support the subtle changes in responses to the two Scenarios over time.
- There were some gender effects, age effects and previous training effects but different destined speciality didn't seem to make any difference.

6.4 UTILITY

It is important that a scale has reliability, validity etc. but a different practical consideration has to do with the 'ease of use' or utility of the measure. Indeed, this second consideration may become the main one if a scale was being considered for regular use on large numbers of respondents. A highly reliable and valid evaluation form could be rendered effectively 'useless' as a routine evaluation measure if it took high levels of skill to complete, or an inordinate amount of time to collate and analyse.

It should be remembered that, in this particular study, the questionnaire had a limited shelf life, since it was designed specifically to measure the effects of the training Unit on three cohorts of course attendees. The intention was to take the 'longer view' and incorporate data from medium term monitoring of the course participants subsequent interaction with the 'real world'. Time constraints and deadlines were

not an issue and there was no requirement or necessity to produce immediate findings that would result in changes to the next running of the module.

Research Question 7 queried the relative usefulness of the different styles of questions incorporated in to the questionnaire. On reflection, it can be seen that there were huge differences in the utility of these different measures and also differences in what could be called the cost-benefit ratio. This is the power and usefulness of results obtained in relation to the time and effort needed to process them.

6.4.1 Likert-type questions

In this case the Likert-type questions were by far the easiest part of the questionnaire to process and analyse. Data input was simply a matter of typing in one number between 1-5 for each of the 24 statements. After checking range, minimum, maximum etc., statistical interpretation was straight-forward. The questions lent themselves to factor analysis and, as has been seen, the factors related closely to the anticipated areas. Generally, the evidence of changes in factor scores over time provided a clear indication of the immediate and longer-term effects on first-year student nurses of the training Unit and associated clinical placements. Hence, the Likert statements were low on cost and high on benefit.

6.4.2 Visual Analogue Scale

The VAS data associated with attribution of blame proved to be the next most easy to interpret, input in to the database and analyse. As previously described, the interpretation required the over-laying of a transparent ruler on to each VAS scale and estimation of the point on the scale where the student's mark bisected the line. Hence the process necessitated the accurate measurement of over 2,100 lines. In most cases this process was tedious but straightforward. It would have been made more difficult if the two lines had been located on different pages.

However, a minority proportion of students, 5-10% maybe, made the calculation of the blame score more difficult by failing to indicate their decision with a single, fine bisecting line, thus \. Instead they sometimes used broad marker pen lines or circles of different radii [sometimes approaching one centimetre]. This obviously slowed down the process since the mid-point of these responses had to be estimated and then measured. Presumably, in future, more precise verbal/written instruction and illustrated guidance could eliminate this difficulty.

Furthermore, the results of the statistical analysis were less uniform. As has been seen the scores for the two scenarios followed the predicted pattern relative to each other. The changes in Scenario One [younger man] scores over time also followed the predicted pattern but Scenario Two [elderly woman] showed interesting changes and an unpredictable trend.

It is difficult to know whether this effect was an accurate reaction to the clinical placements, many of which would involve elderly care settings, an artefact of a poorly constructed scenario statement, or an alien concept in relation to the Unit material. Certainly, the concept of blame attribution was not central to the Unit material. Whilst there was considerable time allotted to violence risk factors, theories and models, and 'causes' of violence, the word 'blame' was not consciously used at all. Thus, the VAS scales were moderate on cost and moderate on benefit.

6.4.3 Open Response Questions

At all stages of data analysis the open responses identifying risk factors in the scenario descriptions proved the most difficult to analyse. For example:-

- the list of 'acceptable' responses was developed along with inclusion rules and amended after reading a number of completed questionnaires [approximately 250 forms]
- a determined attempt was made to delineate the different sections in the list – factors associated with staff, assailant, environment, and task – so that subsequent relative changes in the number of risk factors associated with each section could be monitored
- a separate database was subsequently created to record the number of factors under each heading, staff, assailant etc. The data entered had to be checked so that the total of sub-section

scores equalled the total used earlier, or amendments to the total had to be made. This checking required the creation of additional variables so that totals could be compared and matched

- A second scorer was used to gain a measure of inter-rater reliability

The creation of the second database especially required great concentration and familiarity with the accepted terms and rules for inclusion. The results obtained showed that the number of risk factors identified within the scenarios increased immediately post-Unit and that this effect was maintained for the three-months of follow-up. This effect was statistically significant but not dramatic.

The open-response type of question was the most cognitively demanding to complete. It should be remembered that the questionnaire was administered at different times of day, under different circumstances on the four occasions and this may have introduced some administration-related bias. At t1 and t4 the questionnaire was completed within time negotiated in sessions delivered by other lecturers. At t2 the questionnaire was administered at the start of the first day of the Unit.

At t3 it was administered after the Unit had officially finished, at the end of a long study day involving mental and physical exertion. With only the

psychological pressure to complete the questionnaire separating the student from freedom to leave the classroom, go home etc., one could easily imagine that there would be a tendency to rush this section, or fail to give it the full attention, and so not list all risk factors that were detected. This would depress the Day 3 [t3] score and disguise the immediate full effect of the Unit.

Thus the fact that the t4 [three-month follow-up] scores are higher than t3 scores may indicate continued learning occurring in the clinical placements. Unfortunately, it may also indicate an actual reduction in risk factors identified from t3-t4, undetected because the t3 score was an under-estimate of the true effect of the Unit. It may have proved useful to administer the questionnaire at t3 in the same manner as at t1 and t4, that is to say, by negotiating administration time in another session very shortly afterwards, when there was no 'immediate gratification effect' for scant, or partially considered completion.

It seems difficult to justify the 16+ hours [three working days equivalent] needed to input the individually categorised risk factor scores into the additional database. On this occasion the results were certainly of some use but the time commitment means that it could not be recommended as a regular feature of training course evaluation.

6.5 CRITICAL EVALUATION

The perfect piece of research is still awaited. Chapter 4 established the chosen research design, described the precise methodology and provided a rationale for their selection. It was acknowledged that the research had to take place in the real world of pre-registration nurse education.

A primary constraint was that, in reality, a packed nursing curriculum being delivered to six different cohorts at any time left little room for manipulation. In addition an acknowledgement of student nurses being at heightened risk, subsequently borne out by the results obtained in this study, meant that delaying presentation to create control groups was unethical. Furthermore, the communal nature of nurse education also meant that any control groups were likely to become 'contaminated'.

An interrupted time series design was adopted that allowed the stability of scores on different learning domains prior to the training intervention to be established, and also permitted the follow-up of student post-Unit for the remaining three months of that Trimester. Thus, as planned, the selected research design overcame a number of the recognised potential deficits detailed in Chapter 1, and so distinguished this research from that previously published on a number of counts:-

- Firstly, it will be recalled, follow-up beyond immediate, end of training effects was rare in the literature, and the improved

research design permitted the nature of the interaction between the post-Unit student and the influential forces in the real-world clinical environment to be investigated.

- Secondly, the focus on student nurses, an acknowledged high-risk group, was unusual and certainly an overdue priority for a study of this kind. The typical cohort reported in the literature would most likely be a mix of qualified and unqualified permanent staff.
- Thirdly, the literature review in Chapter 1 showed that the typical training content would more probably include 'hard' physical interventions such as restraint. Hence, whilst care has to be taken in generalising the findings of this study to other sorts of current training programme, there is evidence that its focus on 'soft' skills complemented the evidence already available about restraint skills, and importantly, is set to become increasingly relevant as the 'political' emphasis changes to endorsement of softer skills.
- Fourthly, the combination of three consecutive cohorts created a sample of 243, making it the second largest study found from the international literature review.
- Fifthly, the involvement of the researcher as main lecturer enabled the stability of the Unit content to be ensured for the duration of the study. Subsequent statistical analysis demonstrated the congruity of the three cohorts.

- Sixthly, it was thought that the staggered nature of the study [Cohort One in Trimester three while Cohort Two was in Trimester Two and Cohort One starting Trimester One] would help to highlight any confounding contemporaneous changes in clinical service organisation or clinical contexts.
- Seventh, the study also attempted to monitor the effects of training on a broad range of variables, which included a good number of the learning domains identified by Kraiger *et al* (1993). However, it is conceded that some other studies have attempted a more robust measure of skills development (Collins 1994, Phillips & Rudestram 1995, McGowan *et al* 1999) and/or transfer to the real world by measuring numbers of reported incidents, injuries or costs (Martin 1995, Whittington & Wykes 1996). In the area of skills and competence, this study was restricted to self –assessed measures, although with hindsight a small sample of students could probably have been independently assessed pre- and post- unit, but return to previous placement and monitoring of incident rates was not feasible or appropriate.
- Finally, an attempt was made to interpret changes in these domains using more powerful and convincing inferential statistical analysis [as opposed to raw tally counts or percentages].

In many regards it is felt that the study overcame the deficiencies of previous published work and met the intentions set out at the end of Chapter 1. However there were some difficulties and, on reflection, some areas where perhaps in future things would be done differently. These areas will now be discussed.

6.5.1 Issues With Questionnaire Design

A number of question types were incorporated in the questionnaire and, as discussed above, some of them proved to be extremely time-consuming and labour intensive to analyse. It would be much more straightforward to measure knowledge gain by answers to multiple choice type questions. However, whilst these questions could demonstrate knowledge of risk factors they would not prove information about the absolute number of risk factors identified without prompting, or the anticipated increase in number of risk factors identified following Unit attendance. In addition the multiple choice format tends to rely on prompts rather than un-aided recall, and higher order problem-solving knowledge may be more difficult to assess using multiple-choice questions. Maybe some combination of the two approaches offers the best prospects. Possibly, the inclusion of bullet point icons in the open response boxes would have encouraged the student nurses to list risk factors separately [thus aiding the data analysis and interpretation] rather than sometimes writing passages of text or complete/ incomplete sentences.

The VAS section produced mixed results. Whilst Scenario One produced results that followed the predicted model the results for Scenario Two were apparently inconsistent. This highly regarded question type did produce a fair amount of work in its analysis. Before it was re-used care would be taken to ensure that it targeted a clearly identified aspect of course content.

Many of the Likert-type statement worked very well but those that referred to competence were restricted to self-assessment leaving open the possibility of under- or over-estimation. As previously mentioned, underestimation could result from possible response-shift bias when the respondent changes his or her internal point of reference so that no growth or development is reported, even when a real improvement has occurred.

A similar problem would exist if the respondent were asked to estimate knowledge about a particular area, for example, familiarity with the Law regarding using force against another person. Inclusion of a 'retrospective pre-test' question when completing the post-test would provide additional information about the subjective self-assessment of change. Asking the respondent to rate *on the same scale* both their skill performance or level of knowledge before the intervention and also how it is after the intervention (Arvey & Cole 1991), or estimate current performance relative to any earlier time, say, pre-Unit [better than before, same as before, worse than before etc.] would create additional

opportunities to illustrate change. The former method allows a quantitative comparison to be made while the latter method would allow a qualitative evaluation to be completed.

The risk of over-estimation was also present. The observed gender difference relating to Factor One, and particularly Factor Three scores, wherein males reported themselves to be more confident in their ability to remain safe and have greater practical ability when interacting with aggressive patients, is interesting. Is this a true effect or an over-estimation of confidence and, in particular, competence by males, or a contrasting under-estimation by females?

6.5.2 Questionnaire Administration

In a future evaluation a consistent method of administration would be sought so that, for example, the questionnaire was administered at the same point in the sessions, preferably the beginning, on all occasions. For the purposes of t3, this would mean gaining access to the students in another session as soon after completing the module as possible. Hopefully this would remove any possible administration bias associated with fatigue or students wishing to make a quick getaway.

