Clarkeburn, Henriikka (2000) *How to teach science ethics.* PhD thesis.

http://theses.gla.ac.uk/2852/

How to Teach Science Ethics

Strategies for encouraging moral development in biology (and other) students through the design and use of structured exercises in bioethics.

by
Henriikka Clarkeburn

Presented in candidature for the degree of Doctorate of Philosophy, to the Institute of Biomedical and Life Sciences, University of Glasgow

GLASGOW, 2000

To Jeffrey

Declaration

I declare that the work recorded in this thesis is entirely my own, unless otherwise stated and that it is my own composition. No part of this work has been submitted for any other degree.

<div style="text-align: right;">

Henriikka Clarkeburn
June, 2000

</div>

Acknowledgements

Table of Contents

Tables

Figures

Summary

The aims of this European Commission funded project, carried out at the University of Glasgow, were to develop an approach for the inclusion of ethics in a science undergraduate curriculum and to evaluate the success of that approach. The moral nature of science as an academic discipline and as a professional career justifies the resources spent on novel ethics teaching within a science course. Choices in science - allocation of research funds, selection of research topics, interaction with research subjects (animals, environment, other humans) etc. - often, if not always, include some elements of morality. The dilemmas involved require decision-making which cannot, and should not, be made without reflection on the values that govern science and society at large.

From the student perspective, the ethics curriculum aims to promote and accelerate moral development. In the context of ethics teaching in a science curriculum, moral development consists of two components: moral sensitivity and moral cognitive skills. Moral sensitivity is an ability to understand that moral aspects are as valid as factual data, and to distinguish between the two. Moral cognitive skills consist of an ability to 1) analyse the moral aspects of a situation, 2) differentiate the significant from the insignificant, 3) foresee the moral consequences of actions, and 4) to make moral decisions, in particular when it is necessary to choose between two or more incompatible values.

The core idea of moral development is progression through stages, each signified by certain skills and abilities to analyse and solve moral problems. This development is mostly linear and may continue through adulthood. All those equipped with the basic cognitive skills have an ability to reach the highest stage, although not all will. The progression is generally a move from a dualistic and selfish world-view, where moral networks relate only to other immediate agents, towards increased appreciation of multiplicity and a wider range of moral shareholders. At the higher stages, moral agents are increasingly able to approach moral problems with confidence and the skills required for forming

consistent and sophisticated moral arguments. This stage-like understanding of moral development is based on the works of Kohlberg, Perry and Gilligan.

The approach employed to encourage moral development in this research is based on supporting students' personal involvement in solving moral problems. Student-centred small group teaching is the pedagogical method that allows such involvement. In this research two teaching methods were developed and tested: 1) Structured discussion groups, which involve both individual preparation, based on philosophical readings, and group discussions. 2) Problem-Based Learning exercises. All themes and material used in these teaching sessions were selected to provide links as close as possible with the existing scientific curriculum. For example, genetics students concentrated on genetic testing, and parasitology students on research issues relating to vaccine development and testing.

During the research period of the 1999-2000 academic session, 250 students in Scottish University Level 3 participated in ethics teaching. Another 250 students formed a control group. Ethics was introduced into components of the science curriculum, which are not formally assessed or where the assessment is skill- and participation-based, to avoid any inequality among students.

The ethics curriculum was evaluated by administrating the same Moral Development Questionnaire during the first weeks of term 1 and then again during the last weeks of term 2. The questionnaire consisted of three parts: 1) The Defining Issues Test (DIT) purchased from the University of Minneapolis. DIT is the most widely used test tool for measuring cognitive moral development based on Kohlberg's theory of moral development. 2) A moral sensitivity test consisting of a decision-making scenario in which students were asked to list no more than five considerations which they believed should be reflected upon before making the decision. The students' considerations were analysed and scored according to their depth of understanding of the moral elements in the decision-making. 3) An Osgood scale test on the meta-understanding of moral questions, based on Perry's Ethical Development Scheme.

The results from this work had three main elements:

1. The students' self-reported learning was dominantly in accordance with the aims of the teaching - increased awareness, initiative to think, and practice of moral decision-making skills.

2. The Moral Development Questionnaire results showed significant increases in students' moral sensitivity due to the ethics teaching they participated in.

3. The student population as a whole was using relatively low level moral decision-making tools in an inconsistent manner.

The minimal ethics teaching intervention used in this study was a success as it captured students' motivation and interest and supported moral sensitivity development, which is the first step of moral development. The results show that ethics education is needed to support students' search for adequate moral decision-making tools and their ability to include moral considerations in their general decision-making. The intervention designed in this research provided a methodological approach to successful ethics teaching, while it was inadequate in length to provide significant benefits in moral decision-making to students.

Introduction

The aim of this European Commission funded project, carried out at the University of Glasgow, was to develop an approach for the inclusion of ethics in a science undergraduate curriculum and evaluate the success of that approach.

The most obvious question to ask of any proposal to teach ethics is 'why should one want to undertake such a venture in the first place?' My answer to this question is that ethics should be taught because morality is part of any reflective personal life. This is to say that moral problems are inescapable. It is difficult to think of any possible life plan that would not be determined or conditioned by moral values. Ethical problems arise in all stages of life; they are part of all professions and academic disciplines; and social life is full of moral issues.

Life sciences, as an academic discipline, and career opportunities in science, are therefore included in the field of morality. Choices in science - allocation of research funds, selection of research topics, interaction with research subjects (animals, environment, other humans) etc. - often, if not always, include some elements of morality. The dilemmas involved in these decisions require decision-making which cannot, and should not, be made without reflection on the values that govern science and society at large. To enable scientists to deal with the moral aspects of decisions in science is a worthy goal that may justify the resources spent on ethics teaching within a science course.

The importance of science ethics is highlighted by the growing influence of science in society. The more influential science becomes, the more ethical issues become associated with scientific practice directly, and scientists are increasingly required to participate in the value questions born from new knowledge and new technologies. Edward Hackett (1993) suggests five reasons for the growing prominence of science, which all suggest new areas where ethical judgements have to be made:

> 1. Increasing cost of science. Scientific projects have become increasingly expensive and visible to the public. This causes rising

performance expectations, budgets to swell and mega-projects to emerge (e.g. the Human Genome Project). As science uses scarce (public) resources, it is inviting increased accountability and scrutiny of its affairs.

2. <u>Science is a resource for power</u>. Science is a resource for power offering the possibility of controlling aspects of nature and the attendant ability to alter the world. The most notable accomplishments of 20th Century science include molecular biology and genetics, vaccines and antibiotics, information technology, and atomic physics. All of these have influenced people directly by providing medical benefits, comforts in new forms of power, or accessibility to information. But new science is not solely beneficial; atomic physics allowed for the creation of nuclear weapons causing direct physical and indirect psychological harm. Molecular biology, in particular genetics, and improved information technology have generated fears of misuse, surveillance, and discrimination. It may seem that those who control science also have control over many aspects of the world.

3. <u>Science is a means of legitimising and justifying power</u>. Science is a source of intellectual property, legitimacy, and prestige capable of explaining or justifying actions that might not otherwise win social acceptance. Scientific arguments can be used to disguise political preferences; and the exercise of power - in influencing and invoking scientific argument - conveys the ability to develop or restrict such arguments. Naturally, scientific knowledge can be used not to disguise but to support novel and beneficial policies that would not be accepted without scientific backing, showing their usefulness and desirability. Science, as long as it has at least a relatively independent status, may threaten to reveal such exercises of preference and power concealed in the guise of objective necessity. While knowledge is not the only source of power, it is an important source. Therefore there are motivations to control science and the direction it takes.

4. <u>The high esteem of science is comparable to a religion</u>. Science is the new religion, given the high esteem that has usually only been associated with religion. Life in science is often referred to as a

'calling'. Religious metaphors are used as a rhetorical device to justify or provoke moral outrage or the practices of the sacred institutions of science. Misconduct is labelled sacrilege, dishonest scientists are betrayers of truth or false prophets, genetic engineering is accused of 'playing god', etc..

5. Intensification of the relation between scientific findings and social policy. Scientific research is increasingly involved in the formulation of public policy in the fields of public health, national defence, economic competitiveness, and the resolution of legal disputes. Science journalism has brought to the wider public the new findings of science and thus in its part created a tendency to expect science to produce answers to difficult questions in social and legal policy.

This close interdependency that exists between science and public life influences the social organisation of science, scientists' roles and career opportunities, and even the principles that guide scientific work. For example, cuts in university budgets may increase dependency upon external research funding, and commercial contracts. Scientists may experience organisational pressures to become more entrepreneurial, to undertake more externally funded research, and to perform it according to businesslike practices of accountability, efficiency, secrecy, and the like. This might increase competition between scientists for research funding, which might in turn compromise standards of co-operation, communication, quality, and choice of problems.

On each level of this example we can find moral issues; 'Is there a particular worth in universities being independent research organisations?', 'What is the value of open and accumulative academic information?', 'Who should decide on the research questions?', 'What methods of research are acceptable?', and 'Who should benefit from research results?'. These are moral questions, which require moral answers. To be able to engage in a moral discussion, one needs moral decision-making skills; skills to understand and recognise moral issues, and skills to find solutions to moral problems.

These skills are part of a basic personal repertoire of social skills and they are often acquired latently by observation and participation in social situations. Thus, while moral decision-making skills are not dependent on formal education, their development can be accelerated and supported by it. This is the aim of the ethics teaching developed and evaluated in the research described in this thesis - to find out how formal ethics education can support and accelerate the naturally occurring development of moral skills.

The teaching of science ethics, as I will describe it, does not involve teaching 'right' values or inculcating 'correct' behaviour. It rather consists of learning and practising the ethical skills of recognition, interpretation, understanding, solution-finding, and relating to moral problems perceived in their context and in relation to fundamental personal values. These are skills not only applicable to ethical problems in science, but to all moral dilemmas one may encounter. Therefore, even if one sees science only as a mechanism to produce hard data (with the exclusion of scientists from the moral consideration of applying that data), it is possible to see the value of ethics education within the undergraduate science curriculum. Moral skills, though in great demand within the scientific discipline, are also important skills in personal development towards a mature adult agency in any activity within a democratic society.

Even though ethics has become prominent and an important part of scientific practice, British universities offering a degree in life sciences have so far not included ethics formally in their undergraduate curricula. This is in stark contrast with medical education, where as early as in the 1950's the General Medical Council recommended medical ethics to be taught to all medical students (Boyd, 1987). Even though medical students may expect to face more obvious moral dilemmas in their future interactions with patients, the research related ethical issues in medicine are highly similar to those in the life sciences. There should be no reason why medical students alone should have formal training in dealing with the moral issues relating to life science and biomedical research.

In autumn 1998, I sent an e-mail survey to 58 heads of department/faculty of British universities offering an undergraduate life sciences degree. 37

universities replied to the survey. Only 10 universities reported some ethics in their undergraduate science curriculum. The University of Aberdeen has compulsory ethics for fourth year bioscience students as part of 'Aspects of Research on Humans and Animals', which was also included in the degree examination. The University of Bath offers optional ethics tutorials for 3rd and 4th year students. The University of Bristol conducts a short informative course on the ethical use of animals. The University of Central Lancashire offers an optional self-study module on ethics for 2nd year students. Essex and Exeter Universities include ethics in their compulsory lectures and tutorials with assessment in the form of essays and exam questions. In the University of Plymouth, ethics is discussed in 3rd year tutorials, while the University of Reading offers an optional course in 'History and Philosophy of Natural Sciences'. In the University of St Andrews students discuss ethics in compulsory and optional lectures and tutorials and answer essay questions on ethics in their exams. And finally, in the University of Swansea, undergraduate students in the 2nd year have compulsory lectures that touch on ethical issues. While all these universities have included ethics at some stage of their undergraduate degree, none indicated that ethics would be approached formally or that ethics teaching was used to reach certain goals or that the successfulness in reaching these goals would be measured. Full details of the e-mail survey can be found in 'Appendix I: University Survey'.

The lack of a formal and structured approach to ethics teaching in UK universities highlights the importance of this research. The importance of ethics is recognised and now it is the time to turn these values into action. To increase the benefits of this research and encourage joint efforts between universities on ethics teaching, the course material developed during this project will be made freely available on the Internet for other UK universities.

The research described in this thesis is a somewhat unusual inter-disciplinary project. It is unusual for someone with a traditional philosophy training to be carrying out research in a science department on an educational problem. The structure of the thesis therefore needs to bring coherently together strands from different academic disciplines: philosophy, psychology, pedagogy, and science.

This results in longer than usual theoretical considerations in order to provide a sound basis for the experimental research work.

Before I outline the structure of this thesis, I would like to clarify how I will use the terms 'ethics' and 'moral' in this work. Ethics can mean three different things: 1) it can be used as a synonym for morality; 2) it can be used to refer to morality in a professional context, especially where codes are involved; or 3) it can be used to refer to the second-order study of morality or ethics in the first sense *i.e.* it is the same as moral philosophy. My usage of the terms is closest to the second definition. The aim is to develop an ethics curriculum, which would promote students' moral development. Personally, I might have been inclined to adopt the first interpretation of ethics being synonymous with moral, *i.e.* referring to ethical development instead of moral development, but as moral development is the chosen term in the literature, I considered it prudent to not adopt this practice.

The research described in this thesis is divided into five Parts, each necessary for a careful analysis of the subject area. Part I is devoted to theoretical discussions on the conceptual issues relating to education and philosophy, mainly concentrating on the complex issue of educational aims. It emerges from Part I that ethics education should have intrinsic aims which promote individual moral development in fostering the growth of moral decision-making skills.

Part II concentrates on psychological theories for assessing moral development. Assessment of moral development is important not only for the attainment of academic results, but also in order to justify the importance of an ethics course against the following external limitations:
1. resource limitations, and
2. time constraints.

Firstly, when resources are scarce, they should be used where they produce the best benefits. In other words, even if the goals of ethics education are valid, ethics education may not be able support its place in a science undergraduate

curriculum unless it can be shown to make a distinguishable positive impact on students' moral development.

Secondly, the science curriculum cannot include all areas of science, or even a comprehensive coverage of a specialised area in science. The choice between what is included and what can be excluded while maintaining levels of academic excellence and meeting the external criteria of employers and public bodies, is thus inescapable. Competition for student time within and between departments is intense and proposals for new courses or programmes are likely to meet less than enthusiastic reception from staff members concerned with obtaining minimally adequate time for their own particular areas. This is further exacerbated by competition for scarce teaching resources. For these reasons, new inclusions in the curriculum are often required to satisfy both the criteria of adequate aims and of effectiveness in satisfying these aims.

The question whether ethics deserves greater emphasis in the undergraduate science curriculum has not received unequivocal support from students or staff. Though Downie and Alexander (1989) report high staff agreement on the importance of ethics as part of the undergraduate science curriculum, the support seems to be highly theoretical in the absence of ethics teaching. Student interest in bioethics was shown to be high by Downie (1993), which as such supports the inclusion of ethics in a science undergraduate curriculum. An ability to show that teaching ethics has a significant positive impact might be useful in paving the way for further inclusion of ethics within curricula. Also the diversity of teaching methods used for achieving the aims of ethics education almost guarantees passionate debate and disagreement. In this situation the need for impartial, quantifiable, and objective standards of evaluation becomes ever more urgent.

A further purpose of this evaluation process is to identify three elements and how they influence moral development:

1. Whether the approach adopted in ethics teaching makes a significant change in 'moral development scores' (see section 4.3.1 'Assessment of moral development by use of prototypic statements' for a discussion on what is meant by 'moral development scores') . The two

approaches compared here are Problem Based Learning (PBL) and structured tutorial discussions. No didactic methods are included as they are hypothesised not to be a successful method of teaching ethics to science students (see discussion in Part III - Theories for teaching ethics). Significant differences between teaching methods would give valuable information in making future pedagogical decisions regarding how to teach ethics.

2. Whether the number of hours spent working on an ethics problem is comparable with the advances in moral development. Time is a scarce commodity to both students and teaching staff and an indication of whether the utility rate of ethics education rises up to a certain point and starts to decrease afterwards will give valuable information in planning the optimum exposure to ethics in the curriculum in order to achieve the set aims.

3. From a theoretical point of view, there is an interest in finding any correlation between development of moral decision-making skills and moral sensitivity. If a positive correlation if found, it might be possible to use only one assessment method in the future, with confidence that it will measure both aspects of moral development.

The assessment is therefore conducted for the purposes of research and pedagogical practice. The students assessed will not benefit directly from these research findings, though indirect benefit is to be expected in the form of improved teaching approaches.

Part III discusses the teaching approaches most suitable for achieving the aims set for ethics education in Part I. It emerges from previous research that ethics courses are most successful when they involve student-centred approaches to learning. This approach changes the teacher's role from information provider to facilitator of student development. The skills involved in being a good facilitator are different from being a good teacher and are therefore discussed in detail. Two distinct teaching methods which allow a student-centred approach to learning moral decision-making skills are then discussed; Problem-based Learning and Structured group discussions. Part III also includes a description of the teaching

approaches adopted in this research for teaching ethics to University of Glasgow life sciences undergraduate students during the academic session of 1999-2000.

Part IV describes and discusses the results of the research. First is the qualitative analysis of students' responses to the ethics teaching together with descriptions on how the course material was applied in the class-room. This is followed by an analysis of the Moral Development Questionnaire.

Part V completes the work in a short discussion on the developed curriculum and the results. Part V also includes recommendations for the future based on the discussion.

Part I - Education and Ethics

The desire to teach people to do what is right seems more than justified in a society where crime and dishonesty are commonplace. But ethics teaching does not fit into the boundaries of teaching moral rights and wrongs: in fact it is poorly understood in those terms. The teaching of morally right behaviour, as will become apparent in this Part, is not just an inadequate understanding of ethics education, but an untenable approach in a liberal multicultural university.

Ethics education is not an isolated strand among educational activities, but in its structure comparable to them. In order to clarify the relationships between ethics education and general education theory, Part I will start with a conceptual analysis on basic educational terms: education, training, and indoctrination. Comparisons between education and training refer to the content aims of educational activities, while the comparison between indoctrination and education identify differences in educational methods.

The rest of Part I will concentrate on content aims, leaving the structural issues to be further discussed in Part III. The discussion on educational aims concentrates on the difference between intrinsic and extrinsic aims. This dimension of aims is essential to structuring ethics education. The choice between intrinsic and extrinsic aims has direct influence on the acceptability of ethics education as an educational activity as well as on the pedagogy and assessment used.

By the end of Part I the reader should have a clear understanding of the aims suggested for the research on science ethics education described in this thesis. Also, I hope to have convinced the reader that extrinsic aims are both unacceptable and impractical in ethics education and that ethics education is best organised around intrinsic aims, as they are described here.

1. Education - General Theory

The aim of this project is to develop an ethics learning programme for bioscience students. The more specific objective is to encourage students' moral development. What is meant by moral education or moral/ethical development is not self-evident, and thus the logical place to start this research is conceptual analysis. Successful conceptual analysis of the key concepts, which are concepts of the most polymorphous character, provides a solid ground for the entire research.

This first chapter will concentrate on the analysis of 'education' as a generic concept and on the appreciation of its relation to both content and procedure. The exact nature of education becomes clearer through comparison with neighbouring concepts, 'training' and 'indoctrination', where training can be contrasted with education to illuminate the nature of educational content, and indoctrination can clarify the nature of educational method.

1.1 Education - conceptual analysis

Education is a generic concept and one would be mistaken in considering it as a name for one activity alone. Education is rather an umbrella under which several educational values and practices can be fitted. Instructing students on laboratory practice, encouraging scientific discovery by providing a suitable setting, promoting good manners during departmental meetings, and providing factual information in lecture theatres can all be classified as part of the educational process, even though they are distinctly separate activities with varying aims and embedded values. Still, all educational practices belong to a family of activities united by a sometimes complicated network of similarities. I will here concentrate on the central usage of the term 'education' and leave to less, if not negligible, attention the many peripheral uses the term.

One of the common core elements in education as a concept is that something worthy of the time, effort, and resources should be achieved. This is also the core of

a famous formulation of education by R. S. Peters: "It [education] implies that something worthwhile is being or has been intentionally transmitted in a morally acceptable manner" (Peters, 1970, p. 25). In other words, education has normative implications. This can be viewed from a purely conceptual point of view. Both terms 'worthwhile' and 'morally acceptable' are value-laden statements, but describing education in these terms does not necessarily carry any commitment to the content - what is precisely to be understood as 'worthwhile' or 'morally acceptable'. At this stage all that is said is that education as a practice and discipline has normative implications and that both the value and content of these implications requires a separate analysis and appreciation.

'Worthwhile' as a concept has unmistakably positive connotations. Therefore, in the normal use of the term, it is illogical to say that people have been educated, but that they have in no way changed for the better or that the attempt to educate is not a practice worth attempting. There might be disagreement on what type of changes are changes for the better and therefore what type of educational practices are worth attempting, but whatever they are, the user of the term has placed a positive value on them. This does not indicate, however, that it would be illogical to talk of 'bad or poor education'. In doing so we think that the process of education was carried out poorly or indicate disagreement on the values chosen as the aims of an educational process.

Education is also an 'achievement' term. Built into it is an idea of achieving something that is worthwhile. It is essentially a word for progress. But education is not just an achievement word as it covers both the trying and succeeding; the process and results. Therefore we cannot identify education in terms of worthwhile goals alone, but must supplement it with a definition relating to the process. R. S. Peters provided one in the passage quoted from him above, i.e. it is carried out in a manner which is morally acceptable (Peters, 1970).

Both parts of education, the process and results, are important in classifying whether an activity is to be called education. I will use the method of comparison as a tool to define the parameters of educational activity, both in relation to content

and procedures. Comparing education with activities that resemble education, but are not, will provide useful distinctions and clarity. To do this for content, I will compare education with training, and for clarifying methods I will compare education with indoctrination.

1.1.1 Training and education

Here I will first describe the traditional conceptual difference between training and education, mainly following the argument and distinctions presented by Peters (1970). This clear cut division between training and education is then later criticised as inadequate for understanding education in all its various forms in present times.

According to Peters (1970) training and education have distinct differences. I will look at four aspects of training and education where the contrasting elements help to provide distinction between the activities of training and education:

1) normative vs. descriptive nature

2) knowledge/understanding vs. skills

3) wide vs. narrow cognitive perspective

4) initiating commitment and care vs. (relative) ignorance

The first distinction between education and training is their difference as normative and descriptive procedures. Educational aims have normative implications which are dependant on value systems. As I noted before, education is by definition something worthy in a normative sense. It brings about improvement in the person who is educated, which in more general terms leads to improvements for a society when it consists of educated citizens. Training, on the other hand, is defined as neutral and descriptive. It is most often applied to acquiring skills with no attention paid to the questions of how worthy these skills are or what moral implications they may hold.

Learning to steal is one example. Because stealing is not a worthwhile activity, at least within the common set of values in a society, it is difficult to form a concept

of 'theft education', while on the other hand there seems to be no logical problem in calling it training, learning the skill of stealing (Calman and Downie, 1988). But at the same time, it is possible to form an idea of theft education if we think of a society where some section of the population is forced to steal for livelihood without any alternative, or when stealing is done to possess falsely acquired goods with the intention to benefit those in great need i.e. the Robin Hood ideal. A better example can then be provided with torture. Trying to form a concept of torture education is extremely difficult, if not even impossible. It is difficult to imagine a situation where education in torture would be a worthy goal in individual development or from a societal point of view. On the other hand, torture training possesses no conceptual difficulties as learning the skills of torture requires no normative appreciation of their use. At least not as far as the distinction Peters puts forward.

The second distinction is between knowledge/understanding and skills. Education provides an understanding of principles for the organisation of facts, while training provides skills for a particular task. Education therefore also provides at least some understanding of the 'reason why' of things, while training concentrates on the question 'how'. Taking this distinction a step further, education also transforms, at least in some way, a person's outlook of the world and the knowledge gained influences perception and creates understanding. Training, on the other hand, can be successful without any understanding of 'why', or any associative cognitive changes, only learning the isolated performance of a particular procedure (Peters, 1970). For example, learning how to dissect a tadpole is not education, but training in skills. There is no logical problem in seeing the skill of dissecting a tadpole as necessary for other ends, in this case possibly an educational one, but training in the skill of dissecting is not alone either sufficient or necessary for creating understanding i.e. to be called education.

The third distinction is that for any activity to be called education it must have a wide cognitive perspective. This means that an educational activity may widen and deepen one's understanding of matters other than those included in the activity itself. Education therefore involves wholeness while training has a more limited

scope. For example, consider biochemists who cannot understand why their subject should be of interest to zoologists or medical professionals, as they are solely and exclusively interested in the molecular structure of particular proteins. It might then be more accurate to say that such persons have been trained as biochemists than that they have been educated in the field of biochemistry. A person can therefore be trained in complicated skills and have a command of a complicated but narrow field of information, but the term education would be reserved for activities which broaden and deepen understanding in a wider sense. Training can also work in the opposite direction and rather narrow the consciousness and limit the scope of interest. This might even be required in training for highly specialised skills, like surgery for example (Peters, 1970; Calman and Downie, 1988).

A fourth and final distinction between education and training is that in an educational activity, those who are engaged in it must come to care about what they are doing. Imagine a student who has graduated from a university with a biology or chemistry degree, both activities that can be 'worthwhile' and has a wide perspective, but thereafter shows no interest in the subject. It might be more appropriate to call such a student highly trained in the field of biology or chemistry, but maybe not educated (Calman and Downie, 1988).

The above described distinctions between training and education were originally presented by Peters and are open to criticism. The most apparent weakness of this analysis is the unnaturally sharp distinction between education and training. In reality the two terms seem either less distinct or, at least in some cases, in what is normally called education, overlapping.

First it is important to realise that training can also be broad-based as in training for the ministry of the Church. Secondly, when training is broad-based it is often impossible to reach the aims without wider understanding of the field and a deeper cognitive perspective. For example, being trained in some biochemical procedure may be a precondition for successfully understanding more advanced topics like the importance of nutrition in conception. This can work the other way round as well; to be successfully trained as a good laboratory researcher in biochemistry requires a

larger understanding of the processes in nature. i.e. to be educated in the field of biochemistry (Calman and Downie, 1988).

Despite these criticisms, comparing education and training has provided the following useful distinctions:

> 1) In education we are directly involved in normative implications in defining educational aims, while in training the aims as skills are as such neutral and descriptive, and the judgement on how worthwhile they are is not directly connected with the process of training.

> 2) Education aims to increase understanding and knowledge, in particular to answer the question 'why', while training is more concentrated on improving skills and finding answers to the question 'how',

> 3) Education is logically connected with a wider cognitive perspective, while training can have its interests in a limited skill or discipline.

> 4) Education necessarily involves an interest in the subject, while training can leave a person uninterested in the skills and knowledge gained.

> 5) It is important to realise that training and education are not used in the above described distinctive manner in most every day situations. Rather, training and education in most disciplines are both essential elements and often one cannot be achieved without another. The concepts are overlapping and interrelated.

If we accept the distinction between education and training as separable activities, our view on science education has to be a combination of training and education. Asking a question of 'what is it for?' in relation to different parts of the science curriculum, can be useful in clarifying which parts of the educational programme are best described as education and which training. If it is easy to come up with an answer like 'the course is for developing a laboratory skill X', then it might be

more appropriate to talk about training. While on the other hand, if no apparent narrowly focused answers prevail, as in the case of bioethics courses perhaps, it is more appropriate to talk about education. Naturally there are many courses and teaching components that do not as obviously fall under either description. This should not be seen as a problem, but only as evidence that training and education go hand in hand in science education, both being incremental parts of the process.

Students with science degrees should therefore be both educated and trained. They should have a wide cognitive perspective in how their field relates to other activities and still be trained in particular skills. These skills can be important either for a particular extrinsic purpose, or they are not taught with educational aims in mind because of either the complicated subject matter or lack of true interest for students in that particular field. Different subjects can be taught both with training and/or education in mind. For example, biochemistry can be taught as a tool to understand corporeal processes in humans and other animals i.e. as a subject worthwhile for its own sake. But biochemistry can also be taught as providing essential knowledge with a particular end in mind, for example, laboratory practices, when the activity would be most appropriately described as training rather than education.