A further difference between the time t3 and the other three data points – t1, t2, t4 – relates to the possible 'relationship' between the respondents and the lecturer/ researcher. As identified in Chapter 4 every effort was made to distance the lecturer from the research. The

logo and address of the University of Nottingham, Institute of Work, Health And Organisations was the questionnaire header and the work was presented as being organised from there. No relationship bias was suspected at t1 and t2 since the students had never met the lecturer administering the forms prior to those times [or in the intervening period]. Similarly at t4 the students and lecturer had not met in the period between t3 and t4 [and, as far as most students were concerned, would never meet again]. Any possible allegiance effect was likely to be manifested at t3 when the student nurses respondents and lecturer had just spent a good part of a week together, and in particular, half a day involving physical contact practising breakaways.

Possibly, if the students thought that the evaluation of the Unit was, indirectly, an evaluation of the lecturer, then they may have made an extra effort to complete the questionnaire at this time [t3] and identified more risk factors in the scenarios, rated themselves as more confident, competent etc. Effectively, this would be a counter-bias to the fatigue related one discussed above. Maybe, on this occasion it would have been wiser to have the questionnaire administered by an independent person, although this may have simply highlighted the difference and prompted an even greater perceived defence of the lecturer.

6.5.3 Findings

The study revealed several interesting and important findings, many of which have relevance for other training programmes. For example, the

levels of aggression faced by inexperienced staff, the changes that can occur in personal confidence in maintaining one's own safety, belief in predictability and preventability and self-assessed competence and self-respect. The evolution of scores over the medium term in interaction with the clinical environment is also of interest.

Some of the findings raise more questions and indicate a need for further more detailed study. In particular, the level of involvement of student nurses in verbal and physical violence is a priority area. It is regrettable that in this case 'involvement' was not more precisely delineated. The study was about training evaluation and this particular area of questioning was purely to establish that the need for training existed. With hindsight further specification of the nature of that involvement would have proved illuminating, for example, as the main target or focus of verbal or physical attack, one of several targets, actively involved in the verbal de-escalation or physical management/ restraint, summoning assistance, reassuring or relocating other clients etc.

In addition, there are important specific descriptive details -location where incident occurred [setting, type of ward, time of day etc.], assailant details, violence management approach of staff, whether the incident was officially recorded, the consequences for staff and client [injury type, seriousness, absence, time lost, psychological sequelae, effect on intention to remain in training etc.], and whether de-briefing

was offered or accepted. This information urgently requires a large-scale survey and it is anticipated that results would make uncomfortable reading. One practical difficulty that can be anticipated from the indication of the scale of the problem given in this study is the difficulty in accurately recording large amounts of data about many separate incidents, possibly from retrospective recall of events.

The influence of certain demographic variables on scores of confidence, competence, attribution of blame and identification of risk factors etc. is another area for further investigation. Previous training, age grouping, and gender were all seen to affect student's responses. The ability of the Unit to eliminate, at least temporarily, any statistical difference at t3 between students with and without prior training was interesting and personally gratifying. The more so since much of this previous training would have been for a much longer period of time than the three-day Unit, for example five and ten day courses.

Equally intriguing was the absence of any difference in scores associated with chosen speciality branch. Anecdotally, there is a long tradition of mental health and learning disability nurses seeing themselves as different and separate. This might particularly be the case with regard to the area of aggression /violence and its management. Certainly, in these speciality areas the subject has been on the clinical and managerial agendas for several decades, in comparison to adult and children's nursing, where it is a fairly recent

phenomenon. However, this difference failed to make itself statistically apparent in the scores of junior student nurses destined for these different specialities.

6.5.4 Interaction of Factors?

As previously discussed, there is an interaction between aspects of, say, knowledge, attitude and behaviour in the competent performance of a skill, or knowledge, affect and behaviour in the formulation of an attitude. The model of Kraiger *et al* (1993) placed attitudes, confidence and estimations of self-efficacy within the same 'Affective' learning domain. However, the factor analysis apparently distinguished between them and produced separate, unrelated [orthogonal] factors.

On the face of it there appears to be a conflict here. Are attitudes and knowledge, or confidence and competence inter-related or are they separable? At the risk of appearing contradictory the answer is 'both'. It should be remembered that the questionnaire statements were constructed to relate to particular learning outcomes, which in turn attempted to dissect a complex area [health care staff approaches to violence management] into a series of simpler components. Hence learning outcome statements related to knowledge of prediction, self-confidence, practical ability, non-provocative approach etc. whilst at the same time acknowledging that, in reality a non-provocative approach integrates attitudinal, knowledge and behavioural aspects.

It is possible to distinguish between the five factors. The aspects of violence management are recognisably related but still separable. There is a difference between being confident in one's ability to remain safe at work and, say, feeling competent in interacting with agitated persons. One may feel in no danger but still believe that one lacks or has poor skills. Furthermore, there is a difference between feeling confident in one's ability to remain safe and the aspects of self-respect/ staff rights. I may feel safe while at the same time believing that other staff will inevitably get assaulted, for example. Nevertheless, both of these factors might be placed under the heading of attitudes and located with Kraiger *et als'* (1993) affective domain.

As was seen earlier in this Chapter [Section 6.2.5], the interpretation of the Factor scores in relation to the Research Hypotheses was made more difficult because the Research Hypotheses referred to *attitudes*, to *knowledge*, to *skills competence*. Hence it was observed that evidence about one Research Hypothesis was found in relation to more than one factor. This appeared not to be a difficulty when interpreting the factors in relation to violence and its management but was a little awkward sometimes when interpreting the factors in relation to the Research Hypotheses. The problem is difficult to eliminate since one specifies the Hypotheses at the beginning of the research process and only interprets and identifies the factors at a later stage of the data analysis.

6.5.5 Additional Measures of Change

Chapter 1 showed that the number of violent incidents that occurred or were reported in clinical areas is an unreliable indicator. Nevertheless, this has proved to be a popular dependent variable in the training evaluation literature review. However, there are a number of problems with this usage. The large level of under-reporting makes this measure very unreliable. The admission of one seriously disturbed client can dramatically affect the trends regarding the number of reported incidents. It might have merit if a particular clinical team was training *en bloc* and changes in levels of incidents monitored for a long period. More fundamentally, the ethos of a ward or unit has to be established by the senior staff and it seems unrealistic and improbable that the allocation of a student nurse for a short period of time could make any substantial difference to the level of aggression. Thus it was not felt appropriate to include this as a measure of the effects of the training 'transferring to the work' environment.

In any case, there would have been additional practical difficulties in attempting to deliberately organise this for a particular cohort, bearing in mind the specified clinical requirements of the other five cohorts at different stages of training at any time. Student nurses would be unlikely to return to a clinical area where they had been allocated prior to completing the module. Indeed, under normal circumstances it is something that would be actively avoided, a breadth of placements being preferred to depth.

Rating of role-play or designated skills performance by an independent rater against a list of specified skills components or performance criteria, as described by Paterson *et al* (1992) and Phillips & Rudestram (1995), would have been a possible extension to the existing self-estimation of skill competence. However, as previously identified, the logistical difficulties of including this development in the existing training programme were manifold. As a regular training Unit feature the prospect of including even a short period of role-play [15 minutes] with the concomitant repetition associated with division of students in to small groups for health and safety reasons would add at least a day to the course [15 minutes x 20 groups of 5 students = 5 hours]. It would also require the students to attend a study day for 15 minutes of activity, not a recipe for contented students.

6.5.6 Reliability

Attempts were made to calculate some of the psychometric properties of the questionnaire. It could be that some aspects of the questionnaire were inherently 'fluid' or unstable, and thus unsuited to this sort of interpretation. Nevertheless, efforts were made to calculate internal and external reliability coefficients on several occasions. On most occasions these were in the moderate-low categories and rarely met the highest standards set by Nunnally (1978) and others. As has been discussed, with regard to internal consistency, "a scale can be made to look more homogeneous simply by doubling the number of items, even though the average correlation [among the items] remains the same" (Streiner and

Norman 1989:64). More questionnaire statements would have increased the number available for inclusion under each factor and indirectly increased the internal consistency coefficient but this approach is not really the answer.

Strict criteria were set for the factor analysis. High levels of factor loading significance were set [>0.4]. Statements were excluded from the analysis if they cross-loaded at a significant level on more than one factor, whereas other solutions would have meant that they could have been retained on the highest loading factor or remained in both. Both of these measures would have increased the number of statements listed under each factor. In the event, five of the original 24 statements were removed from the analysis and one further statement [No 19] remained but failed to significantly locate under any factor. Thus, this part of the questionnaire was effectively reduced in length by one quarter.

Possibly, on another occasion some allowance for this sort of loss should be anticipated at the questionnaire design and statement generation stage. The inter-rater reliability testing of the approved risk factors for the two Scenarios showed a very high level of agreement and clearly endorsed the initial work to compile the lists and protocols with mutually exclusive categories.

6.5.7 Follow-up

The addition of the follow-up period distinguished this study from many others and revealed interesting post-Unit effects. Some scores seemed

to improve as a result of the subsequent clinical placements and interaction with the real world, for example, the number of risk factors identified in the scenarios, and Factor Three- personal practical ability. Perhaps in these cases the placements allowed the student to observe and discuss risk factors in situations, practice different approaches in a protective environment, or observe staff successfully predicting, anticipating and preventing incidents. Several others seemed to be maintained at broadly similar levels – Factors One, Two, Four and Five and the attribution of blame to the young man in Scenario One. Other scores seemed to be adversely affected by the unfortunate realities of working with busy, over-worked or under-prepared staff, for example, the scores for Statement 15 and the attribution of blame for the elderly woman in Scenario Two.

Whilst the three-month follow-up period added an extra dimension to the data it also begs the questions “for how long is training effective and when is refresher training necessary?”. Once again further research would be required to monitor the patterns of decay [or otherwise] in levels of knowledge, skill etc. Only a handful of studies that was located in the literature review followed up the effects of violence management training on learning domains past the immediate end of course evaluation. Thackrey (1986) demonstrated that increased confidence was still detectable at an 18-month follow-up whilst McGowan *et al* (1999) using the same ‘Confidence in coping with patient aggression’ instrument showed levels maintained at 6-month follow-up. McKenzie

et al (2000) showed that gains in knowledge about challenging behaviour in learning disability were still present at a 12 month follow-up whilst Collins (1994) showed the maintenance of most but not all changes in attitudes at six-month follow-up, following a five day training programme.

A further chronological issue relates to the length of the training Unit. The literature review showed a broad range of training courses ranging from a couple of hours through to 10 days or longer. Three days seemed to be a typical time period but it begs the questions how long should the training take? [Lehmann *et al* (1983) appeared to cover in five hours what many trainers now take 5-10 days to cover], Is there an optimum time? Many courses appear to be for organisationally convenient periods of time – 1 day, 5 days, 10 days – so that staffing can be covered for a day or a week or two etc. Obviously the needs of different staff groups should be borne in mind but, given the implications for scarce staff training budgets, one is entitled to ask whether shorter training courses could produce similar effects to longer ones?

6.6 IMPLICATIONS FOR PRACTICE

Interpretation of findings leads to a number of recommendations for practice and these will now be discussed. Chapter 1 detailed that many published studies were limited in many different respects, including

- poor research design, at best usually pre- and immediate post-training testing

- poor consideration of reliability and other psychometric properties,
- inadequate consideration of possible contemporaneous changes in service organisation or context
- infrequent follow-up of subsequent interaction with clinical practice
- small sample
- limited range of outcomes measures
- emphasis on immediate subjective reaction level or easily obtained but unreliable measures of behaviour change, for example, number of incidents reported
- very little research on student nurses, despite their prominence in the high risk categories of health workers professional groups
- often basic data analysis techniques,

In general terms, and most importantly, it is felt that this study has demonstrated that evaluation research can be performed on aggression management training courses in a manner that overcomes many of the identified weaknesses detailed in Chapter 1. It is most important that trainers in aggression management, line managers, human resource staff, researchers, and staff training personnel are aware of this so that, in future, they can demand the most demonstrably effective training for their organisations and staff. In this way the most appropriate training can be demanded by employers and training commissioners, and made available for staff. More generally, over the longer term, the overall

quality of aggression and violence management training can be improved and assured. A number of other important, more specific points will now be made.