When I refer to science education I have this larger concept in mind. But there will be times, when it is crucial for me to make a clear distinction between education and training, and I hope to be able to make it apparent when I am using education or training in their more limited senses.

1.1.2 Educational process and indoctrination

The definition of education given at the start of section 1.1 included one statement on method; that education should be carried out in a morally acceptable manner. For most people this would mean that indoctrination cannot be included as one of the educational activities, as the basic procedural tenets of indoctrination are not morally acceptable. Analysing the concept of indoctrination and why it cannot be

accepted as a method of education will clarify at least some of the boundaries of morally acceptable methods of education.

Even though I have included an analysis of indoctrination here primarily as a method of distinguishing the boundaries of acceptable educational practice, it is also important to understand something of its content matter. For something to be considered indoctrination, both the method and content are necessarily of a certain type, and education can share the content but never the method. Indoctrination, as far as content is concerned, is about matters of belief and faith, attitude and values. The fact that a teacher might be able to influence, even radically, students attitudes and beliefs is not a sufficient indication that the student has been indoctrinated. Therefore it is a necessary condition of indoctrination that beliefs should be imparted, but it is not alone sufficient. The question of method becomes crucial in providing a distinction between education and indoctrination.

For an educational process to be classified as indoctrination, the content of the activity must necessarily have three characteristics (Woods and Barrow 1975):

> 1) It must contain a system of beliefs which are interrelated and cover a large section of a general set of attitude and value structures.
> 2) It must be based upon certain propositions or postulates that cannot be demonstrated to be rationally true.
> 3) The type of beliefs that take part in indoctrination are not only unprovable, but there is no shared understanding about the type of evidence that could be used to prove certain beliefs true or false.

Moral and religious education are the most common areas where indoctrination takes place or at least is expected to be possible. Though these subjects are more prone to indoctrinal practices than many others, this is not to say that no other subjects are prone to indoctrinal practices. Even science, which is often taken to be truly factual in its content and therefore not a subject in threat of indoctrination, has elements of belief, faith and attitude built in it. Therefore it cannot be automatically considered immune to indoctrinating practices. Genetic research, for example, can

be strongly influenced by a belief that human nature is determined by genetic constitution, where this belief influences both the choice of research topics and interpretation of data.

Methods of indoctrination vary greatly from brain-washing techniques to subtle under-currents of education that are separate from the overt educational plans. The extremes of brain-washing and physical violations are quite broadly rejected as forms of education, but the further one travels down the line of increasingly subtle forms of educational activity proposed to be indoctrination, the harder it becomes to draw the line between what is indoctrination and what are acceptable forms of education. At the weakest end of the scale is an opinion that all education is in fact indoctrination and that this indoctrination is carried on by subtle strategies - dark sarcasm, choice of study elements, and the hidden curriculum etc. - rather than by overt means. According to this opinion, teachers are without any doubt conveying their biases in teaching and any such subtle passing of attitudes is to be considered indoctrination. An example would be that the current educational system re-enforces sex and/or racial stereotypes that already exist in society just because education is structured in certain ways, because certain types of people influence the teaching etc. and that this qualifies as indoctrination.

If indoctrination as a term was reserved only for crude and overt practices, then there would probably be very few educational institutions anywhere in Britain, or in any other Western countries, which could be said to indoctrinate and there would be no point in looking for a way of distinguishing indoctrination from education. So it is worth accepting that there could be more subtle and less easily recognised processes which would qualify as indoctrination. On the other hand, if all that is commonly called education could in fact be indoctrination, the distinction would become redundant. In believing that there is a middle ground, distinction between these terms is therefore taken as a starting point for an analysis of what are acceptable educational methods and what are to be rejected in terms of indoctrination.

For an educational method to be called indoctrination, it should be a process of inculcating belief on matters of the above mentioned content by non-rational methods. This is where indoctrination comes to a head-on collision with the normal understanding of what are ethically acceptable methods of education - that is, providing truth or at least the best version of it available and allowing for alternative truths to challenge the status quo. Providing lies or distorted truths as the truth is not acceptable. Neither is providing a limited scope for verification. To aim at inculcating unfalsifiable truths, it is necessary to use some form of non-rational persuasion designed to bring about unshakeable belief. An activity with such aims and employing methods of non-rational persuasion, possibly allied to a system of punishment, whether concrete or abstract, cannot be called education, but indoctrination.

The intention of the teacher is also a necessary element for an educational situation to be indoctrinating. It would seem very strange to talk about a person trying to bring about unshakeable belief in a non-provable statement without making a reference to the intentions of the person engaged. The intention may not be obviously present, but unmistakably the person engaged in indoctrination cannot be without an intention to promote the beliefs in his/her students. If a teacher who ceases to present the various views held by different people on a controversial subject (whether it is religious, ethical or political), attempts a particular view on the issues to be taken as the *only* truth by the students, which logically requires overriding the rationality of those same students, then we are entitled to call the activity indoctrination (Woods and Barrow, 1975).

When indoctrinating, the educator is conveying attitudes and biases in a way which overrides the autonomy of the students - their ability to think for themselves. This is the core of activities to be classified as indoctrination: they aim at limiting the use and development of personal intellect, the ability to think for oneself and choose independently and freely between non-factual issues that have elements of faith, attitude and belief unavoidably built into them. This is the opposite to education, which aims at encouraging personal attributes of rationality.

Engaging in ethics teaching is always susceptible to charges of indoctrination, and quite rightly so. Ethics teaching can take the form of indoctrination as it has a subject matter that contains non-provable beliefs and theories and the ethical and moral beliefs held by people have a strong impact on how they lead their lives. I wish to make it clear here, and hope that it will be apparent in all of the stages of this research, that the aims of the ethics teaching in the course of this project will be such that the threat of indoctrination should not cast a shadow on the educational efforts expended. How well I succeed in achieving this aim, ethics teaching that is not indoctrination, will be left to the reader, but at this point, there should be no doubt of the intent.

1.2 Summary

The concept of education can be limited by reference to both the content and methods of the activity. In reference to the content, the neighbouring concept is training. Peters (1970) offers four basic differences between training and education:

      1) normative vs. descriptive nature

      2) knowledge/understanding vs. skills

      3) wide vs. narrow cognitive perspective

      4) initiating commitment and care vs. (relative) ignorance

In reality the division is rarely clear-cut. Educational activities involve elements of both education and training, which in many cases are essential partners in achieving the overall aims. For the purposes of this research, the term education is used to refer to activities which have dominantly educational aims, while they may also contain elements of training in non-essential practices.

In reference to the methods of education, the counter-point is indoctrination. Indoctrination by definition aims to inculcate beliefs of an unjustifiable nature by methods which undermine the students' autonomy to use their rational abilities to weigh different alternatives and to choose for themselves. Ethics education is susceptible to indoctrinating practices, but as I hope to show, ethics education need not be indoctrinative, but can fulfil the criteria for genuine educational practice.

2. Educational aims

Justification for educational efforts is often given by reference to educational aims. Reference to aims is intended to provide adequate reasons to justify the expenditure in both time and financial resources needed for education. Even though the reasons for ethics education stated in the Introduction may win wide acceptance as general reasons why ethics education can support its place as part of science education, the actual aims of an ethics curriculum can still be an issue of disagreement and controversy. Due to the controversial nature of an ethics education, justification for the use of resources and effort requires explicit explanations of the aims of such a curriculum.

The aims discussed here are teacher-centred statements on what a course of study, in this case science ethics, is set to provide - what is the motivation for teaching science ethics and how can it benefit students? As will become obvious, the objective of the teacher-centred aims discussed here is to provide student-centred aims for studying science ethics - what can the students do as a result of their ethics studies? Teachers and students alike need aims for their educational activities and while most of this section concentrates on the motives a teacher or an educational institution could (and should?) have when introducing ethics into the undergraduate science curriculum, the true aims of ethics teaching emerge as student-centred.

Inquiring about aims often includes a question about purposes and motives. Asking for the aims of ethics education could be answered by stating only specific aims like 'the aim of this ethics exercise is to teach the students concepts of utility and justice', but a more adequate answer would include a reference to the motives and purposes of the educational activity, like 'the aim of this exercise is to teach the students the concepts of utility and justice, *because* this will give them tools to understand the nature of moral problems'. Educational aims are logically connected to values through the normative nature of education itself. They are necessarily a reflection on understanding of the term 'worthwhile' i.e. what we value as the end product of education. For this reason it is essential to spell out these normative

elements often embedded in statements regarding motivation and purposes of ethics education.

The value-richness of educational aims provides an additional reason for explaining these aims in detail. Because educational aims are rich in values, we can foresee no obvious agreement on what values should be promoted through these aims. In particular, proposing teaching in a subject even richer in values than many others, ethics, one has to be ready to provide analysis and clarification of the values which the teaching promotes and to be ready to face disagreement and argument regarding their importance and acceptability. Therefore, I believe that stating aims clearly and in detail is important in providing adequate information on the values embedded in the educational activity and in facilitating an opportunity for those disagreeing to voice their opinion.

This chapter is written to support the integration of science ethics into the undergraduate bioscience curriculum, by explaining in detail the aims of including ethics. I will focus my analysis on one classificatory aspect, the distinction between extrinsic and intrinsic aims of education. This is not the only distinction that can be drawn between educational aims, as they can be classified in a great variety of categories. These include distinguishing educational aims according to classification schemes based, as presented by Brezinka (1997), on:

- the degree of complexity,
- whether the aims are ultimate or intermediate,
- the level of abstraction,
- distinguishing between material and formal content (Aims are formal when they are general e.g. to enable a person to behave similarly with the greater moral whole to which he/she belongs, and material when they include specific aims e.g. learning to tie one's shoe laces),
- the number of people for whom specific educational aims are set (personal or collective), or
- whether the aims are intrinsic (good in themselves) or extrinsic (the good is external to education).

Statements on educational aims are a collection of classifications in chosen categories. Not all categories need to be included in each clarification of aims, but often more than one is required for an adequate explanation. Educational aims with great complexity and a high level of abstraction can come to resemble sentences of the following kind "a wish to image the physical-spiritual-intellectual final form of a person which has an aim-giving effect on conscious educational efforts"(Brezinka, 1997, p. 148), which are meaningless to most readers. For practical purposes, these types of aims-statements are less than useful, while they may serve some purpose in clarifying very high-level aims for educational practice and for discussion among people invested in the particular language used in such statements. My approach here will be less abstract and concentrates more on creating an aims statement for the benefit of the learner, not an educational academic.

But why concentrate on the distinction between intrinsic and extrinsic aims? The dichotomy between intrinsic and extrinsic aims is important to an analysis of ethics education, because the choice of either intrinsic or extrinsic aims has a direct influence on the actual teaching choices in ethics. This choice also has implications for the general acceptability of an ethics education as this choice touches on the very fundamental issues of teacher and learner autonomy and plurality of values. Finally, the choice between intrinsic and extrinsic aims for ethics education is likely to influence how successful one is in justifying the efforts and resources required for an ethics curriculum. For these reasons, I provide a detailed discussion on the nature of intrinsic and extrinsic aims, before moving onto describing the aims of ethics education.

The rest of this chapter has the following structure. First, an analysis of extrinsic aims in section 2.1, followed by an analysis of intrinsic aims in section 2.2. The chapter will conclude in a short summary.

2.1 Extrinsic aims of education

According to Hirst and Peters (1970), when educational aims are extrinsic, the
educational process itself is considered neutral and the aims of education are
decided outside the realm of education. Education is considered instrumental to
something that is worthwhile, but which is external to it. It follows from this that
extrinsic aims for a certain educational activity are aims which are only
contingently associated with the actual subject matter or educational process.

Further, extrinsic aims of education also provide an explicit purpose for the
educational activities. It is not just improved skills, but improved skills for a reason.
The aim is thus an identified and pre-perceived utility.

It can be said, for instance, that the aim (and often the justification) of
studying/teaching science is that it will ensure better paid employment in the future.
This would be an extrinsic aim for three main reasons:

1. Higher pay is only contingently associated with the actual educational
   content of a science curriculum, because better pay can be obtained by
   gaining other types of qualifications, and the increase in pay will vary
   with each employer.

2. Whether we consider better pay as a neutral aim or one with certain
   values attached to it has no importance as long as the method chosen for
   achieving that aim is chosen on practical premises, not based on any
   value analysis.

3. Better pay can simply be an explicit purpose for studying/teaching
   science; one the student knows and has named.

So if better paid employment is all a student or a teacher is after when embarking
on a science course, their aims and motives are extrinsic to the subject matter.

In a purely extrinsic view, education plays no part in the selection of the ends it is
to serve. The task of education would then be clear: achieve the given ends in the
best possible manner, i.e. in a way that produces the best results with minimal

effort/cost. So, if the aim for teaching science is to secure better pay for the students after they have been awarded their degree, the course should be designed to match employer requirements for new recruits and geared up for providing students with all the necessary skills for the higher paying jobs.

For educational aims to be extrinsic, it is not necessary that education as such does not participate in the selection of these aims. Educational bodies can participate in the choice of educational aims, but for educational aims to be extrinsic, they cannot promote any particular educational subject or curriculum for any other reason than as the best available method of achieving the already stated aims. For the aims to be extrinsic, one cannot promote science by reference to the value of science as such, but only as means to the identified aim. The same applies to choosing items within the curriculum - they should not be considered important in their own right, but instrumental for an external purpose.

Extrinsic justifications logically vary in how strongly the educational activity (X) is justified in that it brought about the extrinsic aim (Y). Downie *et al.* (1974) describe three levels of such justification: In the strongest sense of justification the occurrence of X is both a necessary and sufficient causal condition of Y. For example, the educational activity of learning a certain laboratory technique is both necessary and sufficient for achieving the extrinsic aim of being able to carry out particular laboratory tests. In a weaker sense, X is only a necessary causal condition of Y, but not alone sufficient. For example it is necessary to learn how to use certain equipment in the laboratory to carry out an experiment, but it is not sufficient, as one needs also to know how to handle chemicals and how to interpret the results in order to complete the experiment successfully. In a still weaker sense X might be conducive to Y - it may make Y more likely or it may be necessary for the best forms of Y, but it is not alone either necessary (Y can be achieved without it), nor sufficient (more than X is required to achieve Y). For example learning about the broader theory relating to the experiment may be conducive to completing the project, though it is neither necessary nor sufficient for that particular project.

In the case of moral education, I believe it to be necessary to include a fourth option - that X is a sufficient, but not a necessary condition for the occurrence of Y. This is a weak form of justification. Supposing that we are faced with two alternatives for reaching the aims of our ethics curriculum and doing X will enable us to reach those goals, but it is not the only available alternative; doing Z would also enable us to attain the same results. In such situation the claim that X was a sufficient causal condition of Y would not by itself give us a justification for doing X. Justification for doing X would then have to be supplemented with other relevant sets of criteria, like additional benefits and lower costs involved in comparison with those of Z.

These forms of justification of an educational activity apply to educational aims on two different levels:

1. Understanding the links between educational activities and the desired outcomes is a useful tool when deciding which educational method to adopt. In particular when the aims can apply to the entire teaching content of the curriculum. For example, what is the causal link between science education and higher pay? This type of justification is unlikely to have the strongest possible causal links, i.e. science education being both necessary and sufficient for highly paid employment. It is rather the weakest justification: science education is conducive to higher pay. One then needs to evaluate whether the link is strong enough to justify the efforts of science education in order to achieve this aim.

2. The causal links can also be analysed for assessing the best possible teaching method for achieving the aims. In this case the aim can be either intrinsic or extrinsic as the analysis is directed at methods of education and not the aims of education.

Competence has become a major element in discussions of educational aims. In most, but not all, cases competence aims are extrinsic. Educational establishments acknowledging and valuing employers' requirements for competent people are likely to design their curriculum so that it will produce graduates with proper competencies for the work place. Because these aims come from outside education

(the employers), and are often presented in a view of education as a neutral process in achieving these aims, they are best classified as extrinsic. However it is important to realise that the externally suggested aims of education are not necessarily extrinsic. An external body can suggest inclusions to a curriculum that promote skills that are intrinsic to becoming a member in a particular profession or discipline or intrinsic in the sense of personal development.

The idea of competence is not a new one in education. What is new is the particular idea of competence that now guides many educational aims - a shift which also makes competence an extrinsic aim. In the past, education in universities has been dominated by a notion of academic competence - a certain capacity to see the world and to engage with it through one or more academic disciplines, and to be able to engage in conversation in a particular cognitive tradition. This idea of competence is similar to the intrinsic aim of 'educated man'. In the current discussion competence is defined according to the needs and demands of employment, which are external to the actual activity of education, i.e. an extrinsic aim for education (Barnett 1997).

In summary, extrinsic aims of education do not promote any particular subject for its own inherent values. Education is viewed in general as a tool for some other, external, purpose. These purposes need not be mundane or material (e.g. earning more money), but for the aims to be extrinsic they should be identified externally to education, to fulfil some purpose not connected with the subject matter, and should be specific about these aims.

2.2 Intrinsic aims of education

The fundamental difference between intrinsic and extrinsic aims of education is that while the extrinsic aims refer to a specific aim external to the activity, the intrinsic aims refer to the benefits to the learner. In other words, intrinsic educational aims do not refer to any particular benefit or goal for which education is necessary, conducive, or sufficient, but rather reference is made to general benefits or goals in

life. Intrinsic aims refer to gaining abilities and knowledge that are good for learners regardless of the kind of life they choose to lead.

We need not consider intrinsic aims to be remote from reality or ignorant of the challenges of life or as vague and impractical. More often than not the successful pursuit of intrinsic aims produces some kind of external benefits. The significant element is that the educational efforts are not primarily justified by these benefits. For example, an ethics curriculum which is justified by reference to broadening the scope of the learner's autonomy (which is an intrinsic aim), may well also make students more valuable employees, just because they are more autonomous individuals. This may produce valuable social benefits (better pay, more secure employment or improved career opportunities). But importantly to the distinction between intrinsic and extrinsic aims, these benefits are not in this example the aims of ethics education, but rather incidental to the intrinsic aim of autonomy (Wringe, 1988).

Reference to intrinsic aims rather than extrinsic aims of education suggests a different approach to the challenges of life - that challenges are not in the primary sense specific problematic tasks within a profession but shared by all professions, and in nature encompassing life in general.

The role of intrinsic aims in higher education has been highlighted by the Dearing Report (1997). Two key recommendations in this report for higher education was to 'inspire and enable individuals to develop their capabilities to the highest potential levels throughout life, so that they grow intellectually, are well equipped for work, can contribute effectively to society and achieve personal fulfilment', and to 'increase knowledge and understanding for their own sake and to foster their application to the benefit of the economy and society'.

Ethics education can provide an example of intrinsic educational aims, in light of theoretical appreciation and the Dearing report. An intrinsic aim of ethics education could be a wish to encourage students' development as moral agents. This aim would be intrinsic (valuable in itself) if the attempt to encourage moral

development is not for any particular explicit reason. An explicit reason for ethics education could be, for example, that a better grasp of moral decision-making tools will provide a better opportunity for the students to find employment and therefore we should teach the students ethics. Another way of expressing this is to state that improved moral agency, when given as an intrinsic aim of ethics education, is not considered as a skill instrumental for any *particular* external activity or justified by reference to any specific benefits. Rather, it is valuable for the learners <u>whatever</u> they choose to do in life. It is learning for *no specific* purpose.

Educational activities worthwhile in themselves possess a number of features which clarify their nature. According to Wringe (1988), intrinsic aims are;

    1) held to be unending in scope,

    2) essentially non-competitive; even though capable of generating intense rivalries, they are essentially non-competitive in respect of their objects (people may make discoveries in their scientific field without depriving the others of an opportunity to do the same);

    3) they are challenging and provide opportunities for a variety of excellencies in seemingly limitless gradations of perfection,

    4) they entail both standards appropriate to the particular disciplines and the exercise of certain generic intellectual virtues such as intelligence, persistence, integrity, clarity, respect for evidence, and non-arbitrariness.

In general an intrinsic educational aim must have a wider perspective than teaching a particular skill, which can be mastered perfectly in a limited period of time. This definition follows the same logic as distinguishing education from training as dominantly different types of activities, a distinction discussed in section 1.1.1. Training does not, by definition, have intrinsic aims while education may have them, though this is not necessarily so. Some take the issue even further, and suggest that an activity should be called education only if its aims are intrinsic: all the rest should be called training. To Peters (1970), for instance, intrinsic educational aims are the only true educational aims.

Describing intrinsic aims as non-competitive and having a variety of excellencies emphasises them as activities of personal improvement. Activities that have intrinsic aims do not, at least not primarily, have explicit external reason for the improvement, for example winning/not losing. Also the process of improvement is not considered to be complete at any particular point in time, thus it is unending in scope. For example, an aim of improving moral decision-making skills, which can be an intrinsic aim for ethics education, could never be considered complete. There is always an opportunity to refine moral decision-making skills. One may choose not to try to improve these skills after a certain point, but there is no logical reason why one could not do so.

A core element in intrinsic aims is that they refer to abilities and knowledge that can be considered 'good' for the learner regardless of the life-plan that person has. The term 'good' in the intrinsic statement of educational aims in naturally value-laden. The theoretical and practical disagreement on the content of 'good' or the source of things 'good', does not necessarily threaten our description of intrinsic aims. As long as the description of 'good' remains a personally chosen value-laden attribute that describes the life one wishes to live, we can understand 'good' in general terms and our description of intrinsic aims is not tied to a particular moral theory. Further, the assumption is that there are some skills and abilities that would be considered good in most life plans and that education can have a role to play in encouraging the development of these skills and abilities in a fundamental way.

Even though we may be able to discuss intrinsic aims without committing ourselves to a certain value-system, it is often important, if not necessary, to be able to provide further explanations why a certain aim is intrinsically good. I will use studying Latin as an example. One possible way of providing the required reasons is to suggest that to teach/study Latin is worthwhile because Latin is an important language in Western civilisation. Naturally this reason may be challenged further- 'there is no goodness in knowing a language that is important to Western civilisation'. One can then attempt to widen the understanding of goodness in order to reach an agreement - e.g. 'teaching/studying Latin is important because it is part of Western culture *and* by learning Latin one can better understand it *and* it is good

to understand one's own culture'. This can yet again be challenged - 'why is it good to understand one's culture?' Maybe an even more fundamental justification could be found, but the process of justification cannot continue indefinitely. Ultimately the chain of reasoning must seek to link back to assumptions which the challenger cannot consistently deny, to some sort of transcendental argument referring to the basic values of being a member of a society, individual moral agent or a 'good' person.

Naturally there might never be an agreement on the core intrinsic aims, but in many cases the disagreement is sophistic rather than genuine. Also there might be great disagreement in how to succeed in fulfilling the core, and often very general, intrinsic aims, even when there is no disagreement about the aims themselves. This disagreement should not be confused with the true disagreement about the aims themselves. My assumption here is that it is possible in most cases to find shared values to support the intrinsic aims of education, at least within the limits of one culture, even one as multiple as Britain, when we might have strong and differing views on how to achieve them.

2.2.1 Values and intrinsic aims

Reference to intrinsic aims is most commonly a reference to a value-statement of something specific being good without any reference to its usage. For the purposes of defining ethics education clearly, I suggest that intrinsic aims make no reference to any particular value-system.

This is not a statement that intrinsic aims are value-free, as no educational aims ever are. Educational aims are always dependant on the society in which they exist. What is pursued through education depends on the personality ideals accepted as valid in a society and/or its sub-groups. Also, from a societal point of view education can be perceived as a means to securing the continued existence of the society and its culture; passing on a basic set of common ideas and sentiments, its basic moral attitudes and dispositions to act, and the special knowledge, abilities and virtues needed for maintaining productivity, and efficiency in all specialised

work domains. From an individual point of view, the value of education depends on personal aspirations, the opportunities society has to offer and how they are linked to education, and personal values. Which ideas, sentiments, moral attitudes, abilities, and virtues are the most dominant is dependant on the time and place.

It follows from this that the important elements within educational aims are established and valid independently of educators and their educational actions. Educators are neither creators of such ideals, nor are they free to arbitrarily choose among the theoretically possible ideals or set ideals actually found in the world (Brezinka, 1997).

This does not indicate that educators have no place in the process of establishing the values affecting educational aims: quite the contrary. Justifying educational aims is an important role in the constant re-identification of social values. The value-dialogue is influenced by educational efforts and educational efforts are influenced by this dialogue. It is therefore important to support this dialogue and to understand its two-way influence on the larger values of society.

In 1990's Britain, the liberal values of freedom, equality and rationality underpin most institutions, including higher education. These ideals may not be fully implemented, but they enjoy wide support as the ideals that should prevail. According to Halstead (1996) these ideals are:

1) Freedom includes freedom of action and opinion, freedom of constraint in the pursuits of one's own needs and interests. This can mean Benthamite utilitarianism of freedom to satisfy one's desire, or to realise one's rationally determined interests or simply to be oneself by being free of physical constraint. The value of freedom excludes totalitarian emphasis on communal unity to an extent that it endangers individuality.

2) Equality is respect for all individuals within the structures and practices of the society i.e. non-discrimination on irrelevant grounds. The value of equality, whether in a sense of formal equality of opportunity (Rawls 1972) or a more expanded idea of equalising life prospects or to distribute wealth and power (more) equitably. The value of equality

rejects slavery or classification of people according to race, gender or social status.

3) Rationality means basing decisions and actions on logically consistent rational grounds. The value of rationality excludes arbitrariness, inconsistency, and the failure to take account of relevant factors. As a value it rules out the uncritical acceptance of dogma.

Support of these values can form the basis of intrinsic aims in education, not because they are self-evidently 'the good values', but because they are accepted as the basic values of our society. There is a consensus, if not unanimity, that we should support these values and that we do not believe them to be good just because of some psychological quirk of our own. We do believe in their goodness, not as the last and unchanging doctrine, but as the firm value base of our society.

The research described in this thesis is set within this value structure. The parameters of this value system are not static and the educational system is not without an opportunity to influence the cultural ideals within which it functions. The process of developing higher education is therefore partly a response to external demands and evolving ideals of personality and outlook of the world, but it also has a role in moulding these ideals and affecting their course.

The teaching material resulting from this research has therefore two kinds of aims. First it is a clear response to the demands placed on higher education by society to encourage student development along these ideals. Second, it has an ideal of developing the skills of the current student generation towards the ideals and values held important within the teaching community. This new type of teaching is intended to influence how students approach their work and through that link influence values in society at large.

2.3 Summary

In this section I have concentrated on one aspect of aims, whether they are extrinsic or intrinsic. This element of educational aims is important when we investigate the

justification for educational resources. Intrinsic and extrinsic aims function differently in justifying educational expenditure. Understanding their nature allows us to provide the most suitable justification for our educational efforts.

Extrinsic aims of education have three main characteristics:

1. They are external to the actual educational activity.

2. Education is considered to be a neutral contributor to the set aims.

3. Education is an instrument for achieving particular aims, not a source of general good.

In training, as described in section 1.1.1, most of the aims are extrinsic, because training as an activity seeks to develop skills and abilities for a particular purpose. Education in general can also be viewed as an instrument for an external set of aims, though that might reduce education to a role that is not wholly acceptable.

Alternatively, intrinsic aims are fully involved with the educational activity itself. They are aims which are considered good in themselves, valuable for no *particular* external and explicit reason. It may be that achieving these intrinsic aims will produce quite significant positive (material) side-effects, but these are only contingent to the educational activity. Intrinsic aims concentrate on the person, not what that person can do.

Based on the above description of aims, it is easy to see why the term 'educational aims' is so frequently used in assessing the justification of education or different educational practices. "What is the course for?" and "What will the students learn" are most accurately answered by reference to intrinsic and extrinsic aims. These answers clarify the scope of our educational activities and whether the justification we can provide is either necessary, sufficient or conducive to the aims we have set for ourselves. Asking questions about the aims of education is a way of providing clarity and focus on what worthwhile aims are.