6.6.1 Relevance of Aggression Management Training

The research clearly demonstrated the overwhelming perceived need by student nurses for training of this type. The Questionnaire contained Statement 12 which suggested that "Staff should be educated about the prevention and management of aggressive behaviour as part of their pre-registration preparation". This statement was positively endorsed to such an extent [100% agreement on Day 1 of the Unit] that it lost any ability to discriminate between respondents and so was removed from the factor analysis.

Furthermore, the study also showed that the training is required as soon as possible after the commencement of training. The data revealed a serious problem with incidents of verbal and physical violence involving student nurses, even in the first year of training. It will be recalled that most incidents were reported in mental health and learning disability settings where over half the students reported being verbally abused and almost half reported involvement in physical violence, many on several occasions. Numbers were somewhat smaller but still concerning in adult, and child settings.

Training authorities and clinical services need to take note of this finding otherwise recruitment and retention will become even more difficult. The

likely consequence is that sensible, desirable candidates will consider a safer career elsewhere, and undesirable candidates will be attracted to the perceived danger. Furthermore, if the frequency of involvement in incidents is due in part to the *inexperience* of the students, as is generally asserted, then maybe there is a requirement for all training courses for nurses and *other health care professionals* to include this sort of material at an early point in the course when the students become exposed to clinical situations involving real patients.

6.6.2 Accuracy of the Level of Incident Reporting

A query remains about the level of accurate, official reporting of these incidents. If these figures were generalised to all student cohorts then, using only modest estimations, there would be hundreds of incidents involving student nurses each year [0.5 incidents per student per placement x 3 placements per year x 600 student nurses = 900 incidents].

There is a national requirement to collate the details of officially reported violent incidents with NHS Trusts and notify the Department of Health. Locally committees are established to review these figures and advise staff on their implications. Certainly, to the researchers' knowledge, no clinical manager has ever notified the Department of Nursing and Midwifery of their concern associated with the level of involvement of student nurses in verbal or physical violent incidents.

The obvious conclusions, given the historical levels of under-reporting, are that student nurses are not reporting these incidents, or are not being encouraged to report these incidents, or are being actively discouraged from reporting these incidents. Whichever is the case, the situation is in urgent need of investigation and coordination between higher education departments and clinical services.

6.6.3 Further Training Requirements - Update And Refresher Training

This study has shown that training can have an immediate positive effect on a number of learning domains. Moreover, it has shown that these effects can be detected over the medium term [three months] and that further changes occur on subsequent interaction with staff and patients in clinical placements [either increases or decreases] over time.

However, in many areas of motor skills performance, for example basic life support training, there is an acknowledged decay in the competence with which skills are performed and a concomitant requirement for period [annual] update. This effect is noted especially if the skills are performed infrequently, as thankfully are violence management skills in many settings. Indeed, one of the earliest studies located in the literature review prophesied the need for regular practice update and refreshers ((Gertz 1980) and this need has been repeated subsequently (Thackrey 1987, Goodykoontz & Herrick 1990).

Hence attendance on an aggression management training course should not be viewed as a one-off event but the start of a series of regular training. This requirement places an extra commitment on Departments of Nursing and training organisations to create resources and curriculum time to provide period update and refresher training. There is also a growing burden on trainers/lecturers to facilitate not only the initial training but also the growing amount of refresher/update sessions. Once again this requirement has cost implications that needs to be planned and managed if lecturers are not to be 'swamped' and, consequently, prevented from teaching on other courses.

An obvious administrative solution would be to group the aggression/violence management update training with other health and safety training, for example, manual handling, and fire prevention training etc. and incorporate them it in to an annual update programme. This package could then be delivered in years One, Two and Three of the student nurse training programme.

6.6.4 Usefulness and Importance of Evaluation Research

The study has shown the changes in several learning domains that occur as the result of attending aggression management training. In this way it has proved to be a powerful supplement to an earlier study of the same training course by the researcher that sought more qualitative data about the usefulness, level of interest and perceived relevance of the Unit and its material. Hopefully, this can act as a spur to other

trainers and encouragement to complete evaluation of their own training. Alternatively, it could stimulate trainers, managers or training officers within NHS Trusts to commission independent evaluation of the training that they purchase or provide for their staff.

As stated in Chapter 1, evaluation is firstly important because of the enormous financial costs associated with this training. Equally, managers need to be reassured that the chosen training or training organisation is meeting the particular identified needs of a staff group or service orientation, ideally expressed as a series of learning outcomes involving aspects of knowledge, attitudes, skills and procedures. Too often this is not the case, and direction is lacking, resulting in ready-made training being blindly purchased and the course content left to the, quite possibly, inappropriate, experience or preference of the trainer. Although the National Audit Office survey (NAO 2003:24) found some evidence that “some NHS trusts have adopted bespoke training courses which they believe have led to improvements in prevention”.

Paterson *et al* (1992:374) asserted that trainers or ‘educationalists’ who provide this type of training “should be obliged to examine the quality of the training they are offering” and, moreover, continued in suggesting that “there is a professional obligation to thoroughly evaluate its effects in order that the best possible instruction is given to staff to enable them to act in the best interests of their clients”. The research has also demonstrated that this requirement to evaluate training need not be an

excessively onerous task. Some forms of data collection can routinely produce evidence of change in a timely and resource efficient manner. Alternatively, the task could be contracted out to independent training evaluators. Either way, the case put forward by Paterson *et al* (1992) become ever more compelling.

Trainers should not be able to dismiss these findings or implications simply because they believe that their course is 'different', or at a higher level, or centres on restraint. Admittedly, the Unit that was studied did not include restraint skills training but had many components that would also be present on other training courses. Hence, it is thought that there is much to learn from its findings.

6.6.5 Low-Level, Soft-Skills Training

The Unit that was studied included skills in de-escalation and breakaway/personal safety but not physical restraint. This effectively located the training within the category of softer skills training. The results of this study have been given greater relevance by several recent publications that have either criticised the excessive focus on 'hard' restraint skills at the expense of 'soft' training for example, de-escalation, customer-care skills or civility training or apologised for the 'misdirected guidance' that has been previously offered.

A National Audit Office survey (NAO 2003:24) highlighted the excessive focus on hard skills and the relatively infrequent inclusion of softer skills

raising concerns about “the number of trusts who fail to make training in situation risk assessment and customer care compulsory” skills in training programmes. More recently the inquiry into the death of David ‘Rocky’ Bennett concluded that, whilst psychiatric care staff may receive adequate training in efficient restraint skills, they are missing out on other skills training, such as attitudinal improvement on racial and cultural needs of clients, and low level diffusion and customer care skills.

Most recently the Nursing and Midwifery Council has commented on the findings of this inquiry in relation to aggression management training. The comments of Rick Tucker, NMC professional advisor on mental health, were presented early in this Chapter but are worthy of repetition here. Rick Tucker, who was also an expert witness at the public inquiry said “there has been an over-emphasis on how to physically manage incidents of violence at the expense of recognition and prevention. This has led to a defensive and reactive culture that has de-skilled many practitioners. We believe the current strategy is failing both patients and carers” (NMC 2004a:7).

Thus, this study is timely and has the potential to influence future training practice since it focused on de-escalation, on non-provocative approaches, on risk assessment, on multi-factorial causation of aggression, on knowing one’s limitations and stopped short of development of skills in physical restraint. It demonstrates the potential

of aggression management training to develop or change knowledge, attitudes, skills etc.

This statement is not intended to be an overt criticism of physical restraint training. The author believes that restraint is needed in some situations but is of the opinion that premature recourse to physical restraint is provocative. Moreover, it is also felt that merely having physical restraint skills may cognitively pre-determine the likely management strategy and bias this towards a 'hard' physical response rather than a 'softer' psychological, inter-personal de-escalation one. It is felt that this process is being referred to when the NMC talks about effectively 'de-skilling' practitioners in the quotation above. The effect has been called the 'Law of the Hammer' and will be familiar to homeowners new to D-I-Y. Buy a hammer and subsequently every DIY job requires a nail rather than the relative finesse of a screwdriver and screw.

Incidentally, the statement can also be viewed as a vindication of the lecturer's stance in refusing to incorporate aspects of restraint into the Unit, despite a steady demand by student nurses to do so. It is believed that, at this premature point in their career, they primarily need to be able to recognise the dangers in situations as they developed [hence the inclusion of risk factors and risk assessment], and feel safe [hence the inclusion of breakaway training]. They also need to recognise their own limitations and be guided by permanent trained staff. In the real

world, where all members of ward and unit teams are not trained, student nurses could then find themselves leading a ward team in the management of an incident, and this was not felt to be defensible.

6.6.6 Utility of Measures

The study has shown that training effects can be accurately demonstrated using a range of different question types. Some approaches lend themselves fairly easily to measurement, and analysis whilst others are extremely labour-intensive. Some seems to produce consistent, robust, reliable, and convincing evidence of change whilst others are inconclusive.

On the basis of the evidence from this study, and the experience of completing it, the researcher believes that the use of Likert-type statements should be recommended. Obviously, time is required initially to carefully create Likert statements that reflect particular learning outcomes relating to preferred knowledge, attitudes or motor skill approaches. However, subsequently data input is straightforward. Analysis and interpretation can be performed at varying levels of sophistication, ranging from basic interpretation in changes of percentage agreement or disagreement with each statement or pre-designated sub-group of statements, through to factor analysis and subsequent inferential statistical analysis of factor scores.

Multiple choice questions or short-answer type questions could be used to measure the acquisition of knowledge and, once again, lend themselves to straightforward marking. The use of open response scenario-type questions, as used in this study, requires both a great deal more preparatory work and time-consuming subsequent analysis. As was shown, they do hold the potential to illuminate the more subtle changes in areas of knowledge, and also measure the higher levels of knowledge, for example, knowledge organisation, mental models and cognitive strategies for problem-solving referred to by Kraiger *et al* (1993) but are very labour intensive. To be fair, they could not be recommended for routine, in-going, routine evaluation of training.

The use of VAS falls between the two extremes. The questions were easy to create and fairly straight-forward and quick [if tedious] to interpret and input. However, the results were inconclusive and contradictory, and for this reason their use would be cautiously suggested. Perhaps further investigation can clarify the strengths and limitations of this approach and advise on the areas where it is can be most effectively employed.

6.7 CONCLUDING REMARKS

Sadly, one thing is certain. Health professionals will be distracted from their work of caring and curing by aggression and violence for the

foreseeable future. Chapter 1 concluded with lists of Research Questions and Hypotheses designed to measure outcomes and ascertain whether aggression/ violence/conflict management-type training 'worked'. Answers to these questions were sought because of the inadequate evidence available in the literature. A secondary broad concern centred on the utility of training evaluation models, particularly the recent work on Kraiger *et al.*

It is believed that this study maintained a technical rigour that overcame many of the reported deficits in previous published studies. It has shown that valuable evaluation research can be performed on pre-existing training courses, using carefully designed data collection methods, and provided evidence that training does 'work' to a large extent. Some types of question have been shown to be more easily administered, more sensitive, or more easily evaluated than others. All have provided valuable evidence about the immediate and medium-term outcomes of a relatively short training programme that contained many typical training components, which should be of interest to trainers, other researchers, and training officers/ commissioners. Given the anticipated pendulum swing back to an emphasis on 'soft' skills it is felt that these results may hold special significance.

The use of Kraiger *et al's* model was useful in creating a questionnaire that allowed many cognitive, skills and affective aspects to be monitored. It wasn't perfect and it did cause some confusion, for

example, when analysing some of the factor analysis results. Equally, it has to be conceded that absence of Kirkpatrick's reaction level means it is not known whether any student found the Unit interesting, useful, enjoyable, relevant etc., or whether they had suggestions for 'process issues', for example, different teaching methods, sequencing of material. Hence, as discussed, a combination of models, as outlined in Chapter 2, probably provides the greatest scope.