3. Aims of Ethics Education

One of the very first decisions to be made about an ethics course is to find an answer to the question 'why this course' or alternatively 'what are we trying to achieve with this course?'. The answer to these questions indicates whether the chosen aims are intrinsic or extrinsic.

In this chapter I will discuss the possible aims of an ethics curriculum. I will start with extrinsic aims in section 3.1 and continue with intrinsic ones in section 3.2. As will become clear, I believe that the aims of an ethics curriculum should be intrinsic. This is not to under-value the external benefits of such educational activities, which I believe to be considerable. The reason why I believe, and hope to show in this chapter, that the true aims of ethics education should be intrinsic is in its general approach to a person, respect for autonomy, and dignity. Also, the practical problems that arise from choosing extrinsic aims for teaching ethics indicate that they are unsuitable.

3.1 Extrinsic aims of ethics education

One of the core issues for extrinsic aims is that they should be clear and definable, often even measurable. The possible definable and measurable aims of an ethics curriculum are 1) possessing a certain opinion or value and an ability to demonstrate that value when requested, or 2) 'ethical behaviour' - a tendency to act in a certain way in response to a specific stimulus. Extrinsic aims are therefore directly involved in the behaviour of an individual or co-operation between individuals, and holding certain 'right' moral values.

Most often, holding a correct view, value, or opinion is not sufficient, but the correct behaviour becomes the most essential element in extrinsic aims for an ethics curriculum. Behaviour is the external sign of mental processes and therefore also of the values held. Behaviour also has a direct impact on social functioning and altering social contacts. It is therefore sensible to discuss the extrinsic aims of

ethics education by including some aspects of behaviour, most notably how education affects behaviour and what type of behaviour is desired. Behaviour can be either actions in a particular situation or more general behavioural patterns applicable to a larger set of possible situations. The following analysis of extrinsic aims of ethics education concentrates thus on behaviour modification.

Extrinsic aims for ethics education can be of three different overlapping types: 1) those concerning the individual, 2) those concerning morality or other phenomena generally assumed to be morally good, and 3) those concerning the good of society (Downie *et al.* 1974). Extrinsic aims of ethics education concern dominantly the success of a student in both personal and work life. Such aims could be, for example, better income based on skills gained through ethics teaching. Aims concerning morality or other phenomena assumed to be morally good relate to activities of a community. Such aims could include speaking the truth or the ability to understand moral problems and to solve them in the best possible manner. The last type of aim - those concerning the good of society - relate to general views on a good society and how it can be best achieved. Such aims may include the ability to understand the moral grounds for the laws of the state and a desire to adhere to them. All these types are overlapping as the aims that are dominantly individual also contribute to the aims that are dominantly social or to those of general morality. The dominance refers to the most direct beneficiary of certain educational aims, while all types are inter-linked and advances in each promote goodness in the others as well.

A call for moral education is often formulated as a need to teach students the difference between right and wrong, often in a sense that the students are then expected to carry this knowledge over to their actions. In other words, the aim is to produce individuals who know what is right and what is wrong and also behave according to this knowledge by doing what is right and avoiding what is wrong. This is a consideration of mainly general morality and socially beneficial aims. While the individual may benefit from the achievement of these aims via better adaptation to society, the benefits are not direct.

This type of moral behaviour education involves three tasks:

1) identification of the appropriate values.

2) the transmission of these values to students (Halstead, 1996),

3) ensuring behavioural patterns in students, identical to the transmitted values.

The task of moral behaviour education proves problematic on several accounts. The first problem is the genuine lack of agreement over the definition of what is right and what is wrong, i.e. what are the appropriate values. There is no shortage of opinion of which values should be the most important, but little agreement between them. Even if people might be able to agree on general moral values - like the value of human life or the importance of justice - this consensus is often dispersed at the very moment someone tries to apply the value to an actual moral problem. Agreeing that justice is an important value is different from trying to decide what justice means when a doctor needs to choose between two patients for kidney dialysis - are we looking at a just method of choosing (first-in-first-served, roll of dice), which would then result in just allocation of scarce facilities or are we more concerned about the results of the selection. In other words, in trying to choose between an old man and a young woman - what criteria should we use in choosing? Would age, social status, medical prognosis, gender, or luck be the decisive factor? There are no easy rights or wrongs in most real-life situations.

Therefore stating that 'teaching the students what is right and what is wrong is the aim of an ethics curriculum' is often not sensible in the situation of multiple interpretations of the same value. Stating that this course teaches the students what is right and what is wrong would be an ambiguous reference to a consensus that does not exist. Not only would it be undermining the plurality of value judgements, but it would also be creating false expectations of what the course contains. One could avoid the latter problem by going to the lengths of stating what are the rights and wrongs to be taught during that course - a difficult task in itself - but this would not give any guarantees that others would be willing to accept those value interpretations as the appropriate rights and wrongs to be taught.

The behaviour approach to education is further likely to fall into the problem of a static view of values. If the aim statement has a right (wrong) value approach, the right and wrong may become fixed and rigid concepts, that are not open to redefinition in the face of a novel or difficult moral situation. This does not correspond with reality where right and wrong are re-defined and re-justified as a normal course of social change. Hence, developing an ethics curriculum with an aim of teaching students what is right and wrong, is an approach almost guaranteed to be one of great controversy and difficulty.

But assuming that a list of values to be taught was agreed upon, the next difficulty would be encountered when trying to transfer these values to the students. Aiming to persuade people (or plainly force them) to adopt certain values is likely to be classified as indoctrination, which is in direct conflict with one of the basic values of liberal society, i.e. freedom of opinion. Indoctrination is not an appropriate educational aim, as was discussed in section 1.1.2.

A further problem with the extrinsic approach to ethics education is the disparity between what one knows to be right and one's actions. It is not straightforward to believe that people always do the right thing as long as they know what the right action is. What people state as their values and how they actually behave seems to have only a loose correlation. Further, verbal agreement with general value statements does not necessarily indicate agreement with a policy statement intended to represent those values (Lockwood, 1976).

A significant amount of research has been carried out in order to verify whether values and actions do or do not correlate. For example, moral development scores (Defining Issues Test scores, more of which in Part II section 4.3) and behaviour have a pattern of significant, but only moderate, relationship. This suggests that other variables also determine behaviour. Moral judgement scores are, rather, linked to behaviour from the perspective of different moral judgement representing variation in the ways people construe situations, which then leads to different actions (Thoma and Rest, 1986).

This problem has been well known from the start of Western philosophy. In ancient Greek philosophy it was called akrasia, meaning the weakness of will. Socrates famously denied the existence of akrasia: 'no one does wrong willingly'. Aristotle on the other hand commented that this statement is 'manifestly at odds with observed facts'. Aristotle continued his argument by stating that those of real knowledge do not do wrong, but only those whose knowledge is blurred by the seeking of pleasure and the like (Rowe 1995). If Aristotle was right, intuitively, and by social observation, it seems that we all have only blurred knowledge!

One of the possible explanations for this lack of strong positive correlation between moral judgement and moral action is that the making of moral judgements is not solely an intellectual activity. Weighing pros and cons is a cognitive process: the conclusions emerge on one side or the other whether we like it or not. These conclusions carry with them an attraction towards the type of action they approve or a repulsion against the type of action condemned by them. Therefore the cognitive weighing of alternatives falls short of an actual decision to act accordingly. To translate moral knowledge into action requires various qualities of character, depends on the circumstances, courage, self-control, perseverance, concern for others, love of justice, strength of will and so on (Downie *et al.* 1974). An ethics course with an objective to build personal character, encourage concern for others, love of justice, strengthen personal will, as well as to teach the students what is right and wrong, how we should love justice (let alone understanding what justice is!) and so forth sounds more than an ethics course. If these are the aims, some truly good and honourable, it would be incredibly hard to imagine a course structure that could even hope to achieve most, if any, of these aims.

A separate problem from adopting behaviour and/or character change as the aim of ethics education is, how to measure the success of such educational efforts. Measuring the effects is particularly important when the educational aims are extrinsic, because if aims are external and specific, there is likely to be a need to know how well these aims were achieved. Whether the need for justifying the success of the educational efforts comes from external pressure of funding bodies, from internal quality assurance, or just from an interest to know how well a course

did achieve its aims, it would be extremely difficult to provide evidence that an ethics curriculum has made a substantial impact on the moral behaviour of students. It might be possible to present students with moral choices and monitor their behaviour in a laboratory, but this is unlikely to be satisfactory if the aim was to change behaviour in general, in all situations and in the long term. Unfortunately, behavioural changes are difficult to measure in real life situations. It is impossible to follow students outside the classroom extensively enough to verify that the effects of ethics education on behaviour are permanent and substantial. This is not just due to the apparent practical problems of recording and evaluating magnitudes of behavioural situations, but there would be difficulties in classifying what is right and wrong behaviour in each situation. Also, we should be concerned with the implications that such large scale monitoring of student behaviour have for the rights of students to privacy and freedom of opinion.

A further problem for an ethics curriculum of extrinsic aims arises if the aims are chosen mainly from the social point of view, to teach the students certain values and modes of behaviour because it is considered to be beneficial for the society or some members or groups in the society. This approach includes a risk of using education as a means to turn students into building blocks of a certain type for a society they have not chosen themselves. In other words, students might become the means to an end and cease to be ends in themselves. This is a Kantian notion of people not to be ever treated simply as means, but always as ends. Modifying a person's set of values and aiming to alter their behaviour for social benefits, is using the entire person for the good of something else than the person him/herself. This is far more than asking a person to do a particular deed for the common good. Aiming to influence a person permanently, to change and choose the values they hold and to expect them to behave according to some predefined pattern, is not compatible with what we normally understand when we talk about personal autonomy and integrity.

In the literature, we can find moral educators supporting the above conclusion that changing student behaviour or values is not an appropriate approach to teaching ethics. Callahan (1980) concludes that: "A legitimate goal in the teaching of ethics

is to help students develop a means and a process for achieving their own moral judgements. If 'moral education' means something more than that - an education in specific moral rules, or specified traits of character - then it is illegitimate." (p. 71).

Because of both value-based problems and problems in measuring the success of an ethics curriculum when the chosen aims are behavioural and/character modification, I will not use them in my set of aims for this project. This leads to exploring intrinsic aims as a justification for ethics education.

3.2 Intrinsic aims of ethics education

Intrinsic aims for an ethics curriculum take a sharp diversion from extrinsic aims by rejecting any particular set of values or moral behaviour as appropriate. These aims cannot be considered good in themselves - at least not in the sense described earlier. Through the analysis of extrinsic aims for ethics curriculum in section 2.1, it became clear that particular values are rarely agreed upon, at least beyond a very general description, and therefore they do not fulfil one of the basic criteria for intrinsic aims - common agreement on their goodness. Again, agreement may be possible regarding general value statements, and the impossibility of agreement refers to more detailed moral statements and interpretations of values. We could logically base intrinsic aims of an ethics curriculum on a general set of accepted values. Unfortunately the positive outcomes of such a course would be limited, as learning of general values per se would not differ significantly from any other cognitive learning of concepts and thus offers no other benefits than better understanding of conceptual intricacies. For these reasons the intrinsic aims of ethics education are poorly understood as the outcomes of cognitive pursuits. Instead of prescribed outcomes, an ethics education can (should?) concentrate on the *process* of moral decision-making.

A possible, and here adopted, intrinsic aim of ethics teaching can be the process of moral decision-making - to encourage the moral development of the learners in making moral decisions, supporting them in developing tools to analyse and solve moral problems and creating opportunities to practice these skills. The emphasis

PAGE NUMBERING AS IN THE
ORIGINAL THESIS

society, a 'good' person would be one who refrains from moral inquiry and decisions. Decisions regarding human interactions and public good would then be made by the leaders, without consultation or need for approval from the citizens involved. Such a society might be a well functioning and a pleasant society, but it would not be a democratic one.

Naturally there are values involved here as no educational activity exists in a neutral, value-free, environment. The values supported here are respect for personal choice of values, freedom of opinion and the importance of self-chosen and rationally defensible values. These values are construct values, not content values, which makes them more defensible within our own social climate. Insofar as democracy means rule by the people, the self-directed choices of citizens should ultimately guide public life. If this requirement is taken seriously, democracy presupposes citizens who are capable of forming authentic convictions - convictions which the person can identify and be conscious of the process that led to that identification - which they can bring to the guidance of society. Therefore, citizens should become conscious of the processes that have affected their value-choices, and develop the capacity for evaluating them as well as those prevailing in society as a whole. Because ability to form authentic convictions is not an inborn talent, the idea of democratic society presupposes an education which fosters independent reflection (Puolimatka 1997). The ethics education described in this thesis holds its place as part of the larger educational effort to encourage people to develop their skills as independent moral agents.

3.2.1 Nature and measurement of moral development

An inherent element in education is the aim to improve, to encourage progress and development - "[I]t would be a logical contradiction to say that a man has been educated, but that he had in no way changed for the better" (Peters 1970, p. 25). Moral education is no different. The aim is to create improvement and in the case of intrinsic aims, improvement in the abilities to make moral decisions - to encourage moral development.

The core of the moral decision-making process is how one perceives moral situations and the method(s) one uses to decide between the perceived courses of action. Particular skills are involved in each step of moral decision-making and moral development refers to the improvement in the use of these skills and abilities.

The approach to moral development adopted here is based on the theories of Kohlberg (1976), Perry (1999) and Gilligan (1982). The fundamental feature of each theory is that moral development is viewed 1) as a linear process where moral abilities can be considered to exist in stages, and the development from one stage to the next is invariant - one can progress to a stage only by passing through the immediately previous one. 2) Development is not automatic - it is possible to delay progress to the next logical stage (even indefinitely) and to regress to a lower stage. 3) This process is considered universal in the sense that the developmental structures of moral decision-making are not dependent on the actual moral decisions made, but on the tools, skills and approaches adopted at each stage.

The theory of moral development is discussed here as a meta-theory - one that provides a framework to discuss the nature of moral development and theories within the discourse. The nature of moral development theory, and the nature of its practical applications, can be illustrated by comparing it with the theories and tests of IQ (Intelligence Quotient), which provides a good point of contrast despite its problematic nature. My comparative analysis follows that of Gibbs and Widaman (1982).

1) Items in IQ tests are chosen principally for their statistical properties and not for their structural or contextual qualities, while the test items in moral development are fundamentally concerned with the correspondence of the test items with the objective nature of moral development.

2) The aims of IQ and moral development tests are different. Where IQ tests seek constant variability of scores *between* individuals, moral development tests measure the maturity of a *particular* individual and in longitudinal studies, the moral development process of *one* individual. Large cross-sectional samples are typically used to establish 'normal' or

average intelligence, while in moral development tests, large samples are used in order to improve understanding of the process of development.

3) An individual IQ score is defined and evaluated relative to average scores, which are often age-related in childhood and stable once in adulthood. IQ scores are also estimated to remain stable in adulthood when one's 'predetermined' IQ has been reached. In moral development, an individual 'score' is used to signify the stage of one's development, which is not hypothesised to be static or at least no concept of predetermined stage of moral development is hypothesised. The aim of IQ testing is, hence, to establish individual differences and moral development tests aim to establish the progress of one person within the individual's realm of development.

The basic questions in moral development research are 1) what are the fundamental processes of moral decision-making, 2) whether there are better or worse ways to make moral decisions and whether we can describe moral development as a process where one rejects worse methods of moral decision-making and simultaneously accepts better ones, 3) can education improve the decision-making process, assuming improvement is logically possible? The last question is included in the research hypothesis of this thesis. I will return to the first question in Part III section 6.4, and concentrate here on the second question - is moral development a logically sensible discourse?

The basic idea in most moral development theories is that a person passes through several moral stages in a particular order, and that not all will necessarily reach the highest stage, but can remain even on one of the lowest ones. This development is not simply a representation of learning a longer and more complicated list of rights and wrongs, but a matter of developing fundamentally more sophisticated conceptions of how moral reality is organised and how to make moral judgements within it.

The three most influential theories of moral development can be seen to describe the same development from different perspectives. Kohlberg concentrates on the

understanding of fairness in hypothetical situations, considering the decision-maker as an independent moral agent. The conception of fairness translates to an understanding of rights and duties - their source and content. Kohlberg's theory deals with 'macro-morality' - questions of the morality of the society and its institutions - not with 'micro-morality' questions relating to everyday life face-to-face situations (Rest *et al.* 1999a).

Gilligan, whose theory was developed as a response to Kohlberg's, emphasises the contextual issues of moral decision-making, an individual within relationships. If the determinants of moral obligation for Kohlberg are principles of justice, for Gilligan they are relationships. The moral conflict for Kohlberg is embedded in the rights of the moral agent and the rights of the other, while Gilligan explores the moral conflict between hurt and care - a reflection on inevitable moral choices that do not inflict only good, but pain as well. In philosophical literature we can find several other approaches/perspectives on moral development.

Shapiro (1999) has developed a model, one he calls a moral spectrum, which describes different moral decision-making methods called prisms. Within this spectrum, Kohlberg's development of principled reasoning encompasses deontological, utilitarian and existential prisms. In other words, it uses a decision-making formula not tied to self-benefit or relationships. The ego prism in Shapiro's model would be included in the lower level reasoning in the Kohlberg model. Gilligan's approach co-incides with prisms of ethics of care and communitarianism. The remaining prism, virtue ethics, is an alternative approach for which moral development models have not yet been developed, but with which character development aims coincide.

Perry's approach, on the other hand, concentrates on the understanding of moral reality - the source of rights and duties, information and values. It is less concerned with the relationships between people (the core element for both Kohlberg and Gilligan), and more on the inner structuring of reality within one individual. The development in each approach entails an integration of rights and responsibilities.

Each moral development theory concentrates therefore on one aspect of moral decision-making - each describing a vital part of the process. Perry's model is an umbrella over the field of moral decision-making. The basic understanding of moral reality is often a pre-requisite for more specific aspects of moral development.

We encounter moral problems both in our private and public lives, and a different understanding of the problems is required. There is no logical reason why both personal (care orientation) and global (justice orientation) cannot be introduced, where relevant, into one and the same reasoning episode. It is possible to imagine a person who is both caring and perceptive of the needs of others and self, and who in addition has finely toned sensitivities for perceiving moral salience in a global sense and perceiving moral problems of great diversity (Flanagan and Jackson 1993).

The progression of moral development according to all three theories is very generally a move away from a dualistic world view, where things are 'black and white' and often where it is me (good) against others (bad). The lower levels in moral development are also characterised by dependency and the need for authoritative figures as the source of answers and values, and a limited scope for social effects - often only the immediate persons are considered. At the start of moral development, people are also incapable of dealing with multiplicity and relativism of facts and unforeseen consequences. Moral choices are based on very concrete indicators, like the likelihood of getting caught and an ability to win a physical encounter. The moral skills required in the lower stages are relatively basic. Only a limited amount of moral thinking is necessary when rules are accepted as given and only minimal implications are considered.

In the middle stages the 'black and white' dualistic view is supplemented by a recognition of greys. Life is no longer as clear-cut, but the source of moral codes is still not considered to be the individual, but rather the society. One is part of a system and following rules is an important part of being a member. Even if there are no absolutes, authority is now considered an authority who chooses within the

relativistic reality. Also, the scope of moral stakeholders expands: one is able to see the wider consequences of moral actions and is interested in them. Often the balance changes dramatically from selfishness to the denial of personal needs and the value of personal desires. The system is, however, still rigid and personal moral commitment to self-chosen values is more or less lacking. The moral skills required are more demanding: one needs to observe and deal with multiplicity and relativism, seek to understand the wider consequences of moral actions and find reasons for moral rules other than immediate (physical) threat.

In the last stages of moral development relativism and multiplicity are accepted as inseparable characteristics of the world, and the individual moral agent is as able as the authorities to make judgements in the moral sphere. One is interested in the wider consequences of moral actions, taking into account people both physically and temporally distant, and seeking to understand both indirect/direct effects of moral actions. The interests of self and others find a balance. The moral skills required are complex because one needs to be able to form and support individually acquired moral convictions and judgements. One needs to be personally able to form logical and coherent moral arguments, and have an ability to foresee several different types of moral consequences.

Not all agree with the idea that Kohlberg, Gilligan and Perry describe essentially different aspects of moral development, but in general understanding, for the purposes of designing an ethics curriculum, I believe this simplification and unification is both defensible and sensible. For further discussion on how these theories overlap, please refer to Appendix II: One Theory.

3.2.1.1 Problems with moral development

One of the problems with this description is that moral development includes aspects of improvement, not wholly a neutral cognitive development, but one that involves personal attributes as well - courage, will, and effort. An implication of this might be that a morally 'advanced' person, showing either faster development

through moral stages or holding an advanced position relative to those of his/her age becomes in a more general sense a 'better person' (Perry, 1999).

But we need not approach moral development by making a normative judgement on the goodness of a person according to their level of moral development. A person moving from a lower stage to a higher stage may be said to be developing not simply because the adopted assumptions are 'better' or more 'true', but because the forms of the later assumptions build upon the earlier ones in a coherent manner. This cannot be said in reverse. This is a skill-based approach. The basic assumption is that people can develop these skills, and that they are able to progress through the moral development stages: all we need to do is to give them opportunities to do so. A person's moral development stage does not, therefore, need to be a normative judgement, but an indication of the personal stage of development, information which can, and should, be used for the benefit of the person whose developmental stage has been recognised. Rest *et al.* (1999a) emphasise that there is not just a difference between schemas, but also a relative developmental advance in moral thinking, because one conception is a cognitive advance over another.

We can call this movement through stages development in a skill-based sense for three reasons (Rest 1983):

1. Each stage in the sequence is progressively more differentiated and integrated.
2. Each new stage employs cognitive operations which are more diverse and equilibrated.
3. Each stage has a more encompassing view on society.

Even if we appreciate a certain relativism of moral values as inevitable, it seems logical to believe that the moral decisions made by a person with an ability to perceive the wider implications of moral actions, and by one who is capable of self-directed thought and of forming logical and coherent judgements, are better than those made by considering only immediate implications and believing in absolute authorities. In any sphere of human development, perceptual, intellectual, social,

emotional, moral and so forth, the implications of development suggest that it is better to grow than to arrest growth or to regress.

The assumption is that the values built into this scheme of moral development are commonly held to be significant. I would argue that it is better for each person to progress to more advanced levels of moral development, that it is better for each person to have a wider understanding of the moral implications of decisions, and better for others as well. Since there is no intention within this research to rank students according to their moral development, to make it a classificatory aspect within the educational establishment, the implications of developmental testing should not be negative on students. The basic assumption is that within the university student population there are very few individuals with capacity restrictions for reaching the highest levels of moral development, and the encouragement of moral development is aimed at the personal improvement of the individual student, not as a measure of effort or success.

Therefore, if we can agree that the higher stages of moral development are preferable to the lower levels, we can also agree that the promotion of moral development is an acceptable and desirable aim of ethics education.

If we accept encouraging moral decision-making abilities as the aim of an ethics curriculum, we need to look at the aspects of moral decision-making in more detail. There are two areas of interest, namely:

1. Moral Sensitivity - an ability to perceive the ethical implications of a situation. It is essential in any situation to be able to identify the moral aspects. Without the initial recognition of moral facts alongside scientific or 'hard' facts it is impossible to make any moral decisions. Moral sensitivity is also about an ability to understand the moral networks and implications of moral actions. Moral sensitivity is also called 'moral consciousness'.

2. Moral Competency - an ability to engage in sound moral reasoning and practical problem solving strategies. The person must be able to make a judgement about which course of action is morally right (or fair, just,

morally good, or adequate) and thus label one possible line of action as what a person ought (morally) to do in that situation. Moral competency is also called 'moral cognitive skills'.

3.2.2 Moral sensitivity

Moral sensitivity is the first step in moral decision-making. Without recognising the moral aspects of a situation, it is impossible to solve any moral problem, for without the initial recognition, no problem exists. Also without an analysis of the moral aspects, it is impossible to move onto making decisions, at least if the decisions are to be made with awareness of the magnitude of the problem and the effects of the decision.

Moral sensitivity is a combination of two different abilities; moral imagination and recognition of ethical issues. Callahan (1980) has described these two aspects in the following way:

1. *Moral imagination* is an ability to perceive a 'moral point of view' - to understand that 1) human beings live in a network of moral relationships, 2) consequences of moral rules can be either happiness or suffering, 3) moral dimensions of life can be hidden or visible, and 4) moral choices are in most cases inevitable and difficult. Moral imagination is thus an ability to see the moral side of the story and an ability to project on the moral consequences of actions. It is like imagination, because it requires one to 'see' something that is not real in a sense that we could touch it, but something that is real in our minds and within our social existence. This level of understanding can be considered as a pre-requisite level for any moral discourse. It is necessary, but not sufficient. Without moral imagination we are not able to engage in discussions on ethical problems. But to have only imagination is like being able to see, but not to act or speak, to be only a passive perceiver of things, but not an active player within the moral network one can see. Therefore, just to have an ability to perceive moral problems is insufficient for making moral decisions, while it is an

essential part of that process. In order to solve moral problems one also needs an ability for conceptual and logical analysis.

2. _Recognition of ethical issues_ is closely linked with moral imagination - it is moral imagination put into action. If moral imagination is an ability, recognition of ethical issues is the application of that ability. It is an attempt to analyse what has been seen, to recognise the value of moral aspects in a particular situation. Recognition of ethical issues is to distinguish between emotional responses to situations and appraisal of facts, moral or scientific. This type of recognition requires an examination of concepts, moral statements, and ethical and moral rules. To be able to recognise ethical issues in this way, one needs to be aware of the moral categories, of the aspects that can be classified as moral and to be able to evaluate their importance to a particular situation.

To be able to imagine and recognise moral problems one needs an ability to accept and recognise multiplicity of the moral reality. The development of moral sensitivity is linked to the developmental positions described by Perry (see Appendix III - Perry's Scheme). The Perry Position 5 is a pre-requisite for appreciating the variety of moral reality, though the ability to do so does not guarantee that one will apply the skills in practice.

After implementing both the skills of moral imagination and recognition of ethical issues, a person should be aware of various courses of action and what are the consequences of each action. This is an acquired ability, _not_ an inborn talent. People may differ in their natural sensitivity to moral problems, but in general moral sensitivity is acquired by exposure to moral problems. Because moral sensitivity is not an inborn talent, it is possible to cultivate it and improve one's ability to understand the moral aspects of a problem in increasing adequacy and precision. Moral sensitivity is an important skill and the difficulties in interpreting a situation as moral and in understanding the implications of moral actions should not be underestimated.

While moral sensitivity has a strong cognitive component, it is not only an intellectual faculty. Recognition and interpretation of moral aspects are also

dependent on situational clues, personal attributes, and affective responses. Rest (1986) provides the following list of the aspects interfering with a rational and clear perception of moral aspects in context:

1. First, people may block from their consciousness certain aspects because the cues in the situation are ambiguous and it becomes difficult to interpret them. Inability to interpret social situations in terms of realising how one's actions influence others is not necessarily defensive, but rather a psychological mechanism to deal with difficult and ambiguous information.

2. Second, research shows that there are distinct individual differences in sensitivity to needs and welfare. For some, blood needs to flow before they recognise the plight of another person, while the other extreme is an oversensitive person for whom every grimace and word takes on momentous moral implications.

3. Third, research has shown that there can be a strong affective response before extensive cognitive encoding. We can have instantaneous dislike for someone or empathy towards another, before we have had any opportunity to analyse the situation. These feelings are not independent of cognition, but they are more primitive cognitions and their accompanying feelings can proceed without waiting for considered reflective judgement and careful weighing of facts. It is important that these first impressions are clearly recognised and that they may not be the most appropriate affective moral reactions in given situations.

Moral sensitivity thus involves identifying what we can do in a particular situation, understanding the consequences to all parties of each line of action, and identifying and trying to understand our own gut feelings on the matter.

Taking into consideration these psychological aspects of recognising and analysing moral situations, an ethics curriculum can proceed to support student development in learning to confront these situations with more reflective thought and control of their initial emotive responses. I will look at the most appropriate methods in Part II.

3.2.3 Moral competency

Making moral choices is not a matter of pick and choose what seems to suit our aspirations in each situation. If we are interested in proper moral arguments and high levels of moral development we cannot use ethical arguments to support our own preferences or opinions unless they coincide with coherent and rational moral argumentation and decision-making models. If we are to take moral reasoning seriously, we cannot allow ignorance of relevant facts, or illogical use of arguments.

As I have discussed before, to be good at morality does not refer to holding the right views. This is similar to being a good scientist, which does not mean that the results and hypotheses of good scientists are always right, but that they observe the physical world closely and patiently, frame hypotheses intelligently and submit them to experimental tests, and are prepared to abide with relevant evidence. Similarly being good at morality is about the procedures one follows in making a moral judgement.

Coherence and consistency are minimal requirements, both in the analysis of ethical propositions and their justification. How deeply these matters of analytical skills are considered in an ethics course is dependent on the situation, the maturity of the students, time available and the overall goal of the course. In any case, something must be said about the importance and nature of these skills (Callahan, 1980).

Ethical enquiry has its own procedures and rules not unlike those in natural sciences. We have a fairly good idea of what rules and criteria govern the activity of science and how scientific arguments and justifications are supposed to work (see for example Resnik, 1998). This does not mean that all scientists follow these rules, only that they exist and that we expect scientists to understand them. In a similar way, it is possible to construct rules and criteria of moral inquiry. The most familiar and fundamental rules of moral argument are shared with other activities, like science (Wilson, 1990):

1) We should adhere to the laws of logic.

2) We should use language correctly.

3) We should attend to the facts.

Adhering to the principles of moral inquiry does not mean that the moral judgement produced after following the rules of moral decision-making will necessarily be correct. The same applies to science; following the rules of science does not guarantee that results would be necessarily true. We may have overlooked some vital fact, or our instruments may not be good enough, and so on.

The fundamental rules of moral decision-making include special skills:

1) one needs to learn to recognise illogical and logical moral arguments. The elements are similar to those of scientific arguments, but not identical. One needs to learn to evaluate moral premises and the thought processes that follow.

2) one needs to learn to separate moral elements from emotional responses and from scientific facts.

3) one needs to be able to evaluate premises and foresee how different logical arguments can be formed by the choice of different premises.

If we believed that moral decision-making is mainly learning a logical problem-solving syllogism and applying that to each case, then we would teach in our ethics curriculum the structure and usage of the syllogism and how to find the necessary facts to apply the syllogism to. This does not seem like a reasonable assumption, if only because such a syllogism has not yet been found, regardless of how hard it has been sought. The basic understanding of logic and knowledge of the facts is therefore important for an ethics course, but pure logic is not enough for proper ethical reasoning.

The most complex skill in moral decision-making is an ability to combine the moral point of view with the limitations of reality and to tolerate the ambiguity of moral situations. One is often forced to choose between two morally unsatisfying options,

to compromise between some basic values, or to accept harm in the presence of greater good. If moral choices were always choices between good and bad, the importance of moral education would be in analysing what is good/right and what is bad/wrong and assist the development of commitment for doing what is good/right. But because reality is more complex, the essential skills need to include an ability to weigh different alternatives, to understand personal values and form logical reasoning steps which provide answers compatible with reality and supporting basic values.

The issue of commitment is part of moral competency. This is commitment to a personal reasoning process which entails acceptance of responsibility for moral choices. This is, according to Perry's scheme of development, the last stage - the recognition of moral relativism is succeeded by a realisation that commitment is an ongoing process of identification with certain values. In the context of this research, this is not a reference to an affective commitment to certain values, but an appreciation of the implications of value choices - freedom to choose is paired with a responsibility for one's choices.

3.2.4 From ability to action

The aspects of moral development described above - moral sensitivity and moral competency - are the first two components of moral decision-making in real-life situations, where actions need to be taken regarding moral problems. The other two components are: 1) the person must give priority to moral values above other personal values - desire to advance career, love, art etc. - such that a decision is made to intend to do what is morally right even when there is a conflict between moral and other personal values, and 2) the person must have sufficient perseverance, ego strength, and implementation skills to be able to follow through on the personal intention to behave morally (Rest, 1986). In other words, in addition to being able to recognise moral problems and decide what is morally the best solution to a problem, one needs to have motivation and ability to implement the results of moral reasoning. A person who demonstrates great ability in one component may not show great ability in another. The components are therefore

independent - ability in one does not indicate ability in another. This independence does not indicate isolation of these components, quite the contrary - each component influences another: the links between them are various and strong, but they are not automatic or identical.

Encouraging motivation to do what is morally right may sound like an appropriate educational aim, at least as long as we hold onto the importance of self-chosen values and thus avoid criticism of indoctrination. Still in this research moral motivation is not included in the aims of ethics education. The lack of consensus on what accounts for the motivation to prioritise moral values over others is one of the reasons for the exclusion. At least eight different perspectives can be identified in the literature (Rest 1986):

1. evolution has bred altruism,
2. there is no special motivation and people merely respond to social reinforcement,
3. shame, guilt and fear motivate morality,
4. empathy is the basis of altruistic motivation,
5. experience of living in loving and caring relationship leads to moral commitment,
6. understanding the interactions in the society motivates an individual to act morally in order to enhance an ideal society,
7. sense of awe and self-subjugation to something greater than oneself motivates moral actions, and
8. concern for self-integrity and identity as a moral agent motivates moral action.

Seeking an educational approach to encourage moral motivation would therefore be plagued with theoretical problems of choosing one, or a combination of, motivational theories - how to justify the choice, and how to interpret the results. Also, aiming to influence a personal aspect like motivation - a motivation to do what is morally right instead of what might otherwise be preferable, even when moral rightness is self-chosen - may be a questionable aim in a multicultural university. Lack of agreement on what encourages such motivation is one problem

and aiming to influence something so personal and something that has clear external implications, is moving again towards extrinsic aims, which were rejected as appropriate in section 3.1.

Similar reasons apply to rejecting the inclusion of the last component into the aims of ethics education. As was also discussed in section 3.1, good intentions are often a long way from good deeds and psychological and religious theories are abundant in explaining this incongruence. Again, an attempt to influence ego strength, or whatever we wish to call it, is prone to the same criticism as influencing motivation.

Still, we may agree that it would be desirable if people were both motivated to act morally and had the personal strength do so and because of this agreement, we might be justified in including them in the intrinsic aims of ethics education. This is not done here because of the practical problems of forming an educational approach to elicit moral motivation and ego strength and the problems of evaluating the success of such educational efforts.

3.3 Summary and conclusions

In this chapter I have looked at two types of aims for ethics curricula, extrinsic and intrinsic. Because of the basic nature of extrinsic aims as external to the educational process and specific for a certain purpose, the extrinsic aims of an ethics course would have to have a direct application in mind. The most obvious application would be 'right' or 'correct' behaviour or moral convictions in general, or within a limited context. The practical problems with this view relate to the difficulties of defining what are the correct values, how to transfer them to the students, and how to ensure that the adopted values influence behaviour in a desired manner. Further, and more substantially, the problem is the limited perspective on an individual as a means to a certain type of result, as unable to choose his/her own values, and that it is undesirable to do so. Because of both the practical problems and because of the view on an individual as a moral agent, I came to the conclusion that it is not acceptable to choose extrinsic aims for an ethics curriculum.

Intrinsic aims relate to an individual learner and have no direct need of application. From this point of view, an ethics curriculum concentrates on the processes of a person in making moral decisions, rather than on the end results of that process. The value statement inherent in this approach is that it is desirable for a person to act as a moral agent, to choose personal values, and learn to support them.

The view on moral development illuminated here includes two basic components; moral sensitivity and moral competency. Moral sensitivity refers to an ability to understand moral networks, the implications of moral actions, and what moral aspects a particular situation contains. Moral competency, on the other hand, is an ability to form logical arguments that combine the restrictions of the reality with the moral aspects. Apart from the skills of being able to appreciate moral premises and understanding forms of logical thought, one of the key skills is an ability to tolerate ambiguity.

Moral development is a course of learning and growing as a moral agent. The early stages are dominated by clear-cut views and limited personal involvement in the moral decision making. From here a person proceeds to appreciate and tolerate relativism and multiplicity and to an increasing extent become an independent and responsible moral agent. If we accept that this type of moral development is good in itself for the person in question and for the society at large, then we should aim at teaching ethics to support that development, to encourage it to take place, and for people to reach higher levels.

In the following chapters I will look at the process of moral decision-making in more detail and at the possible approaches we can take to support this process. What I hope to have done here is to have provided a substantial case for founding the ethics curriculum on the aim of encouraging moral development.

Part II - Assessing Moral Development

The need to legitimise public expenditure is increasingly pressing within academia. Seeking funds for higher education is confronted with similar difficulties to securing funds for research. One response to these pressures is to justify the expenditure by pointing to useful products that will result. This instrumentalist justification has become accepted as a norm in scientific research and educational research is now expected to follow the trend. External requirements, however, are not the only source of pressure to study outcomes of educational innovations. Administrators within the university also require some methods to evaluate curricular innovations that satisfy internal Quality Assurance (Caplan 1980).

The science curriculum cannot include all areas of science, or even a comprehensive coverage of a specialised area. This requires us to choose what is included and what can be excluded to maintain levels of academic excellence and to meet external pressures from employers and public bodies. Competition for student time within and between departments is intense and proposals for new courses or programmes are likely to meet a less than enthusiastic reception from staff members concerned with obtaining an adequate time for their own areas of expertise. This is further exacerbated by competition for scarce teaching resources. Further, it would not be acceptable to provide students with low quality education. They have rights to expect academic excellence in the area of teaching. For these reasons, new inclusions in the curriculum are often required to satisfy both the criteria of adequate aims and effectiveness before their inclusion.

The question whether ethics deserves greater emphasis in the undergraduate science curriculum has not received unequivocal support from students or staff, while most do support the inclusion of ethics on an abstract level (see Downie and Alexander 1989, Downie 1993). An ability to show that teaching ethics has a significant positive impact might be useful in paving the way for further inclusion of ethics within the curriculum. Also, the diversity of teaching methods for achieving the aims of ethics education almost guarantees passionate debate and disagreement. In this situation the need for

impartial, quantifiable, and objective standards of evaluation becomes ever more urgent.

The evaluation process described in this thesis assesses the changes in students' moral sensitivity and moral competency during one academic year. It is a comparative study of one group of students participating in structured ethics exercises and another group following an otherwise identical academic programme. The aim is to identify whether ethics teaching makes a significant difference in the key ethical skills chosen as the aims of ethics education.

Part II is divided into two main chapters. Chapter 4 - "Assessing moral development - the state of the art" - describes the so far achieved elements in assessing moral development. The concentration is on the testing of Kohlberg's moral developmental stages by the use of Kohlberg's Moral Judgement Interviews (MJI) (Kohlberg, 1976) and Rest's Defining Issues Test (DIT) (Rest, 1986), which are both well-established moral development tools. Assessing moral sensitivity is also discussed by reviewing work by Muriel Bebeau in developing the Dental Ethical Sensitivity Test (DEST) (Bebeau *et al.*, 1985) and work carried out in testing student development on the Perry Scheme (Finster, 1989 and 1991; Gray, 1997).

Chapter, 5 - "Assessment procedure", outlines the assessment method adopted in the research described in this thesis. The first section 5.1 considers the importance and formulation of a control group. Second, section 5.2, is a discussion on the inclusion of existing assessment tools for moral competence (DIT and MJI).

Measuring moral sensitivity required development of novel test items, because neither of the previously used methods were suitable for the purposes of this research. A two-prong approach was adopted: 1) development of open ended responses to a non-structured moral problem, and 2) a tick-a-box test according to the Perry Scheme. The development process is explained in section 5.3.

A third assessment type is the course evaluation given to all students participating in the ethics teaching. The design of the evaluation form is described in section 5.4. The chapter finishes with a short summary.

4. Assessing moral development - the state of the art

The choice of aims drives the assessment of any educational initiative. One of the major concerns is validity - is the assessment tool actually measuring what it is designed to measure. In the case of ethics education, a valid test needs to measure the elements of moral development chosen as the aims of the ethics teaching. More specifically it should measure these elements and no other factors.

The key aspects of moral development that were chosen in Part I as the aims of ethics education - moral sensitivity and moral decision-making skills - are not simple concepts with unambiguous descriptions. We are able to give general descriptions of what moral development is and of different levels of that development. Even quite detailed descriptions can be given for moral problem solving methods and abilities in each level. But translating these descriptions into an assessment method that would score the actual and true moral development reliably and with great validity has proved problematic to many moral development theorists.

The first section (4.1) of this chapter describes the basic assessment methods and provides reasons for not including behavioural assessment as part of the testing of moral development in the context of this research. Section 4.2 describes the Moral Judgement Interviews (MJI) developed by Kohlberg to evaluate the moral development stage. Section 4.3 follows on the same theme by outlining the moral development tool developed by Rest to test the moral development stage according to Kohlberg's theory. Rest's main aim was to make testing moral development quicker, easier and cheaper. The scoring of Rest's Defining Issues Test (DIT) is given a detailed account in sub-section 4.3.3. The neutrality and universality of both Kohlberg's and Rest's test tools are assessed in more detail in sub-sections 4.3.4 - 4.3.5. Section 4.4 describes the two moral sensitivity testing methods devised and found in the literature; Bebeau's DEST and McNeel *et al.*'s (1994) moral sensitivity test. The last section outlines the Perry (1999) scheme and testing work that has been carried out to measure the intellectual development side of the scheme, as there is no literature on ethical development testing according to this scheme. There is a short summary at the end of the chapter.

4.1 Basic assessment methods and behaviour as an option

Most moral development measures are based on either verbal or written accounts of the moral decision making process, either in tick-a-box or open questions in a paper based evaluation or interviews in verbal ones. Interviews can be of two basic types: they can be open ended where the subjects are allowed to tell their story uninterrupted or undirected with only minimal guidance or requests for clarification by the interviewer. Alternatively interviews can be structured where subjects are presented with particular questions for which answers are sought. Interviewing subjects or asking written open ended responses produces large amounts of qualitative data, for which a scoring system can then be applied in order to quantify the data. Qualitative methods are often the most appropriate format at the developmental stages of a new theory for moral growth and progress, while quantified methods are often more applicable for justifying certain theories and practices, or when assessing moral development for a particular purpose - here, in investigating the benefits of ethics teaching.

An alternative to these is behavioural assessment. Even if behavioural change *per se* may not an appropriate pedagogical objective in ethics teaching, it is logically possible that behavioural change may serve as evidence to measure other plausible goals of moral instruction - moral sensitivity and moral decision-making skills (Caplan, 1980).

Without disputing the possible link between moral development and behavioural change, there are serious practical and methodological problems associated with this approach:

1. Validity of observations. Behavioural change needs to be recorded in a laboratory setting as extensive real-life evaluations of moral behaviour are both impractical to conduct and/or they pose problems regarding the subject's right to privacy. A significant problem with this approach is that hypothetical moral dilemmas presented in a laboratory situation can shed only limited light on the future behaviour of the subjects in real life situations. Because we are already making inferences from behaviour to moral decision-making skills, adding an inference from a laboratory setting

to actual moral behaviour complicates the interpretation and considerably reduces the explanatory power of this method.

2. <u>Validity of the method.</u> Evaluating behaviour focuses attention on the outcome and effects of moral decision making, not the process. If the aims of an ethics curriculum are process-based, then we need to assess them, not the changes in outcome which are possible subsequent outcomes. As we cannot establish an unambiguous link between moral reasoning and moral behaviour - research evidence supports a moderate, but consistent and pervasive relationship between moral judgement scores and moral behaviour (typically correlations in the 0.3 range, which implies a low correlation; Blasi, 1994) - it seems methodologically unwarranted to draw conclusions about moral decision-making skills based on purely behavioural observations. A more warranted method would include reflective interviews following the behavioural part of the experiment, where insights on the moral reasoning process could be gained.

3. <u>Practical problems.</u> Behavioural experiments are labour intensive and require highly skilled persons to carry them out. For these reasons, behavioural observations are often an unsuitable method due to limited funds and personnel to carry out the research.

Due to the problems in both assessing behaviour and to the only moderate link between moral behaviour and moral decision-making skills and abilities, behaviour is not considered any further as a possible method to measure the success of an ethics curriculum which aims at eliciting moral development.

I will instead look at two different approaches to evaluating the development of moral competence: 1) Kohlberg's original interview method, and 2) Rest's Defining Issues Test (DIT); and two approaches to evaluating development of moral sensitivity: 1) Bebeau's DEST and McNeel *et al.*'s moral sensitivity test, and 2) Perry's meta-ethical theory. These are here considered as possible methods for evaluating the impact of ethics education in the undergraduate science curriculum.

4.2 Kohlberg

Kohlberg's moral development theory developed from interview research in the 1950's. In these interviews, the subject was introduced to a moral dilemma in which the moral agent has two choices of action, and the research subjects are asked to choose the more appropriate action and to justify their choice. After the initial research, Kohlberg and his colleagues used the same method for assessing the moral development stage of their research subjects. The responses are classified by trained judges (clinical interviewers) according to whether the answer is oriented toward avoidance of punishment and deference to authority (stage 1), toward prudent and purely self-centred concerns (stage 2) and so on (See Appendix II: One Theory, for a more detailed description of Kohlberg's moral developmental stages). Kohlberg's method of data gathering was open-ended. In response to Kohlberg's multifaceted stories, a subject may choose to discuss any number of aspects. The interviewer will not ask any probing questions or interject any form of thinking different from the subject's spontaneous views during the interview (Kohlberg 1994).

The open question interviews pose a problem in setting up a reliable scoring system, and the scoring system of Kohlberg's interviews has undergone several transformations in the past decades. Rest (1983) identified five distinct problems with setting up a scoring system:

1. Establishing a unit of analysis is not straight forward. The options are either using sentences, completed thought patterns, or single utterances as the units of analysis. Using sentences as a scoring unit produced dominantly word-related and repetitive data. Small utterances, on the other hand, produced data that were difficult to analyse as subjects said some things that seemed to reflect one stage of moral development and others that seemed to reflect a different one, while complete thought patterns may omit some less developed ideas or less well presented thoughts completely.

2. In open ended interviews, different subjects bring up different topics and touch on different aspects of the dilemma and these elements may not be comparable.

3. How explicitly must the subjects state an idea to be credited with 'having' it?

4. How to decide what is content and what is structure - how formalistic, deep, or abstract should the analysis be.

5. How to combine conflicting cues from various parts of the interview over several dilemmas into a single summary score that describes the subject in general.

One of the scoring methods is a four tier classification system: the interview material for each dilemma is first separated into issues (2 for each dilemma), then into norms (12 for each issue) and finally into elements (17 possible ones for each norm) (Rest, 1983). An alternative method is a two-dimensional scoring grid of 125 elements (25 elements for each 6 stages) (Rest, 1976). Obviously a system of 125 or 428 scoring possibilities for each dilemma is demanding in its use and the scorer will need extensive training to become confident in the procedure. One of the difficulties with the open-ended data gathering system is the requirements it places on the scorer when the answers are not decisive or complete enough for the scorer to decide clearly how the response should be classified. When the subjects have not given sufficient cues to apply a scoring guide, or when the subject's responses do not seem to fit very well into any of the scoring categories, there is not much a scorer can do but guess. It may also happen that the subjects do not report all the thoughts available to them, but only a selection of them. Also, the test results may vary from one scorer to the next. The reliability data for the four-tier system shows that reliability is not a serious threat to the evaluation procedure, as test-retest correlation has been >0.96, which indicates a very robust testing protocol only slightly dependent on the scorer. The most recent scoring system comprises an 800+ page manual (Rest, 1994).

In the current test procedure based on Kohlberg's initial interview approach, any trace of Stage 5 thinking is extremely rare even among adults and Stage 6 practically non-existent. This makes the post-conventional stage in development almost redundant (Schlaefli et al. 1985). This reduces the power of the tool as the developmental steps become less refined, with only three stages readily available for adult subjects (stage 1 is extremely rare with adult subjects). It will be harder to detect developmental

changes if the tool can only produce very broad descriptions of the moral reasoning stage and when most adult subjects are classified within two or three stages.

Due to its labour intensiveness and high requirements for training of skilled interviewers, and the broadness of the evaluation tool, Kohlberg's moral judgement interviews are not considered further as a viable option for measuring the impact of moral education in the undergraduate science curriculum

4.3 DIT/Rest

James Rest began developing his Defining Issues Test (DIT) as an attempt to make assessing Kohlberg's moral stages easier. DIT is based on ratings and rankings of stage prototypical statements in a pen and paper questionnaire. This method involved writing statements which exemplify various ways of thinking about a moral issue - that is, statements which represent a moral judgement stage - and then asking the subjects to react to the statements instead of asking them to give their own views directly.

The technical issue of devising a measuring tool like DIT is the ability to devise statements that accurately represent the moral stages. It cannot be automatically assumed that a statement designed to exemplify some distinctive characteristic of moral judgement does clearly convey the intended ideas to the reader. Rest (1976), however, provides research results that support the reliability of DIT in correspondence with Kohlberg's stage descriptions.

The chief advantages are, however, that it allows the subjects to focus their attention on specific moral judgements and that it systematically studies subjects' reactions to a standardised set of stimuli statements. Also DIT is far less labour intensive than Kohlberg's original interview technique and thus cheaper and simpler to administer. DIT has been widely used in assessing the moral development of students in different subject areas in higher education. The most traditional area of use has been medical subjects - medicine, dentistry, nursing, and veterinary medicine (Baldwin *et al.* 1991; Frisch 1987; Krawczyk, 1997; Bebeau and Brabeck 1987; Self *et al.* 1989). Lately,

DIT has been applied to a wider range of fields including engineering (Self and Ellison, 1998), humanities/social sciences, and sport (Bredemeier and Shields, 1994). So far, I have been unable to locate any DIT studies targeting science students.

Traditionally DIT contains 6 dilemmas, though a 3 dilemma shortened version is also used. The choice of dilemmas was based on interview responses to original stories and on the possibility of creating short and distinct DIT items based on them. In DIT, subjects are asked to read a description of a moral dilemma, to state what they believe should be done and consider 12 statements and then *rate* on a five-point (most important, much, some, little and no importance) scale how important each of these statements is in making a decision about what ought to be done in the dilemma. After rating the items for each dilemma, subjects are asked to *rank* the four most important items from the set of 12 (Rest, 1983). See Appendix IV - Moral Development Questionnaire for more details and examples.

4.3.1 Assessment of moral development by use of prototypic statements

The basic idea with DIT is that people at different developmental stages construe moral problems differently - particularly what they see as the core element of the moral problem and what considerations they subsequently consider most important in making their decision. Different cognitive structures are called schemas in DIT literature. DIT recognises three structures: 1) Personal Interest schema (derived from Kohlberg's stages 2 and 3 - S23), Maintaining Norms schema (derived from Kohlberg's stage 4 - S4), and Postconventional schema (derived from Kohlberg's stages 5 and 6 - S56). The functions of the schemas is essential to human understanding as they facilitate information processing. DIT is designed to activate schemas and record their use. The prototypic statements in the DIT are fragments of lines of reasoning, each representing a particular schema. The assumption is that people working through the DIT will bring to the task those schemas that structure and guide their moral thinking in general. As the participant reads an item that both makes sense and activates a preferred schema, that item is given a high ranking or chosen as of high importance. Alternatively, when the participant encounters an item that seems

too simplistic or does not make sense, the preferred schema is not activated and the item receives low ranking or rating (Rest *et al.*, 1999a).

Prototypic statement methods can measure different levels of idea acquisition. People can recognise and discriminate and thus prefer an idea before they can paraphrase it or before they can spontaneously produce the idea in a response to a dilemma. Thus, using different methods is likely to locate the same subject at different stages. The developmental hierarchy, in ascending order of difficulty, thus appears to be: preference, comprehension and spontaneous use (Rest 1976). Therefore the structure of the moral development test may influence the results, if the different aspects are not considered when the test is designed. DIT scores preference measures, which elevate DIT scores in comparison with Kohlberg's interview scores (Schlaefli, *et al.*, 1985).

Theoretically, the DIT differs from Kohlberg's test not only in the comprehension aspect, but also in the core concept of defining the stages. Although both methods focus on the concept of justice, Kohlberg defined the stages primarily in formalistic terms of reversability, universalisability, prescriptivity etc., relating purely to the structures of the stages. DIT, on the other hand, characterises justice at each stage as following from different concepts of how social co-operation can be organised. Therefore it has more content-related descriptions. In other words, Kohlberg's stage differentiators are more abstract than the DIT equivalents. Correlations between Kohlberg's test (MJI) and DIT generally range from 0.3 to 0.7, depending on the homogeneity of samples (Schlaefli *et al.* 1985).

4.3.2 Validity of DIT

The three most serious threats to the internal and external validity of DIT according to Rest (1983) are:

1. subjects may check off responses randomly without even reading the items, dilemmas or instructions;
2. subjects may pick out items that seem complex and sophisticated, even when they do not understand their meaning; and

3. subjects may try to fake upwards on a recognition task because they do not have to discuss or justify their answers.

Special features have been built into DIT to counteract these problems. One has been the inclusion of M items - items written to sound impressive and sophisticated, but which do not mean anything (e.g. 'whether the essence of living is more encompassing than the termination of dying, socially and individually'). These are used to identify subjects with a tendency to base their choices on the complexity of the item rather than its content. A raw M-score (sum of M item rankings) of more than 8 results in the participant being suspect of not taking proper interest in the test and in most cases the protocol should be discarded from further analysis (Rest, 1986). The M-score seems also to identify individuals who try to fake their answers to be something that they do not really believe in (Rest 1983).

A second internal check is the consistency check. It is carried out by comparing the subject's ratings and rankings. If a subject ranks an item 1st, then the rating for that item should be highest or equal to the highest of all rated items. Similarly the 2nd ranked item should be the second highest rated item, or equal to the highest ranked item. The protocol is considered to be inconsistent if a participant rates items higher than those they rank first or second. This can be due to either a) careless responding, b) random responding, c) misunderstanding of instructions, or d) changing one's mind about an item etc. Inconsistency raises questions about the reliability of the subject's whole protocol, although slight inconsistency is acceptable and understandable. The general rule is to discard protocols that have inconsistencies in more than two stories, or if the number of inconsistencies on any protocol exceeds 8 instances. Also if there is little discrimination in the ratings (all rated 'some importance', for example) there must be a suspicion that the subject may not have taken the test seriously. The rule of thumb suggested is to discard a protocol if two stories have more than 9 items rated the same (Rest, 1986).

In published DIT research approximately 10-15% of the tested participants fail one reliability check or another (Rest et al., 1997).

Further, the underlying structure of the moral stages is emphasised, so that higher-stage statements appear stark and abstract and do not lend themselves to being interpreted as fancier ways of stating the lower stage ideas - e.g. instead of a statement: 'the value of life is more important than property', the statement is: 'what values are going to be the basis for governing how people act towards each other'. Care has also been taken to match each item on word length, complexity of syntax and use of technical/unusual terms. All items are short fragments which do not highlight the action choice aspect (Rest et al., 1997). Furthermore, subjects are able to choose from several items representing their moral stage and thus are not forced to choose a statement of another stage in case one example of a stage's orientation is not suitable (Rest, 1976).

DIT has been validated through four types of studies:

1. Cross-sectional studies to assess whether DIT detects expected longitudinal changes in moral development scores. Moral judgement has been measured by DIT in different age- and educational groups. The older, presumably more advanced groups, show higher scores than the younger, presumably less advanced groups. Longitudinal studies over 10 years have followed the same subjects and re-testing them gives results of increasingly higher scores. Sequential analysis shows that the upward movement is not attributable to cohort effects or to re-testing benefits (i.e. improved scores due to learning benefits gained by sequential testing) (Schlaefli et al. 1985).

2. Correlational studies to assess whether moral development correlates with general cognitive development as hypothesised. Correlational studies show that upward movement in moral judgement is related to increasing comprehension of higher stage concepts (Schlaefli et al., 1985). DIT is also related to other measures of cognitive development, but it is independent of them (King et al., 1989). Higher stages of cognitive development seem also necessary, but not sufficient, for the attainment of moral development (Rholes et al. 1982).