It is sincerely hoped that other trainers/ researchers can take heart from this document, its findings, and the efforts of respondents and researchers that have gone into obtaining them, and strive to develop the ideas and techniques contained within it by application to other examples of training.

REFERENCES

- Allen D., McDonald L., Dunn C., & Doyle T. (1997) Changing care staff approaches to the prevention and management of aggressive behaviour in a residential treatment unit for persons with mental retardation and challenging behaviour. *Research in Developmental Disabilities*, 18(2), 101-112.
- Allen D. & Tynan H. (2000) Responding to aggressive behaviour: Impact of training on staff members' knowledge and confidence. *Mental Retardation*, 38(2), 97-104.
- Allen D. (2001) *Training Carers in Physical Interventions: Research Towards Evidence-Based Practice*. Kidderminster : British Institute of Learning Disabilities.
- Arnetz J. (1998) The violent incident form (VIF): a practical instrument for the registration of violent incidents in the health care workplace. *Work and Stress*, 12(1), 17-28.
- Arnetz J.E. & Arnetz B.B. (2000) Implementation and evaluation of a practical intervention programme for dealing with violence towards health care workers. *Journal of Advanced Nursing*, 31(3), 668-680.
- Arvey R.D., & Cole D.A. (1991) Evaluating change due to training. In I.L. Goldstein and Associates (Eds.) *Training and Development in Organisations*. San Francisco : Jossey-Bass. Ch 4 pp89-117.
- Baker P.A. & Bissmire D. (2000) A pilot study of the use of physical intervention in the crisis management of people with intellectual

disabilities who present challenging behaviour. *Journal of Applied Research in Intellectual Disabilities*, 13, 38-45.

Bandura A. (1977) Self-efficacy: Towards a unifying theory of behavioral change. *Psychological Review*, 84, 191-215. Cited by Kraiger *et al* 1993.

Baron R.A. & Neuman J.H. (1998) Workplace aggression- the iceberg beneath the tip of workplace violence: evidence on its forms, frequency, and targets. *Public Administration Quarterly*, 21(4), 446-464.

Beale D., Fletcher B., Leather P., & Cox T. (1998) *Review on Violence to NHS Staff Working in the Community*. Nottingham : Centre for Organisational Health and Development, University of Nottingham. Commissioned by NHS Executive and Royal College of Nursing.

Beauchamp T.L. & Childress J.F. (1994) *Principles of Biomedical Ethics*. [4th Ed] , New York, Oxford University Press.

Bee F. & Bee R. (1994) *Training Needs Analysis and Evaluation*. London : Institute of Personnel and Development.

Beech B. (1999) Sign of the times or the shape of things to come? A 3 day unit of instruction on 'aggression and violence in health settings' for all students during pre-registration nurse training *Nurse Education Today*, 8, 610-616.

Beech B. (2001) Managing aggression and violence in care settings: a review of the legal and ethical content of staff training courses. *The Journal of Adult Protection*, 3(2), 8-17.

- Bensley L., Nelson N., Kaufman J., Silverstein B. & Walker Shields J. (1995)
Patient and staff views of factors influencing assaults on psychiatric
hospital employees. *Issues in Mental Health Nursing*, 16, 433-446.
- Benson A., Secker J., Balfe E., Lipsedge M., Robinson S. & Walker J. (2003)
Discourses of blame: accounting for aggression and violence on an
acute mental health inpatient unit. *Social Science and Medicine*, 57,
917-926.
- Berk R.A. & Rossi P.H. (1990) *Thinking about Program Evaluation*. Newbury
Park, CA : Sage
- Bloom B. (1956) *Taxonomy of Educational Objectives: The Cognitive Domain*.
New York : Donald McKay. Cited in Kraiger *et al* 1993.
- Bowling A. (1997) *Research Methods in Health: Investigating Health and
Health Services*. Buckingham, Open University Press.
- Bowie V. (2000) Current trends and emerging issues in workplace violence.
Security Journal,, 13(3), 7-23.
- Brantley P.J. & Bruce B.K. (1986) Assessment in Behavioral Medicine In A.R.
Ciminero, K.S. Calhoun & H.E. Adams (Eds.) *Handbook of Behavioural
Assessment*. New York : John Wiley and Sons. Ch 19 pp 673-721.
- Breakwell G. (1997) *Coping with Aggressive Behaviour*. Leicester, The British
Psychological Society.
- Brewer S. (1999) Commentary: The cost of violence to healthcare workers. *NT
Research*, 4(2), 116-118.
- Briggs M. & Closs J.S. (1999) A descriptive study of the use of visual analogue
scales and verbal rating scales for the assessment of post-operative

pain in orthopaedic patients. *Journal of Pain and Symptom Management*, 18(6), 438-446.

Bryman A. & Cramer D. (1997) *Quantitative Data Analysis with SPSS for Windows: a guide for social scientists*. London, Routledge.

Budd T. (1999) *Violence at Work: Findings from the British Crime Survey*. London, Health and Safety Executive.

Burns R.L. (2000) *Introduction to Research Methods*. London, Sage.

Calabro K., Mackey T.A., & Williams S. (2002) Evaluation of training designed to prevent and manage patient violence. *Issues in Mental Health Nursing*, 23, 3-15.

Californian Occupational Safety and Health Administration [Cal/OSHA] (1995) *Guidelines for Workplace Security*. San Francisco, CA. : Division of Occupational Safety and Health, State of California Department of Industrial Relations.

Campbell D.T. (1969) 'Reforms as experiments', *American Psychologist*, 14: 404-29.

Campbell D.T. & Stanley J.C. (1963) *Experimental and Quasi-experimental Designs for Research*. Chicago : Rand McNally. Cited by Robson 1993.

Campbell D. & Stanley J. (1966) *Experimental and Quasi-Experimental Evaluations in Social Research*. Chicago : Rand McNally.

Carmel H. & Hunter M. (1990) Compliance with training in managing assaultive behaviour and injuries from inpatient violence. *Hospital and Community Psychiatry*, 41(5), 558-560.

- Cella D.F. & Perry S.W. (1986) Reliability and concurrent validity of three visual analogue mood scales. *Psychological Reports*, 59, 827-833.
- Cembrowicz S. & Ritter S. (1994) Attacks on doctors and nurses. In J. Shepherd (Ed.) *Violence in Health Care*. Oxford: Oxford University Press. Chapter 2 pp 13-41.
- Chappell D. & Di Martino V. (2000) *Violence at Work*. (2nd Ed.) Geneva : International Labour Organisation.
- Child D. (1990) *The Essentials of Factor Analysis* [2nd Ed] London, Cassell.
- Choiniere M. & Amsel R. (1996) A visual analogue thermometer for measuring pain intensity. *Journal of Pain and Symptom Management*, 11(5), 299-311.
- Clark E. & McCann T.V. (2005) Researching students: an ethical dilemma. *Nurse Researcher*, 12(3), 42-51.
- Clarke A. (1999) *Evaluation Research: An Introduction to Principles, Methods and Practice*. London : Sage.
- Cline M.E., Herman J., Shaw E.R., & Morton R.D. (1992) Standardisation of the Visual Analogue. *Scale Nursing Research*, 41(6), 378-380.
- Collins J. (1994) Nurses attitudes towards aggressive behaviour, following attendance at 'The Prevention and Management of Aggressive Behaviour Programme'. *Journal of Advanced Nursing*, 20, 117-31.
- Comrey A.L. (1978) Common methodological problems in factor analysis studies. *Journal of Consulting and Clinical Psychology* 46, 4, 648-659.
- Cook T.D. & Campbell D.T. (1979) *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago :Rand McNally. Cited by Robson 1993.

- Cox T. & Ferguson E. (1994) Measures of the subjective work environment. *Work and Stress*, 8, 98-109. Cited by Leather *et al* (1998).
- Cox T. & Leather P. (1994) The prevention of violence at work: application of a cognitive behavioural theory. In: C.L. Cooper & I.T. Robertson (Eds.) *International Review of Industrial and Organisational Psychology*, volume 9. Chichester, Wiley and Sons. pp 213-245.
- Crichton J.H.M., Callanan T.S., Beauchamp L., Giasson M. & Tardiff H. (1998) Staff response to psychiatric inpatient violence: an international comparison. *Psychiatric Care*, 5(2), 50-56.
- Department of Health (1999a) *We Don't Have to Take This*. NHS Zero Tolerance Zone Campaign London: Stationary Office.
- Department of Health (1999b) *Working Together: Managing Violence, Accidents and Sickness Absence in the NHS*. London: Stationary Office. (Health Service Circular 1999/229).
- Department of Health (1999c) *Working Together - Securing a Quality Workforce for the NHS*. London: Stationary Office. HSC 1999/079.
- Department of Health and Social Security (1976) *The Management of Violent, or Potential Violent, Hospital Patients*. HC(76)11. London, HMSO.
- Dickson R., Cox T., Leather P., Beale D., & Farnsworth B. (1993) Violence at work. *Occupational Health Review*, 46, Nov/Dec, 22-24.
- Dickson R., Leather P., Beale D., & Cox T. (1994) Intervention strategies to manage workplace violence. *Occupational Health Review*, 50, July/August, 15-18.
- Dimond B. (1990) *Legal Aspects of Nursing*. London : Prentice Hall.

- Dipboye R.L. (1997) Organizational barriers to implementing a rational model of training. In M.A. Quinones & A. Ehrenstein (Eds.) *Training for a Rapidly Changing Workplace: Applications of Psychological Research*. Washington DC : American Psychological Association. Ch 2 pp31-60.
- Duxbury J. & Whittington R. (2005) Causes and management of patient aggression and violence: staff and patient perspectives. *Journal of Advanced Nursing*, 50(5), 469-478.
- Easterby-Smith M. (1994) *Evaluation of Management Education, Training and Development*. (2nd edn). Aldershot : Gower. Cited by Thackwray (1997).
- Elliott P.P. (1997) Violence in health care: what nurse managers need to know. *Nursing Management*, December, 28(12), 38-42.
- Employment National Training Organisation (2003) *National Occupational Standards in Managing Work-related Violence*.
<http://www.empnto.co.uk/projects/wrv/index.htm> Accessed 16th January 2004.
- English National Board for Nursing, Midwifery and Health Visiting [ENB] (1993) *Coping with Violence and Aggression- Prevention and Management*. Circular 1993/11/GMB . London: ENB.
- English National Board for Nursing, Midwifery and Health Visiting [ENB] (2001) *Annual Report for 2000-01*. London, ENB.
- European Foundation for the Improvement of Living and Working Conditions (2000) *Third European Survey on Working Conditions*. Dublin, Ireland: European Foundation. Cited by Hoel *et al* (2001).
- Evaluation Research Society (1980) *Standards for Evaluation*. Washington, DC : Evaluation Research Society.

- Farrell G.A. (1997) Aggression in clinical settings: nurses' views. *Journal of Advanced Nursing*, 25, 501-508.
- Farrell G.A. (1999) Aggression in clinical settings: nurses' views – a follow-up study. *Journal of Advanced Nursing*, 29(3), 532-541.
- Farrell G.A. (2001) From tall poppies to squashed weeds*; why don't nurses pull together more? *Journal of Advanced Nursing*, 35(1), 26-33.
- Ferguson E. & Cox T. (1993) Exploratory factor analysis: a users' guide. *International Journal of Selection and Assessment*. 1(2), 84-94.
- Fidderman H. (2000) Violence in the NHS -the trade union viewpoint. *Health Service Report*, Winter 21-24.
- Field A . (2000) *Discovering Statistics using SPSS for Windows*. London, Sage.
- Fife-Schaw C. (2000a) Quasi-experimental designs. In G.M. Breakwell, S. Hammond & C. Fife-Schaw (Eds.) *Research Methods in Psychology*. [2nd Ed] London : Sage. Ch 6 pp75-87.
- Fife-Schaw C. (2000b) Questionnaire design. In G.M. Breakwell, S. Hammond & C. Fife-Schaw (Eds.) *Research Methods in Psychology*. [2nd Ed] London : Sage. Ch 12 pp158-174.
- Fife-Schaw C. (2000c) Bivariate statistical analyses. In G.M. Breakwell , S. Hammond & C. Fife-Schaw (Eds.) *Research Methods in Psychology*. [2nd Ed] London : Sage. Ch 25, pp350-371.
- Finnema E.J., Dassen T. & Halfens R. (1994) Aggression in psychiatry: a qualitative study focusing on the characterisation and perception of patient aggression by nurses working on a psychiatric ward. *Journal of Advanced Nursing*, 19, 1088-1095.