3. Non-reducibility studies to assess whether moral development measured with DIT can be attributed to a) any other type of development or b) to attitudes and/or political values. a) Multiple regression studies have shown

that when ethics education increases DIT scores, there is no significant improvement in logic scores and *vice versa*: logic education improves logic scores, but not DIT scores (Schlaefli *et al.,* 1985). This indicates the independence of moral development form other variables (here logical development) and that DIT is measuring moral development and no other variable. b) Some results from correlational studies suggest that DIT is not reducible to political attitudes either (Barnett *et al.,* 1995).

This has been contested by Emler *et al.* (1983), who asked subjects first to state their political affiliation (politically conservative, liberal or radical), then complete DIT as themselves, and then again from the perspective of an alternative political view. Political conservatives who obtained relatively low P-scores (the most popular DIT score representing a choice of post-conventional statements: full description in section 4.3.3 below) under normal test conditions, elevated their scores when instructed to respond from the radical perspective. Political liberals, on the other hand, obtained relatively high initial P-scores, and their scores decreased when instructed to respond from a conservative perspective. The latter result is consistent with the view that a subject can fake downwards, because these are forms of thinking that the subject once used and discarded because they are now seen as simplistic. On the other hand, a subject should not be able to fake upwards because the concepts that the subject is using represent the subject's best notion of moral ideals, thus the 'faking upwards' results are more problematic. Emler *et al.* concluded from these results that moral judgement, as measured by DIT, and political attitudes, substantially overlap and therefore that DIT is not a good developmental measure: for Emler *et al.* DIT scores are best explained by variance of political affiliations.

However, several other tests with DIT have shown the contrary evidence that people are not able to fake their results upwards, even when induced to do so (Hau, 1990; McGeorge, 1975; Page and Bode, 1979). A weakness in Emler *et al.*'s study is that they did not demonstrate that politically conservative subjects actually understood those DIT elements that accounted for their higher P-scores when they responded as radicals.

Also, regression analysis has shown that DIT score is related to liberal/conservative views, but that DIT is not empirically reducible to political attitudes alone (Barnett *et al.* 1995). Further, when Emler *et al.* asked conservative subjects to fake their DIT scores as radicals the strongest effect was an increase in A-scores (anti-establishment scores) and only secondarily in P-scores. When asked to fake DIT as radicals, the subjects rejected Stage 4 items because they were the ones they endorsed themselves in the normal test situation: choosing lower stage items was not attractive, because they were viewed as simplistic, and therefore the shift in P-scores could have been due to elimination of lower level items for conservatives faking the scores as radicals. This hypothesis is supported by Barnett *et al.* (1995).

4.3.3 Scoring and analysing DIT results

A variety of scores can be derived from the ratings and rankings of items in DIT. One of the older and more basic ones is the Moral Maturity Quotient (MMQ). It is a weighted average calculated by taking the percentage of stage 1 usage and multiplying it by 1, stage 2 percentage multiplied by 2 etc, and then adding the six products together. A subject that is 100% stage 2 would have MMQ of 200, a subject 90% stage 2 and 10% stage 3 would have MMQ of 210 etc. One of the problems with MMQ is that it can produce results that do not allow clear stage allocation for research subjects. For example, a MMQ 225 could indicate advanced stage two, but looking at the details of the result it can turn out to be a result of 10% at stage 4, 25% at stage 3, 40% at stage 2, and 15% at stage 1, which is not best described as stage 2. A more pressing problem with MMQ is that it may fail to recognise true moral development. For example, a MMQ of 375 can be a result of 70% stage 4 usage, 15% stage 3 usage and 5% stage 2 usage. A similar quotient of 415 can in turn be a result of 20% usage of stage 5, 60% usage of stage 4 and 10% usage of stage 3. The MMQ difference is a relatively moderate 40, but in this example it signifies a major shift in inclusion of post-conventional thinking. A person with a MMQ of 415 would be classified at stage 4, while the stage 4 usage is already in decline and stage 5 usage has emerged. MMQ

would be a very useful scoring method only if the development through stages could be assumed to be a symmetrical step-by-step process (Rest, 1976).

Indexing is a more developed way of analysing DIT data. An index is an overall score by which a participant is characterised (Rest *et al.*, 1997). For the past 20 years, the most widely used index has been the P-index - the relative importance that a subject gives to stage 5 and 6 items, where the P stands for 'Principled moral thinking'. Operationally the P-index is a percentage of top rankings given to stage 5 and 6 items. The P-score is calculated by considering only the items written to represent Stages 5 and 6. In each story, when a P-item is chosen as the most important consideration it is assigned 4 points; if a P-item is chosen as the second most important consideration it is assigned 3 points; third most important, 2 points; and fourth most important, 1 point. All points are then added together with a maximum sum of 9 per story as there are only three stage 5 and 6 items in the choice of 12 statements. The sum is converted to a percentage by dividing the total P score by 0.6 (if the test included 6 stories, by 0.3 if the test included only 3 etc.) - resulting in a score between 0 and 95% (Rest, 1976). Missing data leads to adjusting the P-score on the responses given: for example if a participant leaves out the third rank on one story, the P-score is re-calculated on the basis of 58 points instead of full 60 points, i.e. total P-score converted to percentage by division by 0.58 instead of 0.6 (Rest *et al.*, 1997). The P-index produces the consistently strong trends expected of moral development (Schlaefli *et al.*,1985).

The difficulty with the P-score is that it only recognises statements representing the post-conventional stages. It can then happen that a subject choosing coherently conventional items representing level 4 is given a P-score of 0. This person might have chosen statements representing moral reasoning of stages 2 and 3 a year earlier, and this development would not be captured by calculating a P-score. The use of the P-score is best justified when testing young adults, who can be expected to view moral problems at least partially in post-conventional terms, or are expected to approach that stage in their development in the near future.

The limitations of P-index have led to research to design an alternative index. One of the big shifts has been to replace a stage-focused by a type-based approach. Types are

based on an understanding that people use a variety of stages in their thinking dependent on the situational requirements. The analysis of DIT data should therefore likewise concentrate on identifying under what conditions does a person manifest particular stages of thinking. In this approach, development means that people over time come to use higher stages more and lower stages less. Development is thus a gradual shifting of distributions of stages rather than dramatic moves from one stage to another (Rest *et al.*, 1997).

Rest *et al.* (1997) have developed N2, an index which hopes both to reduce the waste of data built into the P-index and to offer a way to study stage distributions. An N2 score has two parts: the degree to which P items are prioritised (closely related to P-index) and the degree to which the lower stages are rated lower and the higher stages higher. Thus N2 uses both ranking and rating data. N2 calculation consists of three parts:

1. P-items are calculated like P-scores, but no adjustments are made for omitted rankings, though leaving out rankings for a whole dilemma will be adjusted by basing the total score on n-1 dilemmas.

2. The second part of N2 is calculated from the rating data. Discrimination between lower and higher items is calculated by subtracting the average rating given to items representing stages 2 and 3 from the average ratings given to items representing stages 5 and 6. Average ratings are then standardised by dividing this difference by the participant's standard deviation of stages 2+3+5+6. Occasional missing rates are supplied by filling in the average ratings for the story. If rates are missing for an entire dilemma the scores base is adjusted to n-1.

3. N2 is then calculated as a combination score per participant by adding the P-score to the rating data weighted by three. N2 produces similar longitudinal and cross-sectional data to the P-index and it is found to be more sensitive in most instances, for example to measure the impact of educational interventions (Rest *et al.*, 1997).

A second new approach to analysing DIT has been a Type-approach. It is possible for two participants to have the same N2 score with very different mixes of schemas. That

begs the questions: how to describe participants' predominant schema *and* their schema mix so that they characterise 1) developmentally ordered types, 2) that we can see advances of different types in relation to information processing, and 3) relations to different decision-making. A type-model of 6 different types has been constructed to fulfil these criteria (Rest *et al.*, 1999a).

Types are defined by 1) which schema is predominant, and 2) in terms of the extent of schema mix - consolidated in one schema or a transitional mixture of all the schema in which no one schema predominates (CDIT).

1. The predominant schema is determined by calculating average ratings for each 3 clusters of schema items (pre-keyed in the DIT) and the highest average is designated as the predominant one.

2. The extent of the schema mix is determined in terms of the ratio of variance *within* schemas and *between* schemas (Rest *et al.*,1999a).

A detailed guide on how to calculate the schema mix can be found in Appendix V - Computing C -scores from DIT Data (Rest *et al.*, 1997).

Participants are described as consolidated if their CDIT scores are greater than 15.705 and transitional if CDIT is less than that. This is a somewhat arbitrary classification criterion chosen by Rest *et al.* (1999a). Theoretically CDIT can vary between 0 and 100, but in their sample of 505 participants Rest *et al.* (1999a) found a range of .004 to 47.066 with a mean of 11.89 and standard deviation of 9.09.

When these two elements are combined we get six different Types (Rest *et al.*, 1999a):

Type 1	Consolidated	S23
Type 2	Transitional	S23
Type 3	Transitional	S4
Type 4	Consolidated	S4
Type 5	Transitional	S56
Type 6	Consolidated	S56

Figure 1: DIT Types

In their samples of 505 and 44,564, Rest *et al.* (1999a) found 95% of the subjects to fall into types 3 and 5, and only very few subjects to Types 1 and 2. The low number of types 1 and 2 is explained by the sample consisting of subjects in their teens or older, who are expected to have progressed beyond Types 1 and 2. If DIT is administered to children, you would expect a higher percentage of lower Types. The mean schema ratings in types change from a dominant S23 in type 1 use to dominant S56 use in type 6 via dominant S4. What Rest *et al.* (1999a) have also found is that types are developmentally ordered in terms of educational level and age. The correlations between types and age/educational level are weaker than those with N2, but still significant.

The types have also been found to correspond with the ease of information processing, with consolidated types having less trouble than the transitional ones. There is also evidence that the schemas guide the actual decision-making as well. People who use dominantly the same schema also choose similar moral actions (Rest *et al.*, 1999a).

A further DIT score is the U-score, which represents the extent by which subjects utilise the justice-framework in their answering patterns. This was created as a response to the low correspondence of P- and N2 scores to moral action choices (correlation in the low .30s) (Thoma *et al.* 1991). For example, in Heinz's dilemma, if a subject rates an item 'Whether a community's laws are going to be upheld' as high, we would expect this person who focuses on maintaining the law to advocate 'not steal' as an action choice. In contrast, if a person rates high an item 'Isn't it only natural for a loving husband to care so much for his wife that he'd steal?', we would logically expect that person to choose 'should steal' as an action choice (Rest *et al.* 1999a). U-score measures the correspondence between these actual and implied choices and is high when a subject selects as important those DIT items which support their initial action choice, and rate as least important those items which imply alternative choice. U-scores have a potential range of -1 (low utilisation) to +1 (high utilisation). However, large sample estimates of utilisation suggest an actual range of -.40 to .77 (Thoma and Rest 1999).

A lack of consistency between ratings and action choice (low U score) implies the use of multiple moral frameworks, one to determine the decision, and the second to evaluate the stage-based DIT items. U-scores are found to relate to consolidation/transition of subjects. The more coherent the schema (lower C-score) the higher the U-score. Theoretically, people with a predominant, coherent schema are presumed to show greater integrity between item ratings and their choice of action. Presumably, weaker schemas (higher C-scores) cause subjects to change schemas between action choice and rating of items (Rest et al. 1999a).

All DIT scores represent the basic interpretative framework that people naturally and spontaneously bring to moral problem solving - the default schema. A low P-, or N2-score, or lower type means that the person is predominantly not able to conceptualise moral problems by determining what is morally right from the perspective of a society that balances the interests of its participants, optimises the stake of each participant in supporting that society, and eliminates arbitrary advantages and/or influences. Students with low scores are also unlikely to appreciate the insights of ethics professionals and discussions: rather they seem superfluous solutions for problems neither foreseen nor recognised. These students have likewise trouble in extending the principles beyond a discussed case, and they are genuinely confused when ideals conflict. In real life situations they are likely to oversimplify, and although they might have good technical skills and generally good intentions, they are vulnerable to finding themselves involved in ethical problems over their heads (Rest and Narváez, 1994).

Since the start of this research an alternative DIT (DIT2) has become available. It updates dilemmas and items, shortens the test time, and purges fewer participants for doubtful response reliability (Rest et al., 1999b). Unfortunately the decision to purchase of DIT1 was made prior to the availability of DIT2 and thus it was impossible to benefit from the improvements made to the test tool.

4.3.4 Gender variance

Strong claims have been made that moral development methods, in particular Kohlberg's interview method and DIT, are gender-biased. Carol Gilligan has been one of the most dominant figures in advocating that women generally score lower on the Kohlberg moral dilemmas and that this is due to these dilemmas measuring only one type of morality - the rights and justice approach - which is not the dominant moral framework for women (Gilligan, 1982). Some research supports this claim. Bakken and Ellsworth (1990) claim that women show similar age-related moral development to men, but that at each age group, women score consistently lower in the Moral Judgement Interview. Further, they found that women's moral development levels off at stage 3, which is consistent with Gilligan's theory that women prefer to consider contextual and personal effects in moral problems compared with male orientation towards more abstract and generalised rights-based effects. Gilligan's argument is not an acceptance of lesser ability of women's moral reasoning, but a proposition that moral reasoning at a contextual level does not essentially refer to lower moral abilities, but to a different approach. Gilligan's argument is discussed in more detail in Appendix II: One Theory.

The results of Bakken and Ellsworth have not been supported by other research. Thoma's (1984) meta-analysis of 56 studies consisting of over 6000 male and female subjects, yielded no significant gender differences in DIT scores. Education, for example, was found to be 500 times more powerful in predicting moral judgement scores than gender. This could be interpreted as disputing Gilligan's argument of special female moral schema. But it is not necessary to read Thoma's results as evidence against Gilligan's theory. Alternatively, they can be interpreted as evidence simply that women are not disadvantaged in their development towards principled moral judgement, while they may also entertain alternative moral schema, possibly similar to the one Gilligan has put forward.

Some researchers go even further and suggest that women actually score higher in DIT. Baldwin *et al.* (1991) tested medical students on four levels of their medical course, and found that women scored higher in DIT at all levels, and their

development continued strongly throughout the course, while the development of their male counterparts plateaued after the third year. A possible explanation, provided by Baldwin *et al.*, is that women in medicine and in other high achiever professions do not follow the traditional gender patterns, and therefore display male patterns of moral development, and because they have had to solve the conflicting obligations of their gender role identity, they are more prepared to handle the contradictions inherent at higher levels of moral reasoning.

The results of Baldwin *et al.* (1991) do not therefore necessarily suggest female-bias in DIT, but that female medical students have, on average, a higher capacity and interest in moral development. This is not explained in deterministic terms as a biological feature, but as a result of social pressures experienced by professional women in a still largely male-dominant academic culture.

These results give some confidence in the gender-neutrality of DIT. Any gender-biased results, in the light of Thoma's analysis, should not be interpreted as in-built structural biases in the test tool. Gender-biased results should be interpreted as genuine reflections of the research design. In the case of educational research, reasons for gender bias in the post-test results could, for example, be attributed to teaching methods which favoured interactions preferred by female/male students and thus accelerated their moral development more than their male/female counterparts.

4.3.5 Cultural variance

It seems almost axiomatic that different people have different moral values. From this it is tempting to infer that moral development as described above may be culturally dependent and claims of its universality flawed. But both Kohlberg and Rest have argued that certain considerations are so fundamental to human interaction in groups that they are relevant regardless of one's particular culture. Examples of fundamental considerations include: issues of life versus law, conscience versus punishment, and contract versus authority. The specific morals of cultures may be ever-changing, but beneath the surface, structural conceptions remain unchanged (Rest, 1994).

Snarey (1985) has argued that for moral development, as Kohlberg has described it and Rest further incorporated into DIT, to be truly culturally independent and universal, satisfactory proof of four contentions is required:

1. That persons in all cultures inquire about the moral domain and in doing so, ask the same basic questions, or resort to the same basic issues.

2. That in all cultures individuals are found to be upwardly invariant in sequence of moral development without significant regression.

3. That the full range of stages, including the highest, can be found in all types of cultures.

4. That all instances of moral reasoning in all cultures correspond to one of the modes or stages - in other words all moral responses can be scored according to the standardised scoring guides.

In order to prove that all these contentions are correct, moral development has to be studied in a sufficiently wide range of cultures - including different religions, and types of social structure - urban and agricultural, stable and nomadic. Also the research needs to adjust to the culture in order to make the moral problems relevant to the research subject - or at the very least translate the moral dilemmas into the native language of the subjects. Snarey (1985) quotes research using Kohlberg's interview technique in 27 different countries and Rest (1994) the use of DIT in over forty countries. Both sets of research include a wide variety of social, religious and economic structures.

In his review Snarey (1985) found that research supports the first three of the contentions presented above - all cultures ask similar moral questions; moral development progresses in a similar invariant manner; and all stages can be found in all cultures. What seems to emerge from this research is also that the further the culture is from a middle-class Western one, the less the researchers are able to detect higher stages of moral reasoning. Snarey attributes this to a lack of support for the last assumption that all types of principled moral reasoning could be coded using Kohlberg's developmental model. The more sophisticated levels of moral reasoning of, for example, Hinduism or Buddhism, may not be included in the examples of the scoring guide and therefore such responses are not coded and the scores remain low.

DIT research supports the basic assumptions stated by Snarey, and finds less problems with identifying higher stages. This can be explained by DIT using prototypic statements, which give it additional cultural neutrality as there is no room for misinterpreting the higher stage reasoning of other cultures which prevails when moral development is determined through interviews.

The issue of cultural universality is only partially relevant to the current study. The University of Glasgow is a multi-cultural university with students coming from varied social, religious and linguistic backgrounds, but the majority have grown-up in and are from a Western cultural setting and all have a good command of English. Against this background and the research results on the cultural neutrality of DIT, the moral assessment can be considered not to carry any cultural bias.

4.4 Moral sensitivity testing

Moral sensitivity refers to an ability to recognise moral problems in a situation where they are present, but not necessarily apparent. It is also an ability to consider the moral implications of actions not just to oneself or those immediately involved, but in a wider perspective, recognising the impact of moral choices on unidentified individuals.

Research has shown that scores in moral competence (DIT) and moral sensitivity correlate only modestly (in the 0.2-0.5 range) (Bebeau and Brabeck, 1987). It is possible for a person to be skilled at interpreting the ethical issues in a situation, but unskilled at working out a balanced view of a moral solution, and *vice versa*, to be unable to recognise the issues personally while being skilled in solving these problems. Therefore if moral development is understood as development in both moral sensitivity and moral competence, it is necessary to have separate test tools for both aspects, as development in one area cannot be taken to indicate development in the other.

Previous research on moral sensitivity is limited. Attention on moral development has been focused on moral competence measures. Literature provides two, quite similar,

approaches to testing moral sensitivity: the Dental Ethical Sensitivity Test (DEST) and moral sensitivity testing by McNeel.

DEST was created by Bebeau *et al.* (1985) to measure dental students' ability to identify and interpret typical ethical problems arising in dental practice. DEST comprises four recorded dramatised dialogues that might occur in a dental office. The subjects are first asked to listen to the dialogues and later enter the dialogue and assume the role of the dentist and carry on as if he or she were actually in that position. The responses are recorded and later the students are interviewed about their assumptions and perspectives underlying their responses. These interviews are likewise taped. The interviews are transcribed and scored to measure the degree of sensitivity to the responsibilities of the dentist. Seven sensitivity criteria are described for each dilemma and students are scored on a scale from one to three indicating their degree of recognition. The scoring criteria were developed in collaboration with practising dentists and moral philosophers. DEST has proved to be reliable with inter-rater agreement averaging 0.87 and test-retest correlation averaging 0.68 (Bebeau and Brabeck, 1987). The correlation between DEST scores and DIT was found to be 0.2 to 0.5.

DEST is very specific for measuring moral sensitivity in a professional context. The research literature does not entertain considerations of whether professional moral sensitivity can be understood as general moral sensitivity or whether moral sensitivity can develop in relative isolation in different areas of life, and thus one should not extrapolate these results to measures of general moral sensitivity. Further, this approach is only applicable to professions where moral considerations are situated in personal interactions, as in medicine, teaching and law, and to a certain extent in science (fraud, whistle blowing): but this is a less suitable approach for measuring moral sensitivity in a situation where personal interaction is limited (ethics of genetic research, for example).

A similar moral sensitivity test was designed by McNeel (1994). In this research, college students were played four recorded drama situations containing moral problems frequently confronted by students - a) cheating, learning problems, and

racism; b) pressure for sex, date rape, depression, and co-dependency; c) grieving for parent's death, autonomy, career decisions, and parental pressure; and d) alcohol abuse and its consequences, irresponsibility, and broken trust. Before hearing the drama, students were informed that the researchers were interested in what the students noticed and what they paid attention to. After hearing the drama, students took the role of the central character's best friend and spoke into a tape recorder as though they were speaking directly to their friend. Non-directive follow-up probes were used to help the students to express themselves on all the relevant issues they had noticed in the situation. Coding manuals were devised to allow reliable and valid scoring of transcriptions. McNeel found gender differences in the results, but only in some issues. He also found that perception of some moral problems was significantly low - in particular in the date-rape and pressure for sex drama. No comparison between DIT scores and moral sensitivity was presented for the data.

The test approach of McNeel is less tied to professional moral sensitivity, while the approach is similar to Bebeau *et al.*'s DEST in providing scenarios for individual involvement and direct contact with the problem. The results of McNeel also indicate that moral sensitivity is case-dependent, which supports the possibility that moral sensitivity in professional issues may not indicate moral sensitivity in other areas of life.

4.5 Perry scheme

Perry's theory of ethical development (Perry, 1999), developed in the 1950's and 60's, concentrates strongly on the basic conceptions of morality; what are the sources of moral answers, whether there are absolute moral answers, and how one is to commit oneself to different moral values. In the Perry scheme moral development follows the same pattern as intellectual development, though their progress may not be synchronised. Perry's conception of moral development is a pre-requisite for higher levels of moral competence and even moral sensitivity as students who do not accept the possibility of multiplicity and relativity in moral answers cannot successfully consider moral options in their complexity and extent.

Perry's scheme has been dominantly used to measure intellectual development and attitudes towards learning and studying (Gray, 1997; Katung *et al.*, 1999). The approach to measure intellectual development on the Perry scheme is applicable for measuring ethical development as well. Therefore, I will here describe the approach to intellectual development measures, in the absence of any research data on measuring ethical development as Perry has described it.

According to Selepeng (1999), the measures of intellectual and attitudinal development towards studying concentrate on four elements in the Perry model:

A.   Students' perception of their role as learners
   1.   **A type student**: Passive receptor
   2.   **B type student**: Recognises some responsibility for learning, but is confused on what and how
   3.   **C type student**: Sees him/herself as a source of learning
B.   Students' perception of teachers' role
   1.   **A type student**: teacher is authority and source of facts and know-how
   2.   **B type student**: teacher is an authority, and in disputes he/she seeks to find out what the teacher favours
   3.   **C type student**: teacher is authority among authorities, source of expertise
C.   Students' view of knowledge
   1.   **A type student**: knowledge is black and white, factual, non-controversial
   2.   **B type student**: knowledge no longer black and white, which causes insecurity
   3.   **C type student**: knowledge is contextual
D.   Students' perception of exams
   1.   **A type student**: exams require regurgitation of facts, hard work is rewarded
   2.   **B type student**: quantity > quality, important to display maximum knowledge
   3.   **C type student**: quality> quantity, wants room for expression

The assessment of students' development on the Perry scheme has been mainly done with pen and paper measures where students have been asked to agree or disagree (on a five- or six -point Likert scale) with statements that are designed to represent one of the three Perry types. The questionnaires have been validated by experts indicating which statements would different Perry types agree and disagree with; statements with

high inter-rater correlation are included in the final assessment method (Harvey, 1994).

Gray (1997) describes two alternative methods for evaluating Likert-scale Perry questionnaires: the Success index (SI) and the Perry index (PI).

$$SI = \frac{T_C + T_B - T_A}{T_C + T_B + T_A}$$

$$PI = \frac{T_C - T_A}{|T_C| + |T_A|}$$

T_C = total of agreements with C statements
T_B = total of agreements with B statements
T_A = total of agreements with A statements

These two methods have high correlation, which indicates that whatever a Perry questionnaire is measuring, it is a relatively robust tool and several alternative methods of quantifying the answers can be used.

Alternatively, Perry development can be measured by use of a grid-based test, where students are asked to choose one of three statement that best describes them in each of the four elements in attitude change described above (Katung *et al.*, 1999). The answers are analysed by categorising students in six categories - at least three As = dualism; mix of As and Bs = early multiplism; minimum of three Bs = multiplism; minimum of two Cs = early relativism; and minimum of three Cs = relativism - according to how many times they agreed with A, B, or C statements on the grid. Changes in distribution can then be analysed further to discover whether there have been changes in Perry positions.

A third measurement uses an Osgood scale. In this method Perry A and C statements are placed on opposite ends of a scale and students are asked to either agree with one of them or to be neutral in the middle. For example:

I believe it is the job of the lecturer to supply me with all the knowledge I need	strongly agree - agree - neutral - agree - strongly agree	The duty of the lecturer is not to teach me everything, but to stimulate my own thinking

(Selepeng 1999)

One of the advantages of the Osgood scale is that students are encouraged to consider carefully their position along the scale by making them aware of both opposite views. Also in the Likert scale it has been difficult to detect B-types, because C type builds upon the growth already achieved in B: therefore C type students may wish to agree with B statements as well as C statements. In the Osgood scale, Bs will sit close to the middle, choosing dominantly neutrality over strong agreements with either A or C statements and will thus form a unique B-pattern. A further advantage of the Osgood scale is improved detection of As and Cs, as disagreement with B statements in the Likert scale is open to two interpretations - either A or C, while in the Osgood scale, such confusion does not arise.

Research using Perry's scheme has produced strong evidence of developmental change during university years. Typically, large number of first year students are type A, while after the first year the percentage of type A students has fallen significantly and the percentage of students in the C positions has taken a respective rise (Katung *et al.*, 1999). Movement along the Perry scale is dominantly positive: students progress towards the C position, but there are also notable numbers of students who do not progress and a small number of students that display regressive movement towards position A.

A Perry based questionnaire for moral development should measure two distinctive elements in moral development - recognition of moral multiplicity and an ability to cope with moral relativity. I will discuss the methods of developing an ethical Perry questionnaire in section 5.3.1.

4.6 Summary

Kohlberg's theory of development in moral competence was originally combined with Moral Judgement Interviews (MJIs) as the test tool. MJIs are lengthy, require skilled staff and have very complicated scoring guides. For these reasons MJIs are not included in moral development assessment in the research described in this thesis.

Rest's Defining Issues Test (DIT) measures the same developmental theory as MJI, but is a pen and paper test, which can be computer scored. In DIT, which consists of responses to moral dilemmas, the most common measures are the P- and N2-index, and Types, which indicate the usage of different moral decision-making schemas.

Research shows that DIT is a very stable and reliable test tool with large numbers of studies indicating high test-retest reliability and applicability to people in different cultures and of both genders. The drawback of DIT is its narrow scope of measuring only justice and rights oriented moral judgement skills, where universalised judgement is *per se* considered as the highest form of moral decision-making.

The moral sensitivity testing described in this chapter is based on student responses to recorded interviews (DEST) (Bebeau *et al.*, 1985) or recorded stories containing a moral element (McNeel, 1994). The student responses are scored according to the level of recognition of the moral issues.

In the Perry scheme, plenty of research has been carried out to measure intellectual development and attitudes towards studying, while no publications could be found on the use of the Perry scheme in moral development, even though that was the other element Perry originally designed his scheme to represent. The review here therefore concentrated on methods of testing intellectual development, with the assumption that the methods are equally applicable to measuring ethical development. Most methods rely on pen- and paper questionnaires using either Likert or Osgood scales and relying on student responses to Perry-typical statements to represent their own Perry position.

5. Assessment procedure

The choice of assessment method for the IBLS ethics teaching was strongly influenced by time and resource restrictions. With a sample size of over 400 students all to be tested twice, it was impractical to plan an assessment method producing large quantities of qualitative data, or to adopt assessment methods which are laborious and/or expensive to carry out. The time and resource requirements for analysing interview transcriptions, for example, were beyond the budget of this research. Such methods may produce more intricate data, but because of the volume, the assessment method for this research had to rely on pen and paper methods that can easily be quantified.

This chapter will start with considerations over a control group - why it is important and what alternatives there are. This is followed by, in section 5.2, deliberations on including an existing moral competence measure into the moral development assessment tool used to measure the impact of ethics teaching in the research described in this thesis. The following section 5.3 is on the development of novel testing methods to measure moral sensitivity. A two-prong approach is adopted and the first sub-section 5.3.1 describes the development procedures and preliminary results from a tick-a-box assessment questionnaire based on the Perry scheme of ethical development. The second sub-section 5.3.2 describes the development and scoring methods for the open-ended questions targeting moral sensitivity. The last section of the chapter outlines additional testing methods, using a course evaluation to collect more detailed data about students' own perceptions on the benefits of ethics teaching.

5.1 Issue of control group

This is comparative research. The aim is to investigate whether teaching ethics as part of the biological sciences curriculum will support students' moral development significantly better than teaching science without any specifically designed ethics components.

Previous research has shown that formal education is one of the strongest and most consistent correlates with moral development (Rest and Deemer, 1986). Also, there is strong longitudinal data that shows moral development throughout the years in higher education (McNeel, 1994, Rest *et al.*, 1999a). Therefore we can expect moral development of students as they pass through the levels of university education. In wanting to measure the effects of a series of bioethics exercises on the moral development of students, it is essential to have a control group to be able to differentiate between spontaneous moral development and the effect of ethics teaching on moral development. An alternative would be longitudinal studies of students in different (consecutive) years, where one year goes through an ethics curriculum and one does not. Unfortunately the limited time-scale of this research did not allow for such a research set-up.

The control group should be compatible in biographical features (most importantly age and gender) and in academic qualifications and aspirations with the test group. All these features are hypothesised to introduce variation in moral development scores. Most importantly the control groups should have, during the period of interest, very similar academic experiences apart from the ethics teaching. Great variation in the general academic experience would result in considerable interpretation difficulties in extracting the source of change in the moral development scores.

5.1.1 Options for control groups

The three options for a control group to be used to identify the benefits of ethics teaching on IBLS students were: other IBLS students; other science students from the University of Glasgow; or from another university. I will consider these options in turn.

IBLS students. The most obvious and suitable control group is created by dividing the IBLS student body into two groups. One of these groups would follow the science curriculum without the ethics components and the other group the same teaching with the only difference being participation in the bioethics exercises. This method would

secure comparable academic experiences between the test and control groups. Also, allocation into the two groups within a large student body (700-800 on each level) should secure satisfactory compatibility of biographical features between control and test groups.

A more pressing practical problem is to keep up two separate tutorial and laboratory programmes throughout the year, and to provide novel alternative teaching for the non-ethics group. This requires extra labour on the part of teachers and demonstrators, time and resources which may not be easily available. This problem can be largely solved by the extra teaching on bioethics being provided by the primary researcher.

The theoretical/ethical problem of dividing IBLS students into those participating in ethics teaching and into those that do not, is the disparity of their learning experience, which may not be acceptable. The content of teaching can affect students in both groups either positively or negatively, depending on how successful the ethics components are. Either the students in the 'ethics group' will be in an advantageous situation because they receive new and beneficial learning opportunities or they may be disadvantaged for spending their time and dedicating their effort onto something that is not beneficial and thus having less time to devote to material that is being assessed. It will be known only at the end of this research, which of these scenarios is correct. Even though we cannot know the effects of the ethics teaching, we can anticipate at the start of the research that the students participating in ethics teaching are likely to benefit from it. From this expectation a moral question arises: namely, is it right to withhold teaching that is expected to be beneficial to the students in the name of research? The problem is not dissimilar to medical research situations, when treatment estimated to be highly beneficial is not provided to all patients and some are given placebos instead of the drug, that might, if not save their lives, provide them with alleviation of pain or discomfort. But in medicine, as in this research, the effects of the treatment may not be advantageous, but even hazardous, and those belonging to the placebo/control groups may turn out to be the ones 'benefiting' from the research. The difference between the medical example and the ethics education example is that the patients in the medical trials give informed consent to the researchers, and by that consent accept the possibility of being allocated to either group. Informed consent

removes most of the problems relating to autonomous decision-making and a possible right to best possible treatment. But in the ethics teaching case, which is of interest here, acquiring consent is somewhat complicated, but not impossible.

Allowing students to choose which student group they wish to join is likely to distort the results. Research by Spickelmier (Rest and Deemer, 1986) shows that moral development scores correlate positively with students' approach to their work. More precisely, higher moral development scores (using DIT) correlated positively with high motivation towards studies, involvement in university life, clear academic goals, enjoyment of studies, keen approach to new ideas, and participation in the activities of reading and discussing. Therefore we can assume that students with high motivation, and hence those who are more likely to have already accelerated moral development prospects, would be the ones opting for the ethics curriculum, which might be seen as novel, more participatory and challenging. Those who are less motivated and with a lower starting point on moral development would most likely choose the traditional curriculum. Because of this potential bias, the results could be distorted and interpretation difficult.

On the other hand, students could be divided into the groups randomly and informed of the reasons for this division. This could by-pass the ethical problem as long as students strongly objecting to their allocation were allowed to change groups. This presumes that most students would be satisfied with their original allocation. The problem with this is that the students' approach to the ethics components within the curriculum may change as they are told about the untested value of that form of teaching, which in turn may influence the effects the teaching will have on the students. By the same token, it should be possible to maintain student autonomy and reduce any adverse emotional reactions by giving the students information about the research, and outlining its benefits, without jeopardising the clarity of the results.

A way to avoid most of these problems is to form test/control groups within each degree course and teach the test groups in term 1 and the control groups in term 2 with the post test administered in between. This way neither student group is disadvantaged in having/missing a teaching element. It would still be necessary to make sure that the

students are informed of the research nature of the teaching and that they are given an opportunity not to participate in the questionnaire part of the research, if they object to doing so.

Other science students. An alternative to the use of IBLS students as both the test and control group is to use an external control. The external control groups could be of two types, either biology students from another university or other science students from the University of Glasgow. If the external control group was chosen from another university, care would have to be taken that the students in both groups were compatible in all important biographical and academic aspects. Also the university experience would have to be sufficiently similar to enable comparison of the results. Controlling both the biographical and academic variables would not pose an impossible task, while standardising the academic experience in order to enable comparison of the data is extremely complicated. These difficulties are likely to undermine the usefulness of a group from another university as a control group i.e. they would not be a proper control group, but in the relevant aspects a different population. Therefore the use of an external 'control group' formed of students from another university threatens the results of the entire research. Because of this threat, the use of an external control group from another university is not considered further.

Considering other science students from the University of Glasgow, the first issue is whether there is some element in the students choosing different disciplines that influences their moral development, which would not be apparent when judging compatibility of academic and biographical variables. If both groups scored similarly in the pre-test (at the start of the academic year), the worry would be removed. Alternatively if the pre-test results were significantly different between the control and test group, analysing the results of the post-test (at the end of the academic year) would be open to a large variety of interpretations, which is not desirable. Therefore information about the compatibility of the biology students and the control group of other science students in reference to their moral sensitivity and moral reasoning skills is required before the start of the academic year of research interest in order to decide on the acceptability of the control group.

The second issue is incompatible university experience. Different disciplines present students with different types of problems, stimulate their thinking differently and provide them with different extra-curricular activities. All of this may influence their moral development, either positively or negatively. Therefore, even if the pre-test results are similar, in the case of only small differences between the control and test group in the post-test, it would be plausible to locate the source of difference in the incomparable academic experience. This possibility could undermine the entire research set-up. There would be a risk of this research becoming academically insignificant, producing ambiguous results which would not stand up to stringent peer review.

5.1.2 Decision on the control group

To be able to produce results on the effects of an ethics curriculum on the moral development of biomedical and life sciences students, it is vital to have a control group that matches the test group both in biographical aspects and in academic experience. This compatibility can be best secured by forming both the control and test groups out of the same student body so that both groups participate in an identical curriculum structure apart from the ethics components. This compatibility secures unambiguous interpretation of the test results and provides a true opportunity to assess the effects of ethics teaching.

Forming the control group from students from other faculties or departments would increase the difficulty of drawing conclusions from the results and produce a risk of the research becoming insignificant. Therefore such an approach was not undertaken.

The issue of unfairness resulting from differing academic experience was minimised by three elements: 1) students were given an opportunity to choose their group if they were unhappy with the original random allocation into ethics or non-ethics group. 2) Most groups were formed along naturally existing group lines within different degrees. For example, within the degree of Parasitology and Microbiology the students were naturally divided into groups emphasising either parasitology or microbiology and ethics was introduced to one of these groups. The students already had non-

identical academic experiences due to different staff and some different courses. Introduction of ethics also corresponded with teaching of other transferable skills for the non-ethics group. 3) When a natural divide did not exist, the control group received identical ethics teaching after the post-test had been administered.

5.2 Moral competence

Moral competence reflects students' skills in moral reasoning, an ability to consider the perceived moral problem rationally and logically, an ability to take into consideration all relevant facts and prioritise between them in cases of conflicting principles and values. Moral reasoning skills can be assessed using the Defining Issues Test (DIT). Extensive literature supports its use, including the reliability and validity studies discussed in section 4.3.2. Previous research usage provides statistical information on the scores that can be expected for undergraduate science students and assistance in interpreting the gains due to educational efforts. DIT is also efficient and a low cost measuring tool because it can be group administered; it only takes 30 minutes to complete and it can be computer scored.

Despite the very satisfactory track record in validity and reliability of DIT and its economic use in terms of both time and money, it is not self-evident that it is suitable for measuring moral development in every context. Kohlberg's moral development theory and DIT are strongly related to the conception of justice and rights. The measure of moral development is tied to an ability to judge moral issues from a neutral and universal point of view. In the assessment of moral development, this narrow scope of DIT needs to be recognised and the results from DIT tests interpreted in the light of their only assessing one element of moral decision-making skills - an ability to interpret moral problems in universal terms (Rest *et al.*, 1999a).

One of the issues the DIT approach to considering moral development ignores completely is intentions and emotions of love and care. For example in the Heinz drug dilemma (Appendix IV - Moral Development Questionnaire) statement 'Whether Heinz is stealing for himself or doing this solely to help someone else' is classified as a stage 3 item, and in the Newspaper story a statement 'Whether Fred was using the

newspaper to stir up hatred and discontent' is similarly classified as a stage 3 item. If you are not a proponent of a totally consequentalist moral system, the intentions of moral agents should make a difference in your appreciation and analysis of the moral dilemma. Mature emotions towards caring for people one is close to should not be considered incompatible with moral competence, but in DIT, caring emotions are not included in the high scoring items.

DIT measures an ability to consider a wide range of issues in each of the presented moral problems. But because DIT is not an outcome measure - the choice of action does not effect the score - it does not rule out elements like compassion and care in moral action. In the four-component model described by Rest (1986), DIT measures elements in component 2 - an ability to make a moral judgement - not components 3 or 4 which relate to moral action - an ability to give priority to moral consideration and ability to behave accordingly.

In their professional life, scientists are required to think rather in terms of justice than compassion and an ability to consider wider consequences is an essential skill. DIT is designed to measure that element in moral development, and as that type of moral development is a core aim of ethics teaching in life sciences, DIT provides a valid and reliable measurement tool. The lack of importance placed on caring or intention, however, is worth noting at all times when DIT results are discussed, especially when it comes to more biomedical issues and working with animals.

The main draw-back of using DIT for measuring the benefits of short term educational interventions is its design to measure longitudinal change in moral development skills. These changes are rather broad-gauged alterations in thinking over an entire life span and are intended to represent fundamental, underlying structures of social thought rather than fine-grained descriptions of specific concepts and ideas, or detailed and refined changes (Schlaefli et al. 1985). Short term educational efforts are more likely to result in small scale changes, which thus may not be captured by DIT. The change after specific intervention is often only 1/10 or 1/20 of the effective range of DIT - changes within one stage rather than movement between stages (Baldwin et al., 1991; Schlaefli et al., 1985; Frisch, 1987). On the other hand, research has shown that these

gains are both retained and cumulative. The advances in moral development have a tendency to extrapolate to new situations and pervade the students' thinking in circumstances outside the classroom. Therefore, if an educational programme can show that it has produced gains in moral judgement - even if the gains are moderate or even small - then those changes may be worthwhile and the programme may claim to have had an effect upon life beyond the course itself (Rest, 1982).

A complicating issue with DIT is its culture and time dependent design. With the approval of Darcia Narvaéz, who is one of the co-creators of DIT, I changed the vocabulary of DIT slightly to suit British students. The change of American English to British English was carried out in the following stories:

- Heinz: The word 'druggist' was changed to 'pharmacist'; $-signs were changed into £-signs.

- Newspaper: The word 'mimeographed' was removed in front of 'newspaper'.

A further discussion with Darcia Narvaéz considered the change of 'Vietnam war' to just 'war'. Dr. Narvaéz considered this to influence the DIT results to an extent where maintaining their direct comparability with previous studies could be compromised and the word Vietnam was not removed.

Taking into consideration the reliability and validity of DIT, the adjustments carried out to suit British students and the down-side of relatively small changes expected to result, I included a 3-story DIT as my measure for assessing the effects of the ethics curriculum on students' moral reasoning skills. Course coordinators provided stringent time-pressures to keep the testing as short as possible. For this reason it was not possible to include the complete 6-story DIT.

5.3 Novel testing methods

Moral sensitivity has been of lesser interest in moral development assessment in comparison to moral competence measures. Therefore, the choice of reliable and valid assessment methods is relatively limited. The only reported measures imitate in

structure the DEST (Dental Ethical Sensitivity Test) developed by Rest and Bebeau for measuring the moral sensitivity of dental students. DEST is not only laborious and time-consuming to administer, but also designed to measure moral sensitivity in professional students entering a profession with a defined moral code, appreciating patient autonomy, confidentiality, and beneficence. Science as a profession does not have a similar recognised value-base for students to adopt. Attempts to develop such a code have been made (The Institute for Social Inventions proposed one in 1987, and Rotblat (1999) called for one) but no code has been generally accepted. Also, the moral problems in dentistry, and more generally in other health care related subjects, are situated in personal encounters between the professionals and their patients. The moral problems involved have immediacy and intimacy in both their recognition and in the impact of the resolutions. In science, on the other hand, moral problems are more abstract and the consequences more distant and therefore it is often a matter of making a moral decision regarding one's personal conduct in isolation from the direct consequences. These differences in the nature of moral problems in health care and science reduce even further the applicability of DEST type measures to assess the moral sensitivity of science students.

In the absence of a suitable assessment method for the development of moral sensitivity, novel methods needed to be created for the purposes of this research. The core elements of moral sensitivity are the ability to recognise moral issues in problems that contain other - often prominent - elements, and an ability to tolerate moral relativism and ambiguity and to act as a moral agent.

A decision was made to develop a two-prong approach. One element in the moral development questionnaire would test the students stage of moral development on the Perry scale and another element would concentrate on a more straightforward recognition of moral elements in a potential moral problem.

5.3.1 Perry's theory of ethical development

Perry's theory of ethical development describes developmental change in moral sensitivity on the meta-level - the process of coming to realise and accept moral

uncertainty and acquiring skills to deal with it. Perry's developmental methods have previously concentrated on measuring intellectual development and attitudes towards learning and teaching, but the methods of measuring this type of change seems applicable to assessing moral development as well.

The first task was devising descriptions of moral development positions according to Perry's theory. A detailed description of Perry positions can be found in Appendix III - Perry's Scheme, but in short the nine original positions are the following:

1. Basic Duality
2. Multiplicity pre-legitimate
3. Multiplicity sub-ordinate
4. Multiplicity correlate or relativism subordinate
5. Relativism correlate, competing, diffuse
6. Commitment foreseen
7. Initial commitment
8. Orientation in implications of commitment
9. Developing commitment

Instead of devising descriptions of all nine positions individually, I adopted the approach used frequently in assessing intellectual and attitude change employing three Perry types; A, B, and C. The grouping of nine positions into three main types simplifies the use of Perry's scheme in pen and paper tests. To identify students using the nine-position approach seems feasible only when qualitative research methods are employed. As discussed earlier qualitative methods are labour-intensive and thus not suitable for large scale projects with limited resources. The 3-type approach will not have the intricacy of the nine-position approach, but it does capture the essence of development described by Perry.

Type A is formed of Perry Positions 1,2&3 and can be labelled as 'Safety in Dualism'. Perry A students perceive the world in either purely dual terms of 'we-right-good' vs. 'other-wrong-bad', or with diversity as an unwarranted confusion regarding poorly qualified authorities. The right answers to everything exist in the absolute and they are known to the authority whose role is to mediate (teach) them. The only role for

discussing moral questions in class is because the authority (teacher) wants the students to learn and find the right answers themselves. There is no doubt about the existence of the right answer in moral questions and there are only right or wrong answers to moral questions, no better or worse ones. Students either do not doubt that authority knows the right answers to all moral questions or they accept that uncertainty is unavoidable at present, and the authorities will find the rights answers in time. The problem then arises of the legitimacy of the authority to judge the students' answers if they do not *yet* know the right answers themselves. The student's own role and level of personal responsibility consists of obedience, and the responsibility of the authorities is to teach them what is right and what is wrong - even when the students have doubts whether the authorities can actually fulfil their role. Students may feel hostility towards the vagueness of moral answers. Alternatively students may accept moral discussions as mere exercises without strong emotional response or involvement.

Perry type B is a combination of Positions 4, 5 & 6 and can be described as 'Distress in Relativism'. Type B students have accepted the multiplicity and relativism of moral questions, though sometimes reluctantly. Multiplicity is possibly accepted in some areas only, and confusion remains over what is legitimate multiplicity and what is not. Some students come to accept multiplicity, because they believe the authority to want it. True relativism - that knowledge and the appreciation of values (including authority's) are relative and contextual and absolute right/wrong choices are exceptions to the rule - once accepted, threatens the student with loss of identity and disorientation. The student has three ways to deal with the threat 1) to deny any meaning beyond one's immediate passive responses; 2) to exploit the situation deliberately by becoming an active opportunist of the relative reality; and 3) to acknowledge the need for commitment, though not yet making any.

Perry Type C is a combination of Positions 7,8&9, and best described as 'Comfort in Commitment'. The Perry C student resolves the threat of identity loss and disorientation by making commitments to moral values. This process will originate with commitment in one defined area - like choosing one's career - and from there is taken to further and further areas in life. Commitment is soon paired up with

responsibility for the chosen values and how to fulfil them. Commitment will lead to affirmation of identity among multiple responsibilities and to the realisation that commitment is an ongoing process through which the students can express their lives.

The Perry-scheme is open to criticism in being too soft on relativism. If it is interpreted as saying that all values are relative, the criticism seems well-placed. Alternatively, if relativism is interpreted as a characteristic of *applying* values it seems more robust. People are rarely absolutely sure that they have made the only possible right choice in a moral dilemma. We acknowledge alternatives and their potential value. Still we make decisions and commit ourselves to the value application we have chosen. In the particular sorts of moral and ethical questions scientists face as professionals, relativism in the application of core values is often inevitable.

5.3.1.1 Perry Questionnaire design

In order to assess students' progress within the Perry scheme according to the type descriptions collated, in the first instance I generated 68 statements to describe Perry positions - 15 for Perry A, 33 for Perry B, and 20 for Perry C. For each Perry type, the questions covered the five elements significant in moral development according to the Perry Scheme :

1. Source and type of moral answers
2. Role of authority
3. Nature of multiplicity
4. Personal responsibility and relationship with multiplicity
5. Purpose of moral discussions (see Appendix III - Perry's Scheme for full descriptions)

In order to improve the quality of the statements and to compose a list of statements where all types were equally represented, the number of statements was further reduced to 30 (10 of each type). These statements were put together in a random order and sent, together with the descriptions of the Perry scheme, to 8 independent judges (see Appendix VI - Perry Judges for details). Five of the 8 judges were familiar with the Perry scheme before they were asked to evaluate the statements and three had not

previously worked with the Perry scheme, though they had experience in other ethical development theories.

For each statement, the judges were asked to identify which Perry type would strongly agree *and* which strongly disagree with it. The results from the judges were decisive (inter-rater consistency >0.8) on 16 statements out of which 12 were statements designed to represent either Perry-type A or C. The remaining 8 A and C statements, which did not produce conclusive agreement by the judges, were re-phrased to provide a better match with the type descriptions and sent to four new Perry judges familiar with the Perry Scheme, whose judgement resulted in an agreement over the re-phrased statements.

Because the Perry questionnaire was to be designed with computer-scoring in mind, the possible alternative formats were to ask students to respond to Perry statements using either Likert- or Osgood-scales.

In the Likert-scale questionnaire, students are asked to state whether they agree or disagree with a given statement, often on a 3-, 5-, or 6-point scale. The hypothesis for using Likert-scales to measure development on the Perry Scheme, is that type A students would agree with the statements depicting A-type views on moral questions and disagree with C statements - the same goes for C students agreeing with C-type statements and disagreeing with A-type statements. For example a statement 'Moral questions have absolutely right answers just like scientific ones' which is designed to describe the thought pattern of a Perry A-type student, would attract A-type students to agree or agree strongly with the statement and C-type students to state disagreement. Likert scales provide straightforward assessment of A- and C-statements, while B-statement analysis is more complicated.

In the Perry scheme of moral development, the major shift is identified by transformation between A and B types. For this reason, when students reflect on type B statements and indicate acceptance of multiplicity in moral problems they have taken a major step from dualism to multiplicity. For this reason the analysis of B-type statements should not be neglected. Development into Perry C type builds upon this

realisation: it is not as dramatic a change as between A and B, but a more subtle growth in accepting responsibility and finding confidence in relativism. This structural element creates difficulties in designing a Likert-scale Perry questionnaire. If Perry C students build on the foundations of Perry B, it is likely that C-type students are going to agree with the B statements as well as the C statements, as B statements are in accordance with their basic understanding of the moral world - i.e. moral answers are relative and uncertainty is an inseparable element of morality. On the other hand, type B students should not agree with type C statements, which could be used as an analytical approach to finding whether students are B or C. A further problem is with analysing disagreement with a B-type statement, which can be interpreted either as A or C.

To reduce this problem, it is possible to phrase the type B statements so that they emphasise discomfort with moral relativism and the novelty of its recognition, to avoid C-type students agreeing with them. Doing this runs the risk of B-statements being phrased in a negative light, which may deter students genuinely at stage B from agreeing with them. Even though B-stage is a transitional stage towards C, it should not be described in the negative terms of lack of commitment or moral hijacking. Alternatively it is possible to ignore the B-statements in the final analysis, as suggested by Gray (1997) and use the Perry Index (PI) (see section 4.5) as an analytical tool. This raises the issue of how sensible it is to include B-statements in a questionnaire if they are not providing any further information in the final analysis. The only justification for the inclusion of B-statements when they are not analysed would be to offer B-type students statements they can agree with, to avoid discomfort in the testing situation.

An alternative to the Likert scale, is the Osgood scale described in section 4.5. In the Osgood scale, students are presented with opposite statements (A and C types) and asked to indicate on a scale from strongly agree, agree, neutral, agree, to strongly agree, their preference on the stretch of two polar views. This approach removes the problem of C type students agreeing also with B statements, as no B statements are used. B students are identified as students who choose to be neutral between the poles.

As a result of this analysis, I decided to use the Osgood scale for the Perry analysis (see Appendix VII - Preliminary Perry Questionnaire). Altogether ten A and C statements, which had all been rated at minimum 80% agreement among the external judges on their validity in representing the designated Perry type, were included in the questionnaire. Three of the questions related to the source and type of moral answers (questions 2, 4 and 7), one (question 3) related to the role of authority in judging moral answers, two questions represented different perceptions on the role of moral discussions (questions 1 and 6) and three questions described the different approaches to personal responsibility and relationship with multiplicity (questions 8, 9, 10). The A and C statements were randomly allocated on either the left or right hand side of the Osgood scale. The extreme ends of the scales were labelled as 'Definitely my opinion' (A or C), the next ones 'More or less what I believe' (Ab or Cb), and the middle one as 'I can't decide' (B).

5.3.1.2 Preliminary results with the Perry-Questionnaire

The Preliminary Perry Questionnaire was piloted with 70 L3 students from four different honours courses (Genetics and Molecular Biology, Pharmacology, Physiology, and Zoology) and with 60 L1 students in May 1999 during the last week of teaching in Term 3. Students were given an option not to participate in the trial, but only very few students chose to do so. Unfortunately the attendance during the last week of teaching was relatively low in L3, which may have affected the preliminary results as the group was self-selected.

In the trial of the Perry questionnaire, when including all ten Perry statements in the analysis, both L1 and L3 student responses were dominantly C or Cb.

	A	Ab	B	Cb	C
L1 (n=62)	3% (17)	5% (28)	18% (105)	37% (217)	38% (226)
L3 (n=72)	2% (13)	6% (43)	13% (87)	39% (267)	40% (279)

Table 1: Perry response distribution frequencies

The statistical difference between the L1 and L3 results was tested using chi-square, and was significant at p<0.005. The highly significant statistical difference should be interpreted in the light of the general trends of both L1 and L3 students having very similar answering patterns with a strong dominance at the C end of the scale. This indicates that most students have reached at least partially the highest level of Perry moral development.

The results were then broken down into the different elements in Perry moral development, which revealed more diverse developmental patterns in the response frequencies (the results are presented as percentages).

L1	A	Ab	B	Cb	C
I Source and type of moral answers (Q2, Q4, and Q7)	2	3	14	41	40
II Role of Authority (Q3)	12	8	12	37	32
III Nature of multiplicity (Q5)	2	3	7	44	44
IV Personal responsibility and relationship with multiplicity (Q8, Q9, and Q10)	2	8	36	31	24
V Purpose of moral discussions (Q1, and Q6)	2	2	7	34	55

L3	A	Ab	B	Cb	C
I Source and type of moral answers (Q2, Q4, and Q7)	1	6	11	39	43
II Role of Authority (Q3)	7	6	7	36	43
III Nature of multiplicity (Q5)	1	6	9	36	48
IV Personal responsibility and relationship with multiplicity (Q8, Q9, and Q10)	2	10	25	38	25
V Purpose of moral discussions (Q1 and Q6)	0	2	7	41	51

Table 2: Perry elements

First it is noticeable that L1 students chose consistently (apart from element V) more Cb statements than C statements, while the trend is opposite with L3 students (apart from element IV). This can be interpreted as an increased confidence in personal views regarding moral issues gained by two years of formal teaching.

Second, element IV (Personal responsibility and relationship with multiplicity) produced a significantly different answering pattern to all other elements. Both L1 and L3 students responded to element IV items more from a B-type view point than to any other element. This trend was stronger in L1 responses. This result is not surprising considering that the highest levels of development concern issues of responsibility and commitment above all others. The C positions build upon the realisations of multiplicity and relativity gained during the B-positions, and deepen the understanding in particular towards personal commitment and personal acceptance of moral multiplicity. This indicates that while students have reached an understanding of moral issues typical of the C-stage, they are still in the process of forming personal commitment to relative moral values and finding a balance between freedom and responsibility.

Third, it is worth noting the high frequency of pure A choices in element II. This indicates that students otherwise viewing the world from a dominantly type- C- viewpoint, find that mode of thinking less suitable in the context of assessment. A similar trend of retreat to more elementary methods regarding assessment has been found by Gray (1997).

Finally, attention should be drawn to element V, which shows the positive attitudinal capability of students to discuss moral issues during ethics courses. There seems to be no barrier to engagement in a moral discussion which will not lead to a definitive answer. The two discussion questions also provide a consistency check in the questionnaire, as both questions query the same element with different wording. In both levels, the consistency check provides confidence that students answer the questions reflectively and sincerely.