- Fitz-Gibbon C.T. & Lyons Morris L. (1987) *How to Design a Program Evaluation*. Newbury Park, CA : Sage.
- Flannery R.B. (1996) Violence in the workplace, 1970-1995: a review of the literature. *Aggression and Violent Behavior*, 1(1), 55-68.
- Flannery R.B., Hanson M.A., Penk W.E., Goldfinger S., Pastva G.J., & Navon M.A. (1998) Replicated declines in assault rates after implementation of the assaulted staff action program. *Psychiatric Services*, 49(2), 241-243.
- Fletcher T.A., Brackel S. J., & Cavanaugh J.L. (2000) Violence in the workplace: new perspectives in forensic mental health services in the USA. *British Journal of Psychiatry*, 176, 339-344.
- Ford J.D. (1979) Research on Training Counsellors and Clinicians. *Review of Educational Research*, 69, 87-130. Cited in Milne D. (2000) *Evaluation in Mental Health Nursing*. In R Newall & K Gournay (Eds.) *Mental Health Nursing: an Evidence Based Approach*. Edinburgh : Churchill Livingstone. Ch 8 pp121-143.
- Forster P.L., Cavness C., & Phelps M.A. (1999) Staff training decreases use of seclusion and restraint in an acute psychiatric hospital. *Archives of Psychiatric Nursing*, 8(5), 269-271.
- Fox B., Polkey C., & Boatman P. (2002) *Tolley's Managing Violence in the Workplace*. Croydon : Tolley, Lexis, Nexis.
- Gagne R.M. (1984) Learning outcomes and their effects; Useful categories of human performance. *American Psychologist*, 39, 377-385. Cited in Kraiger and Jung (1997).

- Gentry M., Icton J., & Milne D. (2001) Managing challenging behaviour in the community: methods and results of interactive staff training. *Health and Social Care in the Community*, 9(3), 143-150.
- Gertz B. (1980) Training for prevention of assaultive behaviour in a psychiatric setting. *Hospital and Community Psychiatry*, 31(9), 628-30.
- Ghodsian D., Kjork R.A., & Benjamin A.S. (1997) Evaluating training during training: obstacles and opportunities. In M.A. Quinones & A. Ehrenstein (Eds.) *Training for a Rapidly Changing Workplace: Applications of Psychological Research*. Washington DC : American Psychological Association. Ch 3 pp63-88.
- Gibbon B. (1995) Validity and reliability of assessment tools. *Nurse Researcher*, 2(4), 48-55.
- Goldstein I.L. (1993) *Training in Organisations: Needs assessment, development and evaluation*. Monterey, CA, Brooks/Cole. Cited in Dipboye R.L. (1997) Organizational barriers to implementing a rational model of training. In M.A. Quinones & A. Ehrenstein (Eds.) *Training for a Rapidly Changing Workplace: Applications of Psychological Research*. Washington DC : American Psychological Association. Ch 2 pp31-60.
- Goodykoontz L. & Herrick C.A. (1990) Evaluation of an in-service education program regarding aggressive behaviour on a psychiatric unit. *Journal of Continuing Education Nursing*, 21(3), 129-33.
- Gorsuch R.L. (1978) *Factor Analysis* Philadelphia: W.B. Saunders Co.

- Gostin L. (1986) *Institutions Observed: Towards a New Concept of Secure Provision in Mental Health*. London : King Edward's Hospital Fund for London.
- Guba E.G. & Lincoln Y.S. (1989) *Fourth Generation Evaluation*. Newbury Park, Ca.: Sage
- Hagen B.F. & Sayers D. (1995) When caring leaves bruise: The effects of staff education on resident aggression. *Journal of Gerontological Nursing*, 21(11), 7-16.
- Hamblin A.C.(1974) *Evaluation and Control of Training*. Maidenhead : McGraw-Hill.
- Hammond S. (2000) Using psychometric tests. In G.M. Breakwell , S. Hammond & C. Fife-Schaw (Eds.) *Research Methods in Psychology*. [2nd Ed] London : Sage. Ch 13, pp175-3193.
- Harris J. (1996) Physical restraint procedures for managing challenging behaviours presented by mentally retarded adults and children. *Research in Developmental Difficulties*, 17(2), 99-134.
- Hatch-Maillette M.A. & Scalora M.J. (2002) Gender, sexual harassment, workplace violence, and risk assessment: convergence around psychiatric staff's perceptions of personal safety. *Aggression and Violent Behaviour*, 7, 271-291.
- Health Services Advisory Committee (1987) *Violence to Staff in the Health Services*. Sudbury, Suffolk : HSE Books.
- Health Services Advisory Committee (1997) *Violence and Aggression to Staff in Health Services: Guidance on Assessment and Management*. Sudbury, Suffolk : HSE Books.

- Herman J.L., Morris L.L., & Fitz-Gibbon C.T. (1987) *Evaluator's Handbook*.
Newbury Park, CA : Sage.
- Hesketh K.L., Duncan S.M., Estabrooks C.A., Reimer M.A., Giovannetti P.,
Hyndman K., & Acorn S. (2003) Workplace violence in Alberta and
British Columbia hospitals. *Health Policy*, 63, 311-321.
- Hinkle D.E., Wiersma W., & Jurs S.G. (2003) *Applied Statistics for the
Behavioural Sciences* (5th Edn.). Boston : Houghton Mifflin.
- Hoel H., Sparks K., & Cooper C.L. (2001) *The Cost of Violence/Stress at Work
and the Benefits of a Violence/Stress-Free Working Environment*.
Report commissioned by the International Labour Organisation.
Manchester : UMIST.
- Howe T. (1995) Measurement scales in healthcare settings. *Nurse
Researcher*, 2(4), 30-37.
- Huang H.Y., Wilkie D.J., & Berry D.L. (1996) Use of a computerized digitizer
tablet to score and enter visual analogue scale data. *Nursing Research*,
Nov/Dec 45(6) 370-372.
- Hurlebaus A. (1994) Aggressive behaviour management for nurses: an
international issue. *Journal of Healthcare Protection Management*,
10(2), 97-106.
- Hurlebaus A.E. & Link S. (1997) The effects of an aggressive behavior
management program on nurses' levels of knowledge, confidence and
safety. *Journal of Nursing Staff Development*, Sept/Oct 13(5), 260-265.
- Ilkiw-Lavalle O., Grenyer B.F.S., & Graham L. (2002) Does prior training and
staff occupation influence knowledge acquisition from an aggression

management training program? *International Journal of Mental Health Nursing*, 11, 233-239.

Ilkiw-Lavalle O. & Grenyer B.F.S. (2003) Differences between patient and staff perceptions of aggression in mental health units. *Psychiatric Services*, 54(3), 389-393.

Industrial Relations Services (1998) Stamping out violence in the NHS: a survey of policy provision. *Health Service Report*, Summer 19, 2-14.

Industrial Relations Services (2000) Getting to grips with workplace violence: a snapshot survey *Health Service Report*, Winter 29, 2-13.

Industrial Society (1994) *Managing Best Practice Survey*. London : The Industrial Society.

Infantino J.A. & Musingo S-Y. (1985) Assaults and injuries among staff with and without training in aggression control techniques. *Hospital and Community Psychiatry*, 36(12), 1312-1314.

International Council of Nurses (1999) *Guidelines on Coping with Violence in the Workplace*. Geneva : International Council of Nurses.

Jackson D., Clare J., & Mannix J. (2002) Who would want to be a nurse? Violence in the workplace – a factor in recruitment and retention. *Journal of Nursing Management*, 10, 13-20.

Jenkins E.L. (1996) *Violence in the workplace: Risk-factors and prevention strategies*. National Institute for Occupational Safety and Health (NIOSH). Publication No 96-100. Washington DC : US Government Printing Office. Cited by Hoel *et al* (2001)

- Kirkpatrick D.L. (1958/9) Techniques for evaluating training programmes.
Journal for the American Society of Training Directors. 13 (November)
 and 14 (February)
- Kirkpatrick D.L. (1976) Evaluation of training. In R.L. Craig (Ed.) *Training and Development Handbook: A Guide to Human Resource Development*.
 (2nd Edn.) New York : McGraw-Hill. pp18-1 – 18-27.
- Kirkpatrick D.L. (1994) *Evaluating Training Programs*. San Francisco, CA :
 Berrett-Koehler.
- Kraiger K., Ford J.K., & Salas E. (1993) Application of cognitive, skill-based,
 and affective theories of learning outcomes to new methods of training
 evaluation. *Journal of Applied Psychology*, 78(2), 311-328.
- Kraiger K. & Jung K.M. (1997) Linking training objectives to evaluation criteria.
 Ch 6 in M.A. Quinones & A. Ehrenstein (eds) *Training for a Rapidly
 Changing Workplace: Applications of Psychological Research*.
 Washington DC. American Psychological Association. Pp151-175
- Krathwohl D.R., Bloom B.S., & Masia B.B. (1964) *Taxonomy of Educational
 Objectives: The Classification of Educational Goals*. White Plains, N.Y.
 : Longman.
- Kumar R. (1999) *Research Methodology: A Step by Step Guide for Beginners*.
 London : Sage.
- Lanza M.L. & Kayne H.L. (1995) Patient assault: A comparison of patient and
 staff perceptions. *Issues in Mental Health Nursing*, 16, 129-141.
- Leather P., McKeachie S. & Amirkhanian M. (1994) The importance of likeability
 as measure of television advertising effectiveness. *International Journal
 of Advertising* 13(3), 265-280.

- Leather P., Beale D., Lawrence C., & Dickson R. (1997) Effects of exposure to occupational violence and the mediating impact of fear. *Work and Stress*, 11(4), 329-340.
- Leather P., Brady C., Lawrence C., Beale D., & Cox T. (Eds.) (1998a) *Work Related Violence: Assessment and Intervention*. London : Routledge.
- Leather P., Pyrgas M., Beale D., & Lawrence C. (1998b) Windows in the workplace: Sunlight, view and occupational stress. *Environment and Behavior*, 30(6), 739-762.
- Lee L.J. & Sampson J.F. (1990) A practical approach to program evaluation. *Evaluation and Program Planning*, 13(2), 157-164.
- Lee S., Wright S., Sayer J., Parr A-M., Gray R., & Gournay K. (2001) Physical restraint training for nurses in English and Welsh psychiatric intensive care and regional secure unit. *Journal of Mental Health*, 10(2), 151-162.
- Leggett J. & Silvester J. (2003) Care staff attributions for violent incidents involving male and female patients: a field study. *British Journal of Clinical Psychology*, 42, 393-406.
- Lehmann L.S., Padilla M., Clark S., & Loucks S. (1983) Training personnel in the prevention and management of violent behavior. *Hospital and Community Psychiatry*, 34(1), 40-43.
- Lingjaerde O. & Foreland A.R. (1998) Direct assessment of improvement in winter depression with a visual analogue scale: high reliability and validity. *Psychiatry Research*, 81, 387-392.
- Lion J.R., Snyder W., & Merrill G.L. (1981) Under-reporting of assaults on staff in a state hospital. *Hospital and Community Psychiatry*, 32, 497-498.