The Perry Index (PI) and Success Index (SI) discussed in section 4.5 provide less useful ways of analysing the Perry questionnaire using an Osgood scale.

The difficulty with the use of these indexes is their emphasis on Perry A statements and their grouping of Cb and C statements together. In the preliminary Perry results, 23 students out of 41 for whom an SI score was calculated had an SI score of one,

indicating that they did not choose any type A statements. The SI scores not equal to one ranged from 0.43 to 0.8. A result of 0.8 indicates 9 choices of type C statements and one choice of type A statements; the lower scores indicate different combinations of all A and C type responses. Because of the low frequency of A statements chosen by the L1 and L3 students, the Success Index is not a functional measurement of their developmental stage or developmental advances.

In need of an alternative method to SI and PI, to calculate a Perry score, I developed a Personal Perry Score (PPS1). PPS is calculated by assigning each statement choice a representative value: these are then added together for a Personal Perry Score. As these numbers only represent a certain point on the Perry developmental scale, no arithmetic can be performed with them, but they can be taken to represent a certain level of Perry development and averages can be calculated for the group. A simple approach to calculating a PPS is to assign single step values to the five alternative statements on the Osgood scale: the point closest to the A type statement is given a value of 1, the next closest a value of two, the middle option a value of 3, the second closest to type C statement a value of 4 and the closest to the C statement a value of 5. Using this method 79 (36 L1) Perry questionnaires were analysed and the absolute score was divided by ten to give an average score for one statement:

Average PPS1	4.04	sd 0.397	range 3.2-4.9
Average PPS1 for L1	4.06	sd 0.405	
Average PPS1 for L3	4.02	sd 0.389	

Table 3: PPS1

The scores indicate that L1 students have a slightly higher developmental score than L3 students. This result is not supported by the previous frequency analysis, where L1 students were shown to choose lower level items in greater frequency than L3 students. The standard deviation is marginally higher for L1, which indicates slightly more spread out answering patterns for the L1 students.

An alternative Personal Perry Score (PPS2) was calculated by raising the values given to the responses in PPS1 to power two. This was done to increase the weight of C and

Cb responses over the A, Ab, and B responses. The developmental jump from A to B is dramatic and significant, while the developmental tasks beyond B are more subtle, but often require more emotional and psychological labour to achieve. To represent this increased requirement of emotional labour, the higher levels on the Perry scheme were allocated a higher score, giving students who have reached the Cb and C levels 'credit' for their achievement. The new values attached to each response in PPS2 are: A=1, Ab=4, B=9, Cb=16, and C=25. The score was then divided by 10 (the number of questions) to get a question average. The same 79 responses were re-scored using PPS2 and the results were:

Average PPS2	17.2	sd=3.07	range 10.6-24.1
Average PPS2 for L1	17.2	sd=3.16	range 10.6-24.1
Average PPS2 for L3	17.3	sd=2.91	range 10.6-21.6

Table 4: PPS2

Changing the values attached to the different Perry choices did not change the significance of the results. Again the difference between L1 and L3 students are insignificant (p>0.05) and the standard deviations indicate similar answering ranges for both groups.

PPS can also be calculated for the different Perry elements separately. For the 79 responses PPS2 was calculated for element IV 'Personal responsibility and relationship with multiplicity':

Average PPS2 for element IV	15.12	sd=3.56	range 7-25
Average PPS2 for L1 for element IV	14.91	sd=3.79	range 7-25
Average PPS2 for L3 for element IV	15.64	sd=2.87	range 12-22

Table 5: PPS2 for element IV

These question averages show the difference detected between the answering pattern of L1 and L3 students in the frequency analysis given before. The L1 student

answering pattern also has the higher standard deviation, which indicates larger developmental differences amongst L1 students.

These results indicate that PPS1 and PPS2 for all Perry elements are not useful scoring methods for comparing L1 and L3 groups as PPS does not pick up the differences in answering patterns seen in the frequency analysis. PPS1 or PPS2 may on the other hand, provide a tool to measure personal development during the academic year. Measuring PPS2 for particular elements produces results concordant with those found in the frequency analysis and thus is a useful tool for providing numeric data on those aspects of the Perry questionnaire.

The preliminary study with the Perry questionnaire suggests that the tool does measure developmental differences between students, but that it is not a particularly sensitive measure for a student population which dominantly views moral problems from the C position. At the same time the differences between student answers and between L1 and L3 students as groups supports the use of the method as a developmental indicator.

The preliminary study also suggests the need for minor changes to the questionnaire.
1. The questions representing each element should be more randomly distributed to avoid a cluster of personal responsibility questions at the end of the questionnaire.
2. The key to the response types should be changed for the B type from 'I can't decide' into 'Neither of the statements represents my view'. This change was initiated by the students, who indicated a problem with deciding which box to tick, when their difficulty was not that they could not decide between the two polar statements, but that they believed neither of the polar statements represented their view. This re-phrased key is also better related to the Osgood-scale hypothesis that B type students find neither of the statements represent their views on the issue.

5.3.2 Recognition of moral issues

The test of moral sensitivity should measure the recognition of the moral elements and appreciation of them in the decision-making procedure, not the abilities to use higher level moral considerations as the basis for decision-making. Therefore the test elements for moral competence (DIT) are not sufficient for measuring moral sensitivity. Neither will the Perry questions target the interpretative element in moral reasoning, as the Perry questions target meta-understanding of the nature of morality, not moral sensitivity.

To assess the interpretative phase of moral decision making (ethical sensitivity as defined in Part II 3.2.2) it is necessary to use "unstructured problems." A moral problem is unstructured when it does not directly indicate the moral issues involved, either by describing them in the problem narrative, or by giving moral statements to choose as possible solutions or considerations for arriving at a solution. The problems used in DIT are 'structured' moral problems, because the narrative structure describes a particular moral dilemma (e.g. Heinz should steal and save his wife, or should not steal and not save his wife) and the considerations for the decision are all part of the moral deliberation process. An 'unstructured' moral problem is thus a problem scenario which has moral components, but where these components are not self-evidently apparent, and a solution to the problem can be arrived at without ethical considerations.

It is impossible to measure moral sensitivity with a 'tick-a-box' method. Any such method would have to include some level of pre-established moral analysis, which would have taken place before any statements to choose from could have been produced. For example, a test protocol, which gives students an unstructured moral problem and then offers several ethical and non-ethical elements to choose to include in their deliberation, would not test the recognition of moral issues, but the importance students place on these issues. As was discussed earlier in section 4.4, people can recognise and discriminate and thus prefer an idea before they can paraphrase it or before they can spontaneously produce the idea in a response to a story dilemma.

Moral sensitivity is precisely the measure of the spontaneous recognition of moral issues, the interpretation of a situation in moral terms.

It is important to remember that moral sensitivity is not a measure of moral decision-making methods. High scores on moral sensitivity tests do not indicate that one would necessarily have sophisticated methods of solving the problem with appreciation for the moral elements one has recognised. Also the four-component theory of moral action and moral development described in Part II, section 3.2.1 does not demand the abilities of moral sensitivity and moral competence to be highly correlated. Research has shown (Bebeau and Brabeck, 1987) that the correlation between DIT and moral sensitivity tests (DEST) has been consistent, but very moderate at 0.2 -0.3.

Therefore, the nature of moral sensitivity requires the test of moral sensitivity to be qualitative, to allow students to respond to an unstructured problem with only minimal guidelines or pre-established thought-patterns. This type of qualitative data can be collected either verbally in an interview or in a written form. DEST used both methods, which provided equally valid and reliable date (Bebeau *et al.*, 1985), while the interview scores yielded higher estimates of moral sensitivity, as judges felt they had a better opportunity to confirm their judgement from verbal responses. Interviews may produce more data, but they are more laborious to administer. The large student population at Glasgow University dictated the choice of a written test-format.

5.3.2.1 Pilot studies on moral sensitivity

In the first trial, 17 (one response was removed the from analysis due to illegible hand-writing) Zoology students were asked to respond to the following scenario:

> The local health board is planning a pre-natal screening programme for cystic fibrosis (CF). CF is caused by a mutation in chromosome 7. The gene is long consisting of 1,480 base pairs, while most of the mutations are restricted to a relatively small area, more than 200 mutations have been found through-out the gene. The faulty allele (▲F508) that is responsible for 70-80% of CF cases is missing three base pairs, coding for one amino acid in the protein. The rest are caused by different mutant alleles in the same gene. There could be thousands of mutant alleles not yet identified. Approximately one in 20 to 25 in the Caucasian population is a carrier, which means that in about 1 in 625

couples both partners are carriers and they therefore have a 1 in 4 risk to have a child with CF.
In general there are two types of CF, one that affects lungs and pancreas and another that affects just lungs; about 90% have pancreatic insufficiency. Although symptoms range in severity, many patients suffer a prolonged period of increasingly severe handicap. Life expectancy of CF patients in the UK is forty years.
CF test can be carried out by a simple spit test, where the test subject provides a small amount of saliva on a test dish.

Please list below ALL the issues you think should be considered in making the decision on whether to start a programme of pre-natal screening for CF.

Students recognised zero to ten issues based on this scenario, the average being 4.9 issues per response. The difficulty with analysing the responses to this scenario was largely due to the multitude of ethical issues involved in pre-natal screening, in particular screening for a disease like CF which allows the sufferers to lead a relatively long life before succumbing to the disease.

The students picked up the following issues in their responses (if the same issue was mentioned twice in the same response that was counted as one response):

1. Cost of the programme (7)
2. Certainty of the diagnosis (5)
3. Frequency of CF in the area (1)
4. What age can CF be detected (1)
5. Safety of screening procedures to the mother and child (1)
6. Family history (1)
7. Alternative (better) use of the resources (2)
8. Abortion -> should it be an option?, would it increase? (8)
9. Rights of the child (to life) (9)
10. Parents' right to know (1)
11. Need of counselling (5)
12. Increased opportunity for parents to choose (3)
13. Increased parental stress (5)
14. Is this messing with the gene-pool, with the survival of the fittest? (2)
15. Will the test provide alternative information which requires parental decision (1)
16. Compatibility with religious and other beliefs (3)
17. Is this Playing God? (3)
18. Quality of life for the child, parents or siblings (6)
19. Will this lead to screening for other diseases (1)
20. Do people have to suffer debilitating diseases? (1)
21. Care available for CF children (1)
22. Feelings of staff (1)
23. Benefits of early detection to the welfare of CF sufferer (1)

The first six could be interpreted as non-ethical considerations, while the rest are most naturally viewed with at least some attention paid to ethical and moral consequences, though it is possible to view, for example, parental stress in purely factual terms without attaching any moral value to that stress. The students seemed to view the issues from an ethical point of view, though at times this could not be conclusively inferred as the responses were very short (one or two words). In the preliminary analysis, benefit of the doubt was given to consider even these short responses as ethical considerations.

Students recognised an average of 3.1 moral issues, with individual scores ranging from 0 to 6. The sophistication of the answers varied greatly from short remarks such as 'Playing God?' to well-worded and thorough responses such as 'Pre-natal screening in general - if it is done for one condition, and becomes accepted, then it may spread to other conditions - risk of parents 'selecting' their children on the basis of traits they believe are important.' In the preliminary analysis, no importance was based on the quality of responses, but it became obvious that a scoring guide which takes into account the sophistication of the answer needed to be developed.

Because of the multitude of ethical problems in the pre-natal screening story, a decision was made to develop unstructured scenarios which would not have such large numbers of ethical problems present, to provide a more valid and reliable platform for scoring the responses. At the second stage of the pilot study, three stories were given to both L1 and L3 students at the same time as the Perry questionnaire in May 1999. At this second stage, students were asked to list no more than five issues they believed should be considered in making the decision regarding the issues presented in the story. A grid with five boxes was placed below the scenario to indicate the length of the responses expected. Students were also asked in all three scenarios what they believed should have been done (Yes, No or I don't know), and which of the issues they listed should be the most predominant one in the decision-making procedure.

Story 1 - Research laboratory

One of the stories was developed to highlight the problems with making decisions regarding developing preliminary research findings. 17 L1 students and 18 L3 students responded to this story.

> *A small research laboratory has made a break-through in discovering a gene defect that triggers acute childhood asthma together with environmental exposure. The research team has been funded through governmental research councils. The grant is due to run out in a few months time and there are no guarantees of future funding. A large pharmaceutical company has made a bid for the lab, promising to employ the scientists as long as they will sell the patent rights to the company. This would mean a move to a new location and not being able to continue with the current support staff. There is an alternative opportunity to gain further funding from the research council which would allow for the laboratory to stay independent and possibly expand its facilities, but for now there are no guarantees whether such funding will be made available.*
>
> *Should the research laboratory accept the offer?*

All the L3 students gave at least one issue they believed important in making the decision, while five L1 students returned the paper with no issues listed, though three of them had responded to the question whether the laboratory should accept the offer. The three that responded to the question were included in the further analysis, while the other two were disregarded from further data, because it was not clear whether they had even considered answering the question. The same exclusion criterion was applied to all three stories.

The L3 students listed on average 3.5 issues, while the L1 students' average was 2.2. The distribution of responses according to the number of issues stated was the following:

	no issues	1 issue	2 issues	3 issues	4 issues	5 issues
L1	20%	13%	20%	27%	13%	7%
L3	0%	22%	11%	28%	17%	22%

Table 6: Story 1 response rates

In advance I hypothesised that the following moral elements exist in the scenario:

1. How will the support staff be employed in the future?

2. Who decides on the research topics in the future?

3. Whether future research findings are classified information?

4. How can the research findings be used to benefit those suffering from asthma?

5. What is the role of animal testing in the development of the discovery into a drug?

6. Who will get access to the patent?

7. Whether the discovery should be patented?

In 13 student responses, the issue of support staff employment was mentioned and a further six believed the security of employment of the academic staff was an issue to consider. Two students raised the issue of who decides the research topics in the future and only one student was explicit about the issue of classified information. On the other hand, the topic of benefits to asthma sufferers was raised in the student responses 13 times. Some elaborated the issue to include the importance of getting the research done as paramount; others worried about access to the drug ("would the pharmaceutical company be the best option to guarantee access for all?") and who would profit from subsequent research, asthma sufferers or the company. A further three students pondered on the rights of the commercial company to exploit this type of research and another two considered the general benefits of such research ("Is it beneficial to public health?"). None of the students perceived animal testing as an issue to be considered in the decision-making procedure. Eight students raised the issue of patent rights, while only one referred to general considerations of whether patents should be allowed or not.

The students raised a further seven issues that were not considered in the design of the scenario. Seven students considered alternative sources of funding, which might secure the research, but would not have the negative effects of pharmaceutical company take-over. Five students raised the issue of credit: who would be praised for the discovery of a possible drug, the research team or the company who buys them out? A further two students considered whether this research advances science *per se*.

Finally the following three issues appeared once each in the responses: where is the new location, how convenient would the move be, and whether the research would succeed.

Table 7 summarises the data:

Employment of support staff	13	Who decides on the research	2
Benefits to asthma sufferers	13	Benefits of research in general	2
Patent rights	8	Advances science?	2
Other alternatives	7	New location	1
Employment of academic staff	6	Convenience of the move	1
Who gets the credit	5	Facilities in the new laboratory	1
Rights of the company to exploit the research results	3	Animal testing	-

Table 7: Story 1 distribution of responses according to themes

Story 2 - Pharmaceutical milk

The other two stories were adaptations from case-studies found in Bruce and Bruce (1998). This first story describes a research scenario for genetically modifying a cow to produce pharmaceutical milk, described by Wilmut (1998).

> *A research group is planning a project on creating a cow that would produce milk containing a protein that could be used to treat patients with cystic fibrosis. Other pharmaceutical methods to produce this protein have not been successful or they have been very expensive. The plan is to introduce a new gene from another animal into the genetic sequence of the cow that directs the production of the mammary gland to change it from producing normal milk into producing a pharmaceutical milk containing the desired proteins. The new gene will be introduced by nuclear transfer, a technique also used in cloning. The group hopes to develop its research findings into a commercial product.*
>
> *Do you think the research should go ahead?*

This story was given to 23 L1 students and 21 L3 students (three empty L1 response were removed from further analysis). L1 students gave an average of 2.1 issues to the

story, while the average for L3 students was 4.0. The distribution of responses according to issues considered was:

	No issues	1 issue	2 issues	3 issues	4 issues	5 issues
L1	25%	5%	25%	30%	15%	0%
L3	0%	5%	14%	14%	14%	53%

Table 8: Story 2 response rates

The main ethical issues presented by Wilmut (1998) regarding this type of research are:

1. Mixing genes between species - is there a fundamental (deontological) reason why mixing of genes from different species should not be done, even when there is no harm to the host animal and medical benefits from doing so?

2. Risks to humans - risks to both patients and general population if there are infectious agents in the pharmaceutical milk or possible side affects due to slightly different protein structure; could the milk or the carcass of the animal be used as food (legal and labelling issues)

3. Environmental impact - mixing of modified and unmodified animals, ability to destroy modified live-stock problems arise.

4. Animal welfare - the research requires surgery on several animals to obtain each founder animal, which causes anxiety and pain to the animals; the protein may leak from the mammary gland and become active in the producing female, which may cause adverse reactions, even death, in the animal.

In their responses students raised all of these issues. The dominant concerns related to risks to humans and animal welfare. The risk to humans was most often perceived in the form of side affects of pharmaceutical milk to non CF sufferers (19 responses), while one student recognised as an issue the right to refuse treatment, which led the student to consider issues of labelling. The responses (20 responses) which could be classified as considering animal welfare mainly considered the possible discomfort of the host animal, while a few responses included concern for the care of research

animals from a more general perspective, and one student raised the question of whether the milk would need to be tested on animals before it could be released as a pharmaceutical product. The possible ethical issue of mixing genes was recognised by nine students, out of whom eight considered the basic question whether it is right to do so, and one pointed out the issue of safety related to the technique. A further six students recognised the problems of possible mixing of genes as an environmental hazard.

The students also listed issues they considered important in a further four general categories, not proposed by Wilmut. These were:

1. Public opinion - would the public accept this research, and how much education would they require to understand it. This issue was raised by five students.

2. Scientific viability of the research - would CF patients have access to the treatment and would the treatment work i.e. improve their quality of life. These issues were discussed in 10 different responses. The responses ranged from 'will it work?' to 'Will drinking this pharmaceutical milk actually benefit the sufferers of Cystic Fibrosis'.

3. Costs and benefits - this consideration ranged from a very basic cost-benefit idea to considering alternative costs and profit making motives of the companies developing this treatment. Students also considered alternatives to this research in providing the same medical advancement to CF sufferers. 30 issues relating to this theme could be identified in the student responses. The response types varied from 'cost', to 'Who would benefit more: the patient who will be receiving the same treatment or the company who will be getting better profits from cheaper production'.

4. Advancement of science - whether this research would advance scientific knowledge *per se* and if so, we should do it. Two students considered these issues in their response.

Table 9 summarises the frequency of issues raised in the student responses:

Cost and benefits	30	Ethics of mixing genes/cloning	8
Animal Welfare	20	Environmental impact of cross-breeding	6
Risks to humans	19	Public opinion	5
Scientific viability	10	Advancement of science	2

Table 9: Story 2 response rates according to themes

Not all the responses could be classified as ethical, in particular within the cost-benefit and scientific viability categories many responses did not raise any ethical considerations relating to these issues. The scoring guide would have to be able to discriminate between ethical and non-ethical responses within all categories, as well as distinguish between different levels of sophistication in the responses.

Story 3 - Plant virus

A second story adapted from Bruce and Bruce (1998) describes a research situation involving the introduction of genetically modified plant viruses to enhance the nutritional value of the host plant as described by Wilson (1998).

> *A research group is considering a project on developing more nutritious plants by using plant viruses. The aim is to genetically modify these viruses so that when they act on the plants, the plant tissue will produce high levels of novel proteins which will increase the essential dietary value of the plant. Over 900 natural plant viruses have been described by scientists. The viruses studied so far are pathogens in the plant only and humans digest and handle them continuously with no ill effect. The genetic material of natural viruses has not been found to interact with the genes of the host plant. The researchers hope that the new plant varieties could be used in developing countries.*
>
> *Do you think the research should go ahead?*

34 L3 students and 23 L1 students responded to this story. Two empty L3 responses were removed from the analysis. The response rate varied between zero and five. The L1 response rate averaged 3.1. and L3 averaged 3.6. The distribution of response rates was the following:

	No issues	1 issue	2 issues	3 issues	4 issues	5 issues
L1	13%	9%	4%	31%	17%	26%
L3	9%	3%	19%	6%	13%	50%

Table 10: Story 3 response rates

Wilson approaches this technology from the point of view of opportunities, while in turn each opportunity fosters potential harms. Eight pairs of harm-benefit opportunities can be identified:

1. Environmental risk - what are the risks and consequences of the modified virus spreading to wild species.

2. Animal welfare - opportunities to produce cheaper veterinary products, but also potential risks to animals feeding on the plants.

3. Developing country issues - opportunities for more nutritious and healthy products for developing countries with a potential problem of developing country farmers not being able to afford them or becoming dependent on one seed provider.

4. Food safety - plant viruses have not been found to infect humans who consume them, but theoretical potential exists.

5. Commercial driving force - opportunity to high profit margins, which opens a possibility of abuse.

6. Land-use - the cultivation of these crops may increase the pressure on land use for food.

7. Will it work - is this a worthwhile research cause.

8. Ethics of abandoning the technology - would it be ethically right to abandon a technology with great potential because of possible risks.

The students raised all but one of these issues in their responses, only the consideration of NOT doing this research was not explicitly touched on in any of the student responses. Environmental risk was recognised in 31 responses. Some students considered the issue more widely with inclusion of threats to biodiversity and how the virus may affect other genes in the host plant. The issue of animal welfare was raised

in 6 responses. The developing country issues were raised by 25 students, out of which

five considered the technical suitability of these plants for tropical/arid environments.

The responses ranged from 'Whether third world farmers can afford to buy this GM

crop - ethically nice to develop it, but the farmers can't pay enough to recoup the R&D

costs' to 'Developing countries need help'. A further 30 responses recognised the food

safety issue. The responses ranged from 'Is it safe?' to 'I feel there is a safety issue -

do we really know enough about what we are doing to be carrying out such a project'.

Problems with the commercial driving force became apparent in many of the

developing country issues, while only one response raised the issue independently;

similarly the land use issue was raised in only one response. The issue of 'will it work'

was recognised in 36 responses. The students responses relating to this theme covered

a wider perspective than just scientific viability, including considerations of

alternative costs of resources used in this type research (other ways to get the same

benefits) - a better title to represent these responses would be 'Is it worth it?'. The

responses ranged from 'Consider the benefit of increasing the nutritional content [of

the plant], is it very marked? Is it necessary, or justified?' to 'benefits?'

The students also raised the basic issue of genetic engineering - whether there would

be inherent reasons to reject this type of research (3 responses), or doubts about the

safety of the technology (one response). Five students also recognised the issue of

research safety in dealing with viruses i.e. how to make sure the researchers are not

harmed and that the virus will not escape. 21 students likewise considered the source

of funding and the cost of research in relation to profits, while five students raised the

issue of public acceptance of this type of research.

Table 11 summarises the distribution of responses:

Is it worth it?	36	Animal welfare	6
Environmental risk	31	Issues of research safety	5
Food safety	30	Public perception	5
Developing country	25	Issues of genetic engineering	4
Cost and funding	21	Land use	1
		Commercial driving force	1
		Ethics of abandoning the technology	-

Table 11: Story 3 distribution of responses according to themes

Similar to story 2, the responses included both ethical and non-ethical ones, and varied greatly in their depth and sophistication.

The table containing the preliminary data from all three stories can be found in Appendix VI 'Moral Sensitivity - preliminary results'.

At this point a decision was made not to use story 1 further. The decision was based on the number of issues it raised, a characteristic which would provide a complicated platform for developing a scoring guide. This characteristic was considered more significant than the distributional pattern of the number of issues students raised - a pattern which provided good scope for development as a majority of students in L1 listed less items than L3 students, and the frequencies were more evenly distributed between different number of responses than with the other two stories.

The decision between using story 2 or 3 in the final moral development questionnaire was a trade-off between a more concise list of elements (story 2) and a more even distribution of issue frequencies in the responses (story 3). The total lack of L1 responses with five identified issues for story 2 was a reason for concern as the L1 students had completed a lab discussion on cloning, which included considerations of creating mammals to produce pharmaceuticals in their milk, immediately prior to completing the questionnaire. In the other stories L1 students returned responses with five identified issues at rates of 7% and 26%. Also the frequency of L1 responses with no issues identified was highest for story 2, though the no-issue frequency was of similar calibre to that of story 1.

The significance of distribution of issues raised by L1 students for story 2 needs to be viewed from the perspective of their having completed a class discussion on the issues prior to completing the questionnaire. The higher number of no-issue responses may be explained as student fatigue to discuss these issues again, in particular as they had just finished a class exam for the module. However, it seems difficult to find a logical explanation for the lack of five-issue responses in the same group, other than the L1 students not perceiving the moral issues in story 2, despite their exposure to the issues

in the preceding class discussion. The frequency of L3's listing all five issues supports the idea of the issues being recognisable to students.

A preliminary criterion was then applied to make sure that the L3 responses represented a rise in ethical considerations and not only in technical ones. One question was asked for all statements: 'Can you answer this question (only 9 statements out of 100 were not either questions or statements that could be transformed into questions by adding a question word at the beginning or changing the word order) by reference to scientific/technical/financial data alone?' If the answer was yes, the issue was classified as non-ethical. Certain words also classified the issue as ethical. These were: necessary, safe, risk, certainty, worth and benefit.

Story 3 describing research on plant viruses produced a lower rate of ethical responses. From all the issues raised, 56% for L3 and 43% for L1 were classified as ethical. The respective figures for story 2 concerning pharmaceutical milk were 64% for L3 and 70% for L1. Also, the distinction between ethical and non-ethical responses was harder to draw in story 3, as many students raised issues regarding cross-breeding and other forms of spreading, without specifically raising the issues of safety or risks.

Further effort to gain knowledge on the suitability of the stories was completed with a preliminary classification of responses into levels, according to their depth and sophistication. A three-tier system was hypothesised to exist, with the lowest level representing a brief and very general recognition of the issue, the second level detailing the issues concerned and a third level showing wider and more extensive appreciation of the problems involved. Responses to Story 3 (plant virus) were again harder to classify and in some of the thematic categories there were either no issues that could be classified in the lowest recognition or the full recognition tier. A table of sample responses in each category can be found in Appendix IX - Plant Virus - sample responses'. Story 2 (pharmaceutical milk) produced more varied data with responses in most thematic categories and in all levels of recognition.

All these results lead to a decision to concentrate further efforts into developing a scoring guide for story 2 (pharmaceutical milk), as it was found to produce the more varied data which had the quality of clear and uncomplicated scoring.

5.3.2.2 Scoring method

The preliminary scoring method of three tiers was chosen as the framework for the scoring protocol for the moral sensitivity responses. The 44 responses from L1 and L3 students were used as the basis for the scoring guide. All the responses were arranged according to their level of sophistication into four tiers: 1) non-ethical considerations, 2) bare recognition of the ethical issue, 2) some details and implications of the issues recognised, and 4) in-depth and detailed appreciation of the issue. The choice of classification of responses was discussed in detail with two departmental members (Prof. R Downie and Dr R. Downie) to form an agreement on the rating protocol and characteristics that signify responses on each level.

Each of the items was then submitted to pre-established tests of logic suggested by Bebeau *et al.* (1985): Is a criterion logically independent of every other (i.e. could an individual score high in one, but not the other?) This test was carried out in order to limit the number of criteria used for the scoring. The eight original categories recognised for the story were:

Costs and benefits	Ethics of mixing genes/cloning
Animal Welfare	Environmental impact of cross-breeding
Risks to humans	Public opinion
Scientific viability	Advancement of science

A logical dependence was found between 4 items. First 'environmental impact' and 'animal welfare' were found to be logically dependent as long as environmental impact was described as an issue of cross-breeding and escape of the genes into non-intended flocks. A second logical dependency was found between the categories of 'scientific viability', and 'cost and benefits'. Scientific viability concerned the benefits to CF patients which can be logically interpreted under the more general heading of costs and benefits.