- Love C.C. & Hunter M.E. (1996) Violence in public sector psychiatric hospitals: benchmarking nursing staff injury rates. *Journal of Psychosocial Nursing*, 34(5), 30-34.
- Lyon C. (1994) *Legal Issues Arising from the Care, Control, and Safety of Children with Learning Disabilities who also Present Severe Challenging Behaviour*. London : Mental Health Foundation. Cited by Paterson et al (1997).
- MacKay C. (1994) Violence to health care professionals: a health and safety perspective. In T. Wykes (Ed.) *Violence and Health Care Professionals*. London : Chapman & Hall. Ch 1, pp 9-22.
- Martin A. (1990) The case for self defence. *Health and Social Services, Journal*. 88, 697.
- Martin K. (1995) Improving staff safety through an aggression management program. *Archives of Psychiatric Nursing*, IX(4), 211-215.
- Martinson R. (1974) 'What works? Questions and answers about prison reform' *Public Interest*, 35:22-45.
- Maxfield M.C., Lewis R.E., & Cannon S. (1996) Training staff to prevent aggressive behaviour of cognitively impaired elderly patients during bathing and grooming. *Journal of Gerontological Nursing*, 22(1), 37-43.
- Mayhew C. & Chappell D. (2002) An overview of occupational violence *Australian Nursing Journal*, 9(7), 34-35.
- McColl E. (1993) Questionnaire design and construction. *Nurse Researcher*, 1(2), 16-25.
- McColl E., Jacoby A., Thomas L., Soutter J., Bamford C., Steen N., Thomas R., Harvey E., Garatt A & Bond J (2001) Design and use of

- questionnaires: a review of best practice applicable to surveys of health service staff and patients. *Health Technology Assessment*, 5(31).
- McDonnell A. (1997) Training care staff to manage challenging behaviour: an evaluation of a three day training course. *The British Journal of Developmental Disabilities*, 43(2.85), 156-162.
- McDonnell A., McEvoy J., & Dearden R.L. (1994) Coping with violent situations in the caring environment. In T. Wykes (Ed.) *Violence and Health Care Professionals*. London : Chapman & Hall. Ch 10 pp189-206.
- McDowell I. & Newall C. (1987) *Measuring Health: A Guide to Rating Scales and Questionnaires*. Oxford : Oxford University Press.
- McGowan S., Wynaden D., Harding N., Yassine A., & Parker J. (1999) Staff confidence in dealing with aggressive patients: A benchmarking exercise. *Australian and New Zealand Journal of Mental Health Nursing*, 8, 104-108.
- McKechie S., Leather P. & Ozuygun S. (1997) Viewers' reactions to financial services commercials. *Journal of Financial Services Marketing* 1(4), 375-388.
- McKenzie K., Paxton D., Patrick S., Matheson E., & Murray G.C. (2000) An evaluation of the impact of a one-day challenging behaviour course on the knowledge of health and social care staff working in learning disability services. *Journal of Learning Disabilities*, 4(2), 153-165.
- Mezey G. & Shepherd J. (1994) Effects of assault on health-care professionals. In: J. Shepherd (Ed.) *Violence in Health Care: A Practical*

- Guide to Coping with Violence and Caring for Victims*. Oxford : Oxford University Press. Chapter 1 pp 1-11.
- National Audit Office (1996) *Health and Safety in the NHS Acute Hospital Trusts in England*. London : The Stationary Office.
- National Audit Office (2001) *Educating and Training the Future Health Professional Workforce for England*. London : The Stationary Office.
- National Audit Office (2003) *A Safer Place to Work: Protecting NHS Hospital and Ambulance Staff from Violence and Aggression*. London : The Stationary Office.
- National Institute for Occupational Safety and Health (2002) *Violence: Occupational Hazards in Hospitals*. Cincinnati, OH : NIOSH.
- Nolan P. (1990) Psychiatric nursing – the first 100 years. *Senior Nurse*, 10(10), 20-23.
- Nolan P. (1993) *A History of Mental Health Nursing*. Cheltenham : Stanley Thornes.
- Nolan P., Soares J., Dallender J., Thomsen S., & Arnetz B. (2001) A comparative study of the experiences of violence of English and Swedish mental health nurses. *International Journal of Nursing Studies*, 38, 419-426.
- Nunnally J.C. (1978) *Psychometric Theory*. New York : McGraw-Hill.
- Nursing and Midwifery Council [NMC] (2004a) Changing a culture: managing violence in mental health *NMC News*, April, No 7. London : NMC Publications.
- Nursing and Midwifery Council (2004b) The role of the NMC in managing violence *NMC News*, July, No 8. London : NMC Publications.

- Nursing Times (2003) News item Nurses shocked at killing of colleague on psychiatric ward. *Nursing Times*, 99(25), 2.
- Occupational Safety and Health Administration (OSHA) (1999) *Workplace violence* <http://www.osha.gov/oshinfo/priorities/violence.html> Web page accessed 12th June 1999.
- Oppenheim A.N. (1992) *Questionnaire Design, Interviewing and Attitude Measurement*. (2nd Edn.) London : Pinter Publishers.
- Pancer S.M. & Westhues A. (1989) A developmental stage approach to program planning and evaluation. *Evaluation Review*, 13(1), 56-77
- Parahoo K. (1993) Questionnaires: use, value and limitations *Nurse Researcher*, 1(2), 4-15.
- Parahoo K. (1997) *Nursing Research: Principles Process and Issues* London : Macmillan.
- Parkes J. (1996) Control and restraint training: a study of its effectiveness in a medium secure psychiatric unit. *Journal of Forensic Psychiatry*, 7(3), 525-534.
- Paterson B., Turnbull J., & Aitken I. (1992) An evaluation of a training course in the short term management of violence. *Nurse Education Today*, 12, 368-375.
- Paterson B., Tringham C., McComish A., & Waters S. (1997) Managing aggression and violence: a legal perspective on the use of force. *Psychiatric Care*, 4(3), 128-131.
- Paterson B. & McComish S. (1998) The physical management of violent behaviour. *Psychiatric Care*, 5(6), 228-231.
- Patrick J. (1992) *Training: Research and Practice*. London : Academic Press.

- Patton M.Q. (1981) *Creative Evaluation*. Newbury Park, CA : Sage.
- Patton M.Q. (1986) *Utilization-Focused Evaluation* (2nd Edn.) Newbury Park, CA : Sage.
- Patton M.Q.(1996) 'A world larger than formative and summative' *Evaluation Practice*, 17(2), 131-144.
- Patton M.Q. (2002) *Qualitative Research and Evaluation Methods* (3rd Edn.) Thousand Oaks CA : Sage.
- Pawson R. & Tilley N. (1997) *Realistic Evaluation*. London : Sage.
- Perone S. (1999) *Violence in the workplace* Research and Public Policy Series, No 22, Canberra, Australian Institute of Criminology. Cited in V. Bowie (2000) Current trends and emerging issues in workplace violence. *Security Journal*, 13(3), 7-23.
- Phillips J.J. (1990) *Handbook of Training Evaluation and Measurement Measures*. London : Kogan Page.
- Phillips D. & Rudestram K.E. (1995) Effect of non-violent self defence training on male psychiatric staff members' aggression and fear. *Psychiatric Services*, 46(2) , 164- 168.
- Polit D.F. & Hungler B.P. (1995) *Nursing Research: Principles and Methods*. Philadelphia : JP Lippincott and Co.
- Poster E.C. (1996) A multinational study of psychiatric nursing staffs' beliefs and concerns about work safety and patient assault. *Archives of Psychiatric Nursing*, x(6), 365-373.
- Poster E.C. & Ryan J. (1994) A multiregional study of nurses' beliefs and attitudes about work safety and patient assault. *Hospital and Community, Psychiatry* 45(11), 1104-1108.

- Poyner B. & Warne C. (1986) *Violence to Staff: A Basis for Assessment and Prevention*. London : Health and Safety Executive.
- Quinones M.A. (1997) Contextual influences on training effectiveness. In M.A. Quinones & A. Ehrenstein (Eds.) *Training for a Rapidly Changing Workplace: Applications of Psychological Research*. Washington DC : American Psychological Association. Ch 7 Pp177-199.
- Reid M.A. & Barrington H. (1997) *Training Interventions: Managing Employee Development* (5th Edn.). London : Institute of Personnel and Development.
- Roach L. (1997) Violence at work *Nursing Standard*, 12(5), 23-24.
- Robson C. (1993) *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. Oxford : Blackwell.
- Robson C. (2000) *Small-Scale Evaluation: Principles and Practice*. London : Sage.
- Roger K.A. & Chappell D. (2003) *Preventing and Responding to Violence at Work*. Geneva : International Labour Office. [Cited on Institute of Conflict Management web site]
- Rosenthal T.L., Edwards N.B., Rosenthal R.H., & Ackerman B.J. (1992) Hospital violence: site, severity, and nurses' preventive training. *Issues in Mental Health Nursing*, 13, 349-356.
- Rossi P.H. & Freeman H.E. (1993) (5th Edn.) *Evaluation: A Systematic Approach*. Thousand Oaks, CA : Sage.
- Rossi, P.H., Freeman H.E., and Lipsey M.W. (1999) (6th Edn.) *Evaluation: A Systematic Approach*. Thousand Oaks, CA : Sage.

- Royal College of Nursing (1994) *Violence and Community Care Staff: A Royal College of Nursing Survey*. London : RCN.
- Royal College of Nursing (1997) *Trainers in the Management of Actual or Potential Aggression: Code of Professional Conduct and Minimum Training Standards*. London : RCN.
- Royal College of Nursing (1998) *Dealing with Violence Against Nursing Staff: An RCN Guide for Nurses and Managers*. London : RCN.
- Royal College of Psychiatrists (1998) *Management of Imminent Violence: Clinical Practice Guidelines; Quick Reference Guide*. London : Royal College of Psychiatrists.
- Royce J.R. (1958) The development of factor analysis *The Journal of General Psychology*, 58, 139-164.
- Rutman L. (1977) *Evaluation Research Methods: A Basic Guide*. Beverly Hills, CA : Sage.
- Ryan J. & Poster E. (1993) Workplace violence. *Nursing Times*, 89(48), 38-41.
- Sanderson G. (1992) Objectives and Evaluation. In S. Truelove (Ed.) *Handbook of Training and Development*. Oxford : Blackwell Business. Ch 5 pp114-142.
- Scriven M. (1967) 'The methodology of evaluation' in R.W. Tyler, R.M. Gagne & M. Scriven (Eds.) *Perspectives of Curriculum Evaluation*. Chicago : Rand McNally. pp. 39-83.
- Scriven M. (1991) 'Beyond formative and summative evaluation'. In G.W. McLaughlin & D.C. Phillips (Eds.) *Evaluation and Education: At Quarter Century*. Chicago IL : University of Chicago Press. pp. 19-64. Cited By Clarke (1999).

- Shacklady J. (1997) Violence in the community. *Journal of Community Nursing*, 11(10), 12-15.
- Shepherd M. & Lavender T. (1999) Putting aggression into context: an investigation into contextual factors influencing the rate of aggressive incidents in a psychiatric hospital. *Journal of Mental Health*, 8(2), 159-170.
- Sheridan M., Henrion R., Robinson L. & Baxter V. (1990) Precipitants of violence in a psychiatric setting. *Hospital and Community Psychiatry*, 41(7), 776-780.
- Smith J.C. & Hogan B. (1988) *Criminal Law*. (6th Ed) London : Butterworth.
- Smith-Pittman M.H. & McKoy Y.D. (1999) Workplace violence in healthcare environments. *Nursing Forum*, 34(3), 5-13.
- St Thomas Psychiatric Hospital (1976) A program for the prevention and management of disturbed behavior. *Hospital and Community Psychiatry*, 27(10), 724-727.
- Stecher B.M. & Davis W.A. (1987) *How to Focus on Evaluation*. Newbury Park, CA : Sage.
- Streiner D.L. (1993) A checklist for evaluating the usefulness of rating scales. *Canadian Journal of Psychiatry*, 38(2), 140-148.
- Streiner D.L. (1994) Figuring out factors: the use and misuse of factor analysis. A checklist for evaluating the usefulness of rating scales. *Canadian Journal of Psychiatry*, 39(3), 135-140.
- Streiner D.L. & Norman G.R. (1989) *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford : Oxford Medical Publications.

- Suchman E. (1972) "Action for what? A critique of evaluative research" in C.H. Weiss (Ed.) *Evaluating Action Programs* Boston : Allyn and Bacon.
- Cited by Rutman L. (1977) *Evaluation Research Methods: A Basic Guide*. Beverly Hills CA : Sage.
- Tabachnick B.G. & Fidell L.S. (1989) *Using Multivariate Statistics* [2nd Edn.], New York : Harper Collins.
- Taylor D. (2000) Student preparation in managing violence and aggression. *Nursing Standard*, 14(30), 39-41.
- Thackrey M. (1987) Clinician confidence in coping with patient aggression: assessment and enhancement. *Professional Psychology: Research and Practice*, 18(1), 57-60
- Thackwray B. (1997) *Effective Evaluation of Training and Development in Higher Education*. London : Kogan Page.
- Training Agency (1989) *Training in Britain*. London : HMSO.
- Turnbull J., Aitken I., Black L., & Paterson B. (1990) Turn it around: Short-term management for aggression and anger. *Journal of Psychosocial Nursing and Mental Health Services*, 28(6), 6-12.
- United Kingdom Central Council for Nurses, Midwives and Health Visitors [UKCC] (1986) *Project 2000: A New Preparation for Practice*. London : UKCC.
- United Kingdom Central Council for Nurses, Midwives and Health Visitors UKCC (2002) *The Recognition, Prevention and Therapeutic Management of Violence in Mental Health Care*. Report prepared for the UKCC. London : Health Services Research Department, Institute of Psychiatry.

- Waltz C.F., Strickland O.L., & Lenz E.R. (1991) *Measurement in Nursing Research*. (2nd Edn.) Philadelphia : FA Davis.
- Warr P., Bird M., & Rackham N. (1970) *Evaluation of Management Training*. Aldershot : Gower.
- Watson H.E. (1995) Psychometric analysis of assessment tools. *Nurse Researcher*, 2(4), 56-68.
- Weiss C.H. (1972) *Evaluation Research: Methods of Assessing Program Effectiveness*. Englewood Cliffs, NJ : Prentice Hall.
- Weiss C.H. (1997) How can theory-based evaluation make greater headway?. *Evaluation Review* 21 , 501-524.
- Wells J., & Bowers L. (2002) How prevalent is violence towards nurses working in general hospitals in the UK? *Journal of Advanced Nursing*, 39(3), 230-240.
- Whitley G.G., Jacobson G.A., & Gawrys M.T. (1996) The impact of violence in the health care setting upon nurse education. *Journal of Nursing Education*, 35(5), 211-218.
- Whittington R. (1994) Violence in psychiatric hospitals. In T. Wykes (Ed.) *Violence and Health Care Professionals*. London : Chapman & Hall. Ch 2, pp 23-41.
- Whittington R. (1997) Violence to nurses: prevalence and risk factors. *Nursing Standard*, 12(5), 49-56.
- Whittington R. & Wykes T. (1992) Staff strain and social support in a psychiatric hospital following assault by a patient. *Journal of Advanced Nursing*, 17, 480-486.

- Whittington R. & Wykes T. (1996) An evaluation of staff training in psychological techniques for the management of patient aggression. *Journal of Clinical Nursing*, 5, 257-261.
- Wilkinson C.L. (1999) An evaluation of an educational program on the management of assaultive behaviors. *Journal of Gerontological Nursing*, 25(4), 6-11.
- Wright S. (1999) Physical restraint in the management of violence and aggression in in-patient settings: A review of issues. *Journal of Mental Health*, 8(5), 459-472.
- Wykes T. & Whittington R. (1994) Reactions to assault. In: T. Wykes (Ed.) *Violence and Health Care Professionals*. London : Chapman & Hall. Ch 6, pp 105-126.
- Wynne R., Clarkin N., Cox T., & Griffiths A. (1997) *Guidance on the Prevention of Violence at Work*. Luxembourg : European Commission, DG-V. Cited by Beale *et al* (1998).

APPENDICES

Appendix One - Draft Strategy For Training In Aggression And
Violence For Pre-Registration Students (Summary)

Appendix Two - Revised violence questionnaire [different layout
because of altered page margins]

Appendix Three - Completed Request for Access to Students Form

Appendix Four - 100% Stacked for percentage responses to 24
statements at four time points

Appendix One

DRAFT STRATEGY FOR TRAINING IN AGGRESSION AND VIOLENCE FOR PRE-REGISTRATION STUDENTS (SUMMARY)

INTRODUCTION

The increase in aggressive and violent incidents in health care settings is clearly acknowledged. The increased risks for all staff, both trained and in training, is recognised widely, for example by the English National Board and more recently by the new Secretary of State for Health, Frank Dobson.

The English National Board (1993) specify that all pre-registration courses for nurses and midwives must include specific instruction in the theory of aggression and violence and helpful interventions. It further specifies that training should be provided by Control and Restraint Instructors who possess recognised teaching qualifications and have completed defined courses of study that emphasise non-physical strategies, anatomically safe physical restraint and who stress that physical restraint should be a last resort.

At the moment this instruction is distributed throughout the Common Foundation Programme and certain Branches. In the CFP, preparation for mental health placements includes advice about the role of the student nurse regarding aggression and violence, what to do and what not to do.

Some theory is also offered on models of aggression and the context of challenging behaviour. In the Mental Health Branch students have usually received some instruction and practice of breakaway techniques.

On the basis that the Department of Nursing and Midwifery now has a lecturer who is an accredited trainer with the RCN Institute and who satisfies the ENB conditions it is appropriate to propose a strategy for providing pre-registration students with a comprehensive and better coordinated unit of instruction on aggression and violence that includes theories of aggression and violence, health care settings and societal contexts, and that incorporates different levels of intervention skills. The programme is summarised below and on the next page.

INITIAL INPUT

The initial content could be presented in 3 days within the CFP programme in Trimester 3

Day 1 - in groups of 40 - 60

Day 2 and Day 3 - in groups of up to 25

REFRESHER DAYS IN BRANCHES

0.5 days per group - in groups of up to 25

Any further higher level input that was required for particular branches i.e. 5 day restraint / management of aggression and violence courses would need to be agreed separately.

PROGRAMME

Day One: Content

Philosophy of Prevention Self Awareness Personal Values Rights
Theories of Aggression and Violence
3 Hours

Assessment (of Self, Patient and Family, Environment)
3 Hours

Day Two: Content

Planning and Interventions
Roles, Prevention Strategies, Care Planning Training Readiness
2 Hours

Skills: verbal and non-verbal therapeutic responses
4 Hours

Day Three: Content

Role play Scenarios of difficult clients/ patients
3 Hours

Breakaway Training
3 Hours

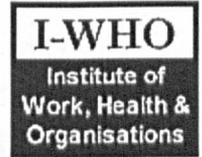
bb//02/ 12/97

Appendix Two - REVISED VIOLENCE QUESTIONNAIRE [DIFFERENT LAYOUT BECAUSE OF ALTERED PAGE MARGINS]



Institute of Work, Health & Organisations

Nottingham University Business School
Jubilee Campus, Wollaton Road, Nottingham NG8 1BB, UK
<http://www.i-who.org>



*A World Health Organization Collaborating Centre in Occupational Health
European Agency's Topic Centre on Stress at Work*

PREVENTION AND MANAGEMENT OF AGGRESSION AND VIOLENCE IN HEALTH SETTINGS TRAINING QUESTIONNAIRE I.D No: __

Personal Details	Group / Intake:.....	Gender: M / F	Age: ____
Destined Branch : (please ring)	Adult Mental Health	Learning Disabilities Midwifery	Child
Previous Training: Have you previously received instruction in any of the following aspects of the management of violence? :-			
• Theory of management of aggression and violence		Yes / No	
• breakaway skills training		Yes / No	
• physical restraint		Yes / No	
Experience of Violence: During this Trimester have you been involved in any 'violent incidents' whilst on your placements?			
Verbal Aggression (Abuse/Intimidation) Yes / No If 'Yes', How many? ____			
Physical Aggression/Violence Yes / No If 'Yes', How many? ____			
What placement were you on? (Please ring)			
Learning Disabilities		Adult Child	Midwifery Mental Health

SCENARIO 1

Read the following brief scenario and then answer the questions related to it.

It is late on Friday evening and the A&E Department is already very busy. A 30-year-old man arrives in an intoxicated and dishevelled state. He has a number of cuts to his face and hands and quickly becomes agitated. He shouts obscenities and mentions something about the Patient's Charter. On seeing the man's behaviour, a male nurse approaches and attempts to put his hand on the man's arm in an effort to placate him. The young man pushes him away and becomes even more vociferous.

Question 1

List below all the factors that you believe contributed to the young man's aggressive behaviour:-

Question 2.

Estimate, by making a mark at a point on the line below, to what extent you view the young man to be to blame for the incident:-

Not at all
to blame



Totally
to blame

-oOo-

SCENARIO 2

Read the following brief scenario and then answer the questions related to it.

Mrs Smith is a 75-years-old woman with a long history of psychotic episodes who lives in a residential home. She now suffers badly with arthritis in her knees and hips and needs staff assistance when using the toilet.

On this occasion, having just used the bathroom, she rings the bell to summon help. Most staff are in a hand-over meeting between shifts and no-one answers her call. After waiting a few minutes she starts shouting for assistance and banging on the cubicle door. A female member of staff, about to finish her shift, arrives to assist the lady back into her chair as quickly as possible. She pulls the old lady to her feet rather quickly, saying "There is no need to shout – I have others to attend to as well you know!". At this point the old lady swears at the member of staff and lashes out at her.

Question 3

List below all the factors that you believe contributed to the old lady lashing out:-

Question 4.

Estimate, by making a mark at a point on the line below, to what extent you view the old lady to be to blame for the incident:-

Not at all
to blame



Totally
to blame

Please indicate your level of agreement - strongly agree [SA], agree [A], undecided [U], disagree [D] or strongly disagree [SD] - with each of the following statements, by placing a tick in the appropriate column.

Statement	S A	A	U	D	S D
1. Health professionals should accept that being assaulted is an inevitable part of the job					
2. Mentally ill patients are always likely to behave more aggressively than others					
3. It is unacceptable for nurses to protect themselves when being physically assaulted by a patient					
4. Most aggression and violence by patients is predictable					
5. When staff members are assaulted and have no injuries, there is no need to report the assault					
6. Prediction of patient assault is within the competence or ability of practising psychiatric nursing staff					
7. The members of staff who are physically assaulted are generally those who are least competent in their job					
8. Staff with an authoritarian manner are more likely to be assaulted					
9. Staff with a yielding and submissive manner are more likely to be assaulted					
10. We under-estimate how much people with mental illness or learning disabilities are responsible for their behaviour					
11. Most aggression and violence by patients is preventable					
12. Staff should be educated about the prevention and management of aggressive behaviour as part of their pre-registration preparation					
13. When a patient becomes increasingly aggressive I get so nervous that I can hardly think straight					
14. I am able to demonstrate a non-provocative approach towards, and stance in front of, a patient					
15. The student nurse's role in a violent incident is to remain in the background and let qualified staff deal with the situation					
16. I can describe the main principles of breakaway techniques					
17. I am confident of my ability to remain safe at work					
18. I feel confident in my own ability to manage a patient's behaviour as it becomes more aggressive					
19. It is always better to intervene sooner rather than later in aggressive situations					
20. I am confident of my ability to see the potential for violence in health care situations					
21. I am able to talk in a calming and reassuring way to a					

verbally aggressive patient/relative and manage the situation					
22. I am fully aware of the legal framework within which self defence is legitimated					
23. I am confident of my ability to protect myself using legally permissible force if attacked by a patient or visitor					
24. Staff have a right to take legal action against patients who have assaulted them					

**Appendix Three - COMPLETED REQUEST FOR ACCESS TO
STUDENTS FORM DEPARTMENTAL RESEARCH COMMITTEE
MEETING – 9TH MAY 2000**

**Request for access to student cohorts for research purposes- additional
information**

Please find below information detailing a request for access to a number of pre-registration diploma student cohorts.

**Training Student Nurses in the Prevention and Management of Aggression
and Violence in Health Settings**

Background.

Violence in the NHS and health care sector has gained a deservedly high profile over the last few years. It is now acknowledged as a serious problem that impinges on financial budgets, staff physical and psychological well being, staff recruitment and retention and quality of care. It is also clearly a major health and safety issue with serious implications for managers and staff.

The Department of Health has demonstrated its determination to tackle this area through a number of recent policy statements and initiatives, two of the most notable being the recent Zero Tolerance campaign and the review of restraint being undertaken by the Institute of Psychiatry.

Education and training to develop knowledge, change attitudes and increase inter-personal and psychomotor skills is now widely advocated as

the way forward to reduce the number of violent incidents (HSAC 1997). Consequently, the creation of a lucrative business opportunity has resulted in a large number of organisations and individuals offering their services to health providers. However evaluation of the effectiveness of this training activity has generally been distinguished only by its paucity. At the same time, surveys have shown that student nurses, as a staff grouping, have the highest risk of sustaining all categories of violence, from verbal abuse through to assault with a weapon. Yet this group has largely been excluded from any staff training programmes.

Training Initiative

It is against this backdrop that, two years ago, the Department of Nursing and Midwifery, Keele University introduced a 3-day Unit of instruction for all pre-registration student nurses during the first year of their course. The unit emphasised the development of knowledge, attitudes and skills and some efforts have already been made to measure its effectiveness (Beech 1999).

Evaluation

The serious measurement of effectiveness of any educational input would need to distinguish between predetermined, anticipated changes and random changes in knowledge, attitudes, and skills; changes induced by other sources; the relative robustness and permanence of any change when exposed to the reality and competition of clinical practice and clinical staff.

Research Details

I am registered for MPhil /PhD studies within the Institute of Work, Health and Organisations (I-WHO), University of Nottingham and the research detailed below is planned as part of this post-graduate study.

My post-graduate studies now provide me with the opportunity to perform evaluation research using a quasi-experimental, repeated measures design that can systematically obtain robust data which sheds light on the role of training in determining knowledge, attitudes and skills. For this reason I request permission to access a number of student nurse cohorts and administer a confidential, non- anonymous questionnaire at a number of time points pre- and post- Unit of instruction.

Specifically, I request permission to access:

Group 09.1999 in Trimester 3 and 4

Group 02.2000 in Trimester 1, 3 and 4

Group 09.2000 in Trimester 1, 3 and 4.

Research Study Title: An evaluation of the effectiveness of a unit of instruction on prevention and management of aggression and violence in health settings.

Aim : To determine the effects of a unit of instruction on the aggression and violence-related knowledge, attitudes and behaviours of student nurses.

Data Collection Instrument: A questionnaire is currently being finalised which will contain a number of statements and questions that are designed to ascertain aspects of knowledge and awareness, attitudes and beliefs, self-esteem and confidence and skills. These questions obviously relate closely to the specified aims and objectives of the Unit.

Dissemination Strategy: Initially in documents to satisfy MPhil/PhD requirements. Subsequently, it is intended to publish individual articles whilst adhering to criterion 8 of the 'Conditions of Access' document. A signed copy of this document is enclosed.

Thank you for your consideration of this request.

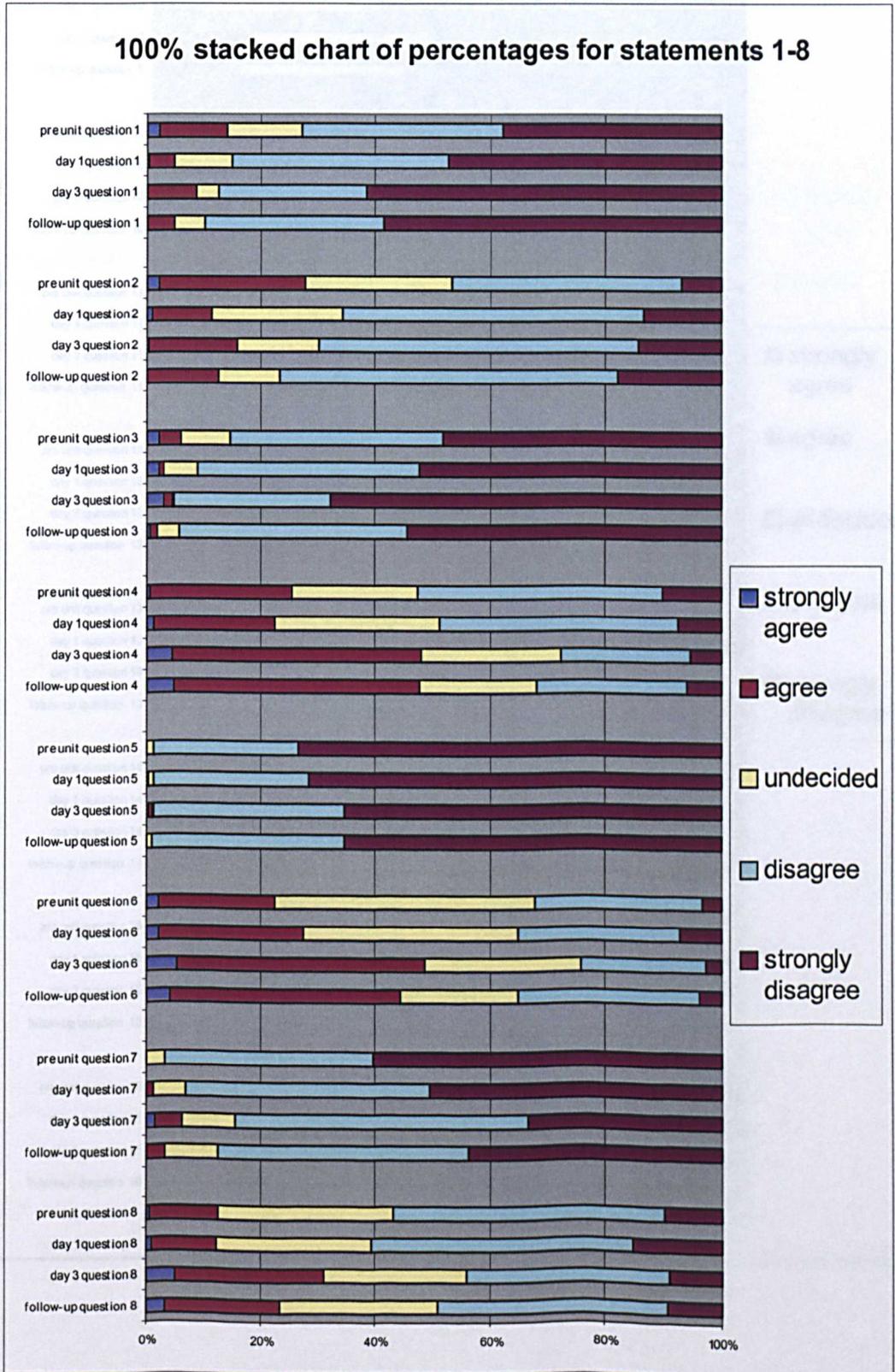
Bernard Beech

Lecturer,

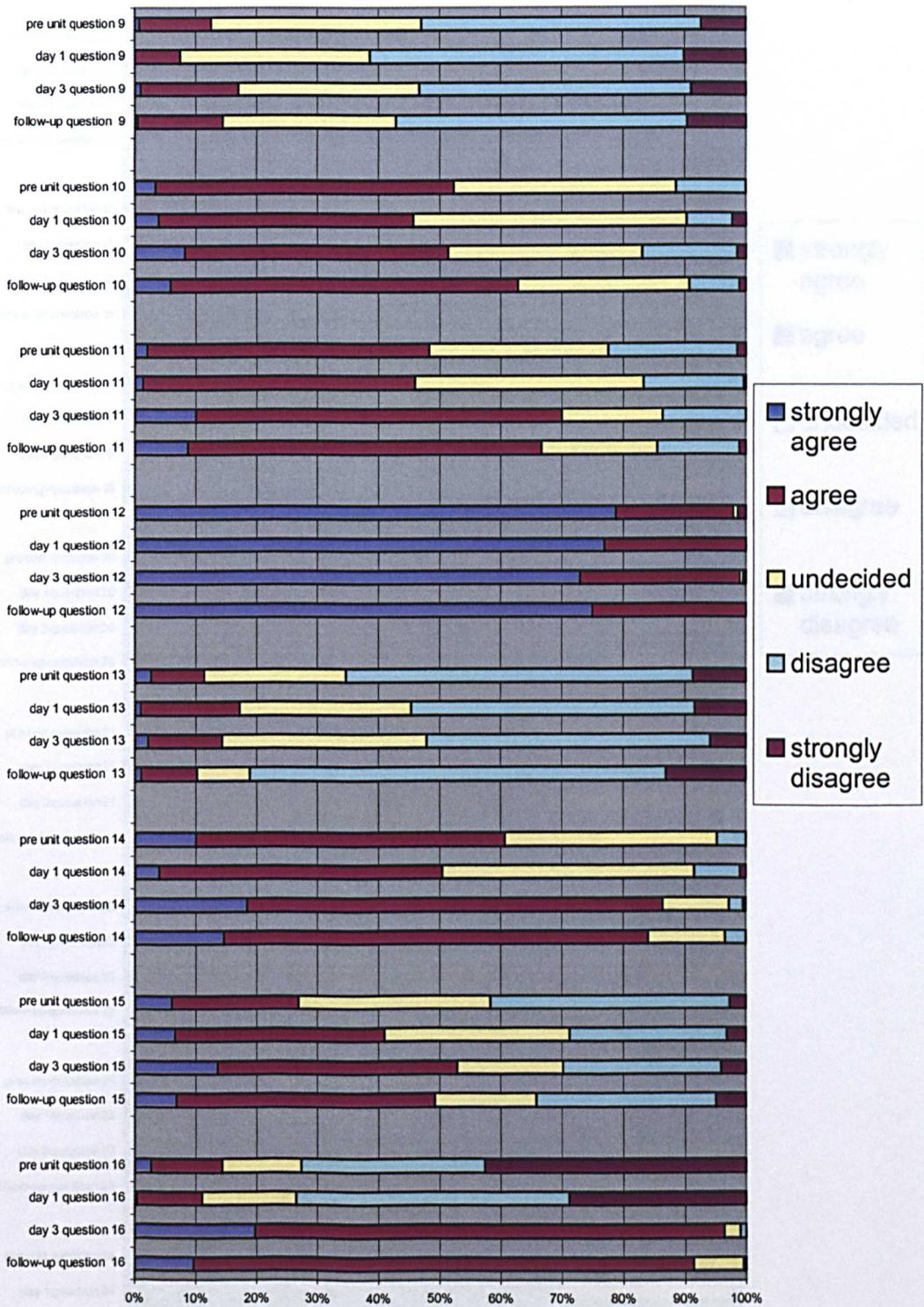
Department of Nursing and Midwifery, Keele University

25th April 2000

Appendix Four – 100% Stacked for percentage responses to 24 statements at four time points



100% stacked chart of percentages for statements 9-16



100% stacked chart of percentages for statements 17-24