For practical reasons, the categories of 'advancement of science' and 'ethics of mixing genes/cloning' were combined to cover all theoretical/basic values issues relating to the research proposal.

In the end the scoring guide consisted of four items:

1. Costs and benefits
2. Risks to humans and animals, including issues of animal welfare
3. Public opinion
4. Basic value issues

Each item was broken into three levels of recognition where each level had several sub-sections to give a scoring key for all the horizontal elements within each category. The scoring guide can be found in the Appendix X - Scoring Guide. For scoring purposes, each level of recognition was assigned a score from 1 to 3 so that the first general recognition scored one point, the second more advanced recognition two points, and the most advanced category response three points. Non-ethical issues raised accrued zero points.

The preliminary results from the 20 L1 students and 21 L3 students were scored using the scoring guide. The average score for the whole sample was 4.7 (L1 responses with zero issues raised were included in the scoring). The L3 average score was 6.0 and the L1 average score 3.4. When the responses not raising any issues, but indicating an opinion, were removed from the analysis, the L1 average was 4.8 and the overall average 5.4. The personal scores ranged in L3 from 11 to 1 and in L1 from 10 to 0, when the maximum score would be 15.

Out of all the scored responses, L3 students had 17% classified as level 3 items, 47% as level 2 items and 36% as level 1 items. The L1 students had a smaller number of scored responses, but they had a higher percentage belonging to the level three category (22%), while 44% were classified as level 2 and 34% as level 1.

The student responses on whether the research should go ahead varied greatly between L1 and L3.

L3 Opinion		L3 moral sensitivity average (range)		L1 Opinion		L1 moral sensitivity average (range)
L3 yes	29%	8 (3-12)		L1 yes	80%	4.6 (1-9)
L3 no	19%	5.8 (1-13)		L1 No	0%	
L3 I don't know	38%	7.5 (2-14)		L1 I don't know	20%	7.5 (2-13)
L3 empty	14%			L1 empty	0%	

Table 12: Sensitivity test pilot results - action choice

In other words, 80% of L1 students believed the research should go on, only 29% of L3 students were of same opinion. Again while none of the L1 students were opposed to the research, 19% of L3 students were. A sense of uncertainty was entertained by 20% of the L1 students, with the corresponding figure for L3 students being 38%. A further 14% of L3 students did not indicate their preference.

Even though within L1 and L3 the students did not seem to have an opinion on the acceptability of the research programme related to their moral sensitivity scores, the difference between L1 students and L3 students as a group was very notable.

5.3.3 Correlation of PPS2 and Moral Sensitivity score

The Perry moral development scale and the moral sensitivity score are logically independent elements of moral development. Where the Perry questionnaire measures development in the meta-understanding of moral issues and personal relationship with commitment and responsibility, moral sensitivity measures an ability to interpret a situation in a moral respect. Developmental advances in one area will not necessarily indicate developmental advances in the other. Similar independence was found by Bebeau and Brabeck (1987) between moral competence (DIT) and moral sensitivity (DEST).

The correlation between the scores for moral sensitivity and Perry for those students for whom moral sensitivity scores were calculated (responding to the story of

pharmaceutical milk) was found to be non-significant (p=0.1418, Pearson correlation coefficient), as can be seen in Figure 2.
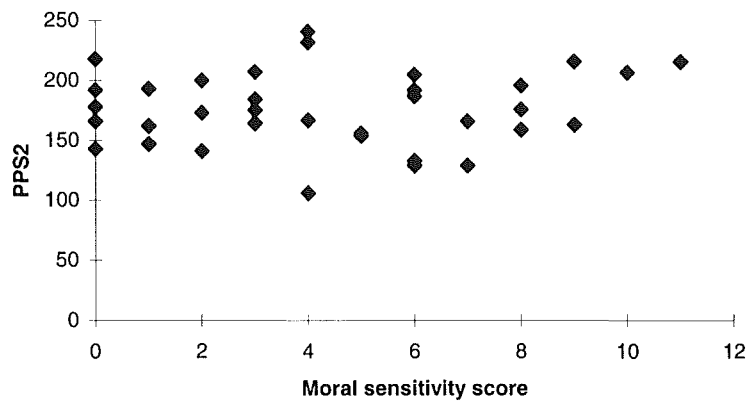
Figure 2: Moral sensitivity and PPS2 correlation

5.4 Course evaluation

At the end of each moral teaching programme the students were asked to give feedback on their reactions to the course. A standard feedback form was filled in by all students. The students were asked three open-ended questions following the example of Bebeau and Thoma (1996):

1. What is your view is the most important issue in X?

2. What was the big point you learnt from the discussion today?

3. What is the main unanswered question you leave the discussion with today?

The scoring of these responses is discussed in chapter 9.

5.5 Summary

A control group is essential for research on the effects moral education has on moral development. Without a control group, it is impossible to differentiate the impact of moral development that occurs naturally and moral development gained through ethics

teaching. The alternatives for control groups considered here were: a different cohort of science students at Glasgow University, science students in other universities, non-science students at Glasgow University, and a control group formed from students within the same cohort at one university. The last option was chosen to produce the most accurate results possible, despite the organisational difficulties.

A decision was made to include DIT as a measure of moral competence in the testing tool used for this research. Despite the theoretical misgivings of the tool, it provides a well tested and valid measure for certain types of moral development and the results would be comparable with other moral education courses, which is a true advantage. Also it will provide interesting studies of the results with comparison between moral sensitivity and moral competence and how they are promoted by ethics teaching.

Perry's theory of moral development offers an opportunity to test students' perception of the nature of moral problems and how these perceptions change during the academic years. An Osgood scale questionnaire was devised to capture this part of students' moral development. Ten statements representing each of the Perry types were sent to independent judges who were asked to indicate which Perry type they believed would agree and disagree with each statement. The A and C statements which had higher than 0.8 inter-judge agreement were collected for a Perry Questionnaire comprising 10 statements. This questionnaire was piloted with 70 L3 students and 60 L1 students in May 1999. The preliminary results indicated that students were already in L1 dominantly operating with Perry C type concepts when dealing with moral issues. The results were broken into five elements, out of which only one (Personal responsibility and relationship with multiplicity) showed significant differences between L1 and L3 students in the answering patterns. This difference could not be captured by using a Perry Index (PI), or by calculating Personal Perry Scores (PPS). PPS on the other hand, was believed to provide a useful tool for tracking personal developmental changes during the research project.

For measuring moral sensitivity, a decision was made to create a story which required students to give open ended responses, which would be analysed to give a moral sensitivity score. The first attempted story was abandoned after it was piloted with 18

L3 students, because it generated too may different types of responses to allow for a reliable scoring system to be developed. In the next stage three new stories were piloted with both L1 and L3 students, approximately 20 for each level and each story. These stories were first analysed according to the number of issues they raised and the frequencies students' raised each issue, and how many issues out of the maximum five the students raised. After this stage one story was abandoned after it produced a scattered response pattern. The other two were further analysed according to whether the responses were ethical or not. One of the stories came out of this stage with considerably easier scoring patterns and this story, revolving around genetically engineered pharmaceutical milk, was chosen as the story to be used in the moral assessment tool. This story was further analysed to create a scoring guide which allowed each of the issues to be scored according to the level of recognition.

Part III - Teaching Ethics

Can ethics be taught? The issue is an old one. Doubts of whether ethics can be taught often find resonance with those who support the behavioural aims of ethics education - teaching people to be 'good' and to do the 'right' thing. The question is also prominent when the aims of ethics education involve building character and adopting virtues. But the question is less prevalent if ethics education is approached from the direction of skills and abilities. If ethics consists of skills and abilities to recognise, analyse and solve moral problems there is no reason to believe such abilities and skills could not be taught equally as well as other types of skills and abilities. We do not doubt whether we can teach students the skills required in mathematics, analysing literature, or understanding biochemical reactions - similarly we should not doubt whether we can teach the skills of moral decision-making. Teaching skills of moral decision-making may require particular approaches in order to be effective, it may not be easy and quick, but these problems should not cloud the vision that it is possible to teach skills in moral decision-making - while it may be impossible to ensure that these skills are put to use outside the classroom.

The way I have described the aims of ethics education in Part I is compatible with the view that ethics is not out of reach of education. After the methods developed in Part II for measuring the success of ethics teaching, Part III is an investigation into *how* to teach ethics.

Even if teaching moral decision-making skills is perceived possible, it is not obvious what type of teaching would be most successful. This is an issue familiar to all disciplines. Successful ethics teaching is dependent on at least four elements: 1) the choice of timing - at what age/educational level ethics is introduced; 2) the duration of the ethics curriculum; 3) the choice of pedagogical approach; and 4) choice of topics. Part IV is divided into three chapters which cover all of these areas.

Moral development is not dependent on formal education alone and section 6.2 provides a short overview of the influence other life-experiences have on moral development. This is followed by a detailed look into teaching set-ups that have been found to accelerate moral development - what are the criteria for good ethics teaching in method, duration, and timing and what are the pre-requisites for successful moral development.

Chapter 7 concentrates on describing two teaching approaches which, in the light of the findings of the first chapter, incorporate the elements of successful ethics teaching programme. These are Problem-Based Learning (PBL) exercises and Structured Discussion Groups. The chapter starts off with consideration of what a student-centred teaching approach in ethics means for the teacher, who now needs to adopt the role of facilitator. This is followed by an analysis of moral decision-making. A framework model is developed in this chapter to provide a platform for incorporating different teaching approaches into ethics teaching. This is followed by a discussion of how PBL and Structured discussion groups can be used to promote the Moral Decision-Making (MDM) model.

Chapter 8 explains in detail how the research project described in this thesis has been structured and what are the links between the research set-up and the knowledge gained in the previous chapters. The latter part of the chapter describes the decision-making procedure applied to the choice of topics.

At the end of Part III, a clear picture should have emerged on the opportunities for teaching ethics and how those opportunities have been seized in this research. This is highlighted in a short summary.

6. Theories for teaching ethics

This chapter focuses on the elements that research has found to encourage moral development. As a preliminary to the research review, section 6.1, starts by developing further the thoughts already discussed in Part I on the aims of ethics education. In this section the emphasis is on a detailed analysis of the differences between teaching ethics as skills in comparison with teaching ethics as rules. In theory both approaches are compatible with the intrinsic aims described previously in Part II, but I hope to prove that the skills-based approach has significant advantages over the rules-based one. This is followed by section 6.2, which is an overview of research carried out to identify how life experiences promote moral development. This creates a back-drop for the educational efforts discussed later in this chapter. This is followed by a review of research findings regarding four essential elements of ethics education: section 6.3.1 concentrates on the pre-requisites for moral development, section 6.3.2 on what is the best timing for an ethics course, section 6.3.3 on the most suitable duration of ethics courses, and section 6.3.4 on the most successful methods of ethics teaching. How we view moral decision-making has a major influence on the skills we perceive as necessary for successful moral development and how we believe moral education can best encourage them. To answer these questions, a Model of Moral Decision-Making (MDM) is developed in section 6.4. This model provides a framework for discussing moral education practices.

6.1 Skills vs. rules

The aims of ethics education were set in Part I to be intrinsic. Intrinsic aims were described as aims to improve the students' ability to deal with moral problems for no particular and measurable future benefit. The teaching approach to achieve these aims could be both skill- or rule-based, though I hope to show here that the skill-based approach is superior to the rule-based one for several reasons.

The division into skill- and rule-based approaches to ethics education runs parallel with a traditional philosophical divide between those who claim that moral action should be governed by rules and those who believe it should be governed by skills in context. The rules view justifies particular norms and moral actions by latching them on to universal rules. The validity of these rules depends on the status of universal or transcendental moral principles justified by considerations from the point of view of all. The skills view, on the other hand, localises ethics by describing it as skills moral agents are able to exercise in everyday life. In the skill-based approach rigid rules are dispensed with and replaced by the competence of moral agents to analyse and resolve moral problems within their context.

The role of ethics teaching from the rule-based perspective is to teach children/young adults universal ethical norms and rules, which the students can then apply to the moral problems they encounter in life. I considered some of the problems with this approach in section 3.1 - the difficulties in choosing a set of principles acceptable to all, or even most; the difficulty of applying universal rules to complex real-life moral problems; and the inconsistent and insufficient links between teaching rules and changing actual behaviour. The rule based approach is further weakened by the fact that actual moral problems are poorly solved if the only consideration is whether a moral action satisfies a rule. The rule-based approach is contrasted by one genuinely progressing to master the difficult relationships between moral values (rules), and the intricacies of the situation, and to master a combination of habits, attitudes, and feelings. Quoting Løvlie (1997): "[Moral] Judgement is dependent upon sorting out the threads of the moral web, taking both social, psychological and moral facts into consideration … justifying action by appeal to principles does not add significantly to the solution of particular moral problems." (p. 410).

The skills approach to moral education emphasises the faculties of the students. The aim is to enhance students' abilities to recognise, analyse and solve moral problems, not by direct application of a set of moral rules and principles, but by careful independent reflection on the situation and consideration for those

involved. The skills view can be considered a more sophisticated guide to moral actions than rules, as it requires personal involvement, judgement made by the agent, rather than an impartial and unreflective application of rules. This is not to say that moral rules are redundant, for they are not. Dreyfus and Dreyfus (1986) describe rules important for novices in any particular field, an importance which decreases proportionately with an increased personal competence. Their example is from driving - where a novice driver performs according to explicit rules - how to handle the gears, when to look at the mirror, etc. - an expert driver makes immediate and often sub-conscious adaptations to traffic patterns. The same can be applied to moral situations. In novel and complex situations, we need to rely on some rules to start with, but as we become more accustomed to moral problems, we are both more able to make our own decisions and more intuitively follow adequate decision-making patterns.

The ethics education described here will follow the skill-based approach with recognition of the importance of moral principles and decision-making rules to students who are not yet experts in moral decision-making but rather taking their early steps as independent moral agents.

6.2 How life experiences affect moral development

Moral development is not dependent on formal ethics teaching. It can occur in the absence of formal moral education and the role of moral education is rather an acceleration and encouragement of moral development that is inherent in personal development *per se*. Therefore, before studying the specific elements in moral education that have been found to accelerate moral development I will briefly discuss other elements affecting moral development, which create the back-drop for moral education.

In a search for life experiences that correlate with development in moral judgement, formal education was found to be one of the strongest and most consistent correlates (Rest and Deemer, 1986). DIT has been the most commonly used method of tracking down the sources of moral development and the results

show a strong trend of increasing scores with increasing years of formal education. When an average DIT P-score for adults is 40, adults with only high school education score an average of 28.2, while practising medical physicians score 49.5 and those with doctoral degrees in philosophy or social sciences score 65.2 (Rest 1982). This evidence shows a trend of increasing P-scores with increasing years of formal education, but it does not provide answers to what years of formal education actually represent in terms of experiences, subject activity, or processes. In other words we need to analyse why formal education fosters moral development.

Rest and Deemer (1986) entertain five alternative explanations:

1. The environment in higher education socialises certain attitudes and forms of verbalisation (This presupposes that DIT measures, at least partially, attitudes and their verbal expression).

2. Some particular skill or piece of knowledge is taught in higher education and the upward trend of moral reasoning scores reflects more students learning these skills over time. (This explanation assumes that DIT scores reflect some particular skill or piece of special knowledge.)

3. Higher education indirectly imbues students with a general perspective or world view. (This presupposes that the DIT scores reflect some particular world view dominant in higher education.)

4. Higher education provides a generally stimulating environment in which individuals work out their own ideas about morality. This view does not presume that the higher education environment is promoting any particular ideology (as is assumed in the alternatives 1 and 3), but rather promotes reflection and self-discovery. (This alternative assumes that self-constructed development in moral judgement inevitably leads to more principled thinking i.e. higher P-scores.)

5. It is not the higher education as such that makes the difference, but rather correlation between moral reasoning scores and formal education indicates something about the people who seek extended education. Perhaps, according to this alternative, the people who choose higher education are those who are predisposed to be more reflective, who

seek intellectual stimulation, and who are self-motivated to develop cognitively.

Logically any of these explanations (or a combinations of them) could account for the empirical findings that development of moral judgement is related to years in formal education. The critical experience that fosters growth could be talks at tea breaks, course readings, extracurricular activities, life style of students, the formal ethics teaching, or the reflective personalities of higher education students.

Research in the 1970's (reviewed by Rest and Deemer, 1986) tried to locate which element of the higher education experience explains the correlation with years in formal education and the upward curve of moral reasoning scores. One type of study was to link certain activities, interests and life styles with moral judgement. In general these studies indicate that subjects with higher moral judgement scores tended to be better read, more knowledgeable and active both academically and socially. Nevertheless, the power of the trends using these experience-measures was not much higher than the correlations with the more simple variable, years of formal education. Another type of study was carried out to find if there was a linkage between certain types of personal experiences (e.g. 'attending a course which presented material from different perspectives', 'experiencing brutality or suffering' or ' experiencing a significant decision involving a family member') and moral reasoning scores. The results yielded inconsistent results or low levels of significance. Yet another type of research was direct interviews asking the subjects what, in their own view, had influenced their development in moral thinking. The responses were identified with categories like 'new real world experiences', 'formal instruction, reading and study', or 'involvement in community affairs'. Again the results were inconsistent and weak in explanatory power.

After this type of research failed to pinpoint any particular experience in formal education that correlated with increasing moral reasoning scores, another approach was constructed by Spickelmier at the University of Minnesota in collaboration with Rest (Rest and Deemer, 1986; Spickelmier's original work is

an unpublished manuscript). It was a set of structured interviews that included questions on a wide range of experiences that were hypothesised to influence development. The topics of the interviews included the following: living situation; peer relations; relationships with faculty; academic involvement; learning preferences; extracurricular activities; exercising responsibility; relationships with the opposite sex, family; occupational goals; personal goals and values; experiences of trauma or crises; and post-college experiences. Questions were of the following type: 1) 'Describe a typical evening in each of your college living situations?', 'What did you or your roommates tend to be doing?', 'What did you talk about?'; 2) Give me the first names of some of your friends in college. Describe them to me. Compare your friends with other students in your school; 3) Can you tell me about a book, a lecture, a particular intellectual problem that impressed you during college?'; and 4) How dependent were you on your parents for financial support and how did that change over the college years?'. A coding guide was then developed for a dozen dimensions. The three academic ones, which also were most predictive of moral judgement development, were:

1. Socialising Environment. This code characterised how fully the student was immersed and involved in the college milieu versus simply attending college and being influenced strongly by non-college socialising forces.

2. Academic Success. This code was to differentiate those who had formulated a clear academic goal, did well in academic work, and made normal progress toward their academic goals, in contrast to those students who did not have clear goals and were frustrated in achieving them.

3. Educational Orientation. This code represents the degree to which the students worked hard at their studies, enjoyed academic life and the world of ideas and activities of reading and discussing, and chose friends who where similarly serious students.

Spickelmier's study was small and non-random (students from one institution only), so the results do not carry a strong generalising power, but they are still

interesting. First, moral judgement scores seemed to correlate with life style, i.e. those who had high DIT scores in the first year tended to be higher also in the academic variables. Second, certain life experiences correlated with the moral judgement scores: in other words, students who scored higher on these life experience variables showed greater gains in moral judgement scores between the test scores of first year and two years after college. Spickelmier's results support explanations 4 and 5 above. This indicates that moral development does not depend on some skills or attitudes fostered in higher education, but rather that moral development is more dependent on personal orientation towards learning and development in general. If this is true, the correlation with higher education may be only contingent in the sense that higher education attracts people more likely to make moral development, not that higher education itself creates the developmental advantages. On the other hand, it can be viewed that higher education encourages the inherent abilities for moral development and only those with the opportunity/desire to participate in higher education receive these benefits. Deborah Deemer's research follows this line of thought.

Deemer (Rest and Deemer, 1986) picked on the questions raised by Spickelmier's research. Deemer conducted a study involving over 100 subjects tested over 10 years, from a diversity of backgrounds. The subjects were first tested at high school in the early 1970s with DIT and then again ten years later with another DIT test and a structured interview. The sample consisted of subjects of various educational levels, family situations and located in both cities, countryside and semi-urban areas. The interview questions were adapted from those used by Spickelmier to accommodate non-college subjects. Deemer and Rest concluded from these interviews that moral judgement appears to be a by-product of general social development, not a special result of particular moral courses, moral crises, or moral leaders. The most important life experiences that correlated positively with moral judgement scores were:

I.   **High School codes:**

        A.   <u>Academic Orientation.</u> Subjects that studied hard, made good grades, enjoyed reading, and took challenging courses where scored 'high' on this code.

      B.   <u>Further Education Encouragement.</u> A high code represents encouragement by parents and/or school personnel to continue with higher education.

II.   **Interim Period:**

      A.   <u>Educational/Career Orientation.</u> A high score is given to students who work hard at their studies, enjoy learning and engaging in ideas, work toward academic goals, and choose friends who are serious about their studies.

      B.   <u>Continued Intellectual Stimulation.</u> This is a code for characterising the extent of intellectual stimulation provided by one's environment over the ten-year period of the study. High scores are given to those who hold a secure and challenging job or whose community involvement is rich, friends or spouse are enlivening or who have made family life particularly stimulating.

III.   **Young Adulthood codes:**

      A.   <u>Career Fulfilment.</u> A high score is given to subjects who find their jobs challenging and meaningful, who are committed to their careers, and have chosen their own career path.

      B.   <u>Civic Responsibility.</u> This code measures the extent to which a subject is concerned about the community and actively contributing to the welfare of the community. High scores reflect leadership roles, intermediate codes of acting on request and low scores an opposition to civic participation.

      C.   <u>Political Awareness.</u> This code reflects the extent a subject is interested, informed and articulate about macro issues in society. High scores are given to those who read several sources of information and participate frequently in political discussions, medium scores to those who watch the news on TV, read some newspapers and participate occasionally in political discussions, and low scores to those who seldom read about political issues and take little interest in them.

In sum, the moral judgement scores of young adults correspond with career or education encouragement, interest in studies and career, involvement in community life and political awareness.

An important characteristic of Deemer's study is that some of the codes emphasise personal characteristics and some environmental ones. All codes presuppose a complex, reciprocal interaction between the individual and the environment. It is hypothesised that people, to a greater or lesser degree, help to make their own environments and often self-select themselves into situations of challenge and opportunity, while at the same time in order to develop, people need certain environmental support and advantages. Development proceeds most when the person seeks development and the environment supports and fosters it. No unidirectional or simple causal relationship between personal characteristics or environment is presupposed; events do not themselves cause development and personal characteristics do not cause certain situations. Rather, there are mutual influences of individual characteristics interacting with the environment - situational opportunities and difficulties.

The aim of my research relates to these issues directly. Formal education provides a good base for moral development which is shown by research to produce significant advances in moral cognitive skills. Regardless of how we view this advantage to come about - as primarily from the greater potential of students proceeding with higher education or as higher education itself creating an atmosphere of learning and encouragement - ethics teaching can have a role as a further catalyst of moral development. It can provide an opportunity and source of encouragement for interaction with moral issues and thoughts of others which could spark intellectual interest, involvement and skills to go on and develop one's moral agency independent of formal teaching. This research is designed to find out to what extent certain forms of moral education can provide this opportunity for accelerated moral development, a stepping stone on which to accumulate further development.

6.3 How education encourages moral development

It would be very bold indeed to assume that all educational approaches relating to ethics would actually encourage moral development. Worse still, it is logically possible to have ethics education that discourages, or even regresses, moral development. Ethics programmes with negative effects would not promote students' own decision-making skills or recognition of moral aspects, but would reduce their freedom in making decisions in a manner not dissimilar to indoctrination.

For an ethics education programme to be successful, it has to get four elements right. First, it is necessary to design the course to meet students at the general developmental stage they are at. A programme is unlikely to be successful if it either grossly over- or under-estimates the students' logical, verbal or intellectual abilities. These issues will be discussed in section 6.3.1. Second, people are more open to moral development at certain stages in life, and moral education should try to utilise the already existing eagerness and readiness by offering the educational programme for people most likely to take an interest in it. Previous research has also shown that moral development does continue into adult years (Rest, 1982), which increases the opportunities of moral education. These issues will be discussed in section 6.3.2. Third, the length of a moral education programme has been found to influence the results the programme has. The right length is a balance between too little and too much. These parameters are discussed in section 6.3.3. Finally the teaching methods used in moral education have a direct impact on moral development gains. The key element that emerges from reviews of previous educational programmes is the importance of nurturing students' personal moral agency. Successful pedagogical programmes are reviewed in section 6.3.4.

6.3.1 Pre-requisites for moral development

Moral development is not an independent variable, but associated with other forms of personal development. There are thresholds in development which

require one to have passed a certain level in one area before further development can be gained in another. For moral development these threshold skills are cognitive. Without basic cognitive skills one is not able to recognise, analyse and judge a moral situation. Therefore cognitive development facilitates moral development, but is not sufficient for it. Alternatively, absence of basic cognitive skills can explain the limits of moral reasoning.

One of the most basic pre-requisites for moral development is logical reasoning. According to Piaget (Inhelder and Piaget, 1958) there are three major developmental stages of logical reasoning: the intuitive, the concrete operational, and formal operational. At the level of concrete operational, entered around the age of 7, children can make logical inferences, classify things, and handle quantitative relations about concrete things. Most, but not all, will enter the formal operational stage in adolescence, at which they can reason abstractly - consider all possibilities and relations between elements in a system, form hypotheses, deduce implications from the hypotheses and test them against reality. Since moral reasoning clearly is reasoning, advanced moral reasoning depends upon advanced logical reasoning. One cannot, for example, reach post-conventional stages of Kohlbergian moral development before one has reached the stage of formal operational logic.

In addition to logical development, one needs other personal abilities to be able to recognise, analyse and judge moral problems. The pre-requisite social ability is role-taking. It can be described as a stage-like development parallel to moral development, which progresses in stages of social perception or social perspective. Corresponding to the three levels of moral judgement, Kohlberg (1976) postulates three levels of social perspective as follows:
1. Concrete individual perspective (corresponding with Kohlberg's pre-conventional stage)
2. Member-of-society perspective (corresponding with Kohlberg's conventional stage)
3. Prior-to-society perspective (corresponding with Kohlberg's post-conventional stage).

These role-taking stages describe the level at which the person sees other people, interprets their thoughts and feelings, and sees their role and place in society. These stages are very closely related to moral stages, but are more general, since they do not deal with just moral issues, but with all social interactions. Also, to make a moral judgement at a certain level is more difficult than simply seeing the world at that level. But seeing the problem is necessary before any decisions can be made and therefore one needs to reach the corresponding level of social perspective before one can reach the level of moral development. The social perspective stages are therefore necessary conditions for reaching a certain moral stage, but not sufficient.

Apart from the cognitive pre-requisites for moral development, emotive elements are also crucial. For students to benefit from education, they need a desire to learn, motivation to develop, and self-discipline to work towards the developmental goals. This motivational pre-requisite can be partly induced, or at least encouraged, by the teacher or peers (Entwistle, 1998), but if a student is adamant about the decision not to learn, the efforts of others are easily wasted.

6.3.2 Timing

Because moral development is not an isolated strand of development, but deeply associated and dependent on other areas of development, it is important for successful moral education to be given in a form that corresponds with the abilities of the student and during periods when moral development most naturally occurs. Ethics education that disregards the natural conceptual frameworks of students and assumes either too much or too little cognitive and social capacity is likely to seem artificial and irrelevant to the students. Learning does not take root from such teaching efforts, the students development is not enhanced and in some situations it can even regress.

Moral development tests invariably show an age trend when measured with DIT. Moral development is not restricted to childhood or adolescence, but continues

well into adulthood. Older subjects show higher stage responses when measured against Kohlberg's stage theory. A meta-analysis of 55 studies by Schlaefli *et al.* (1985) showed that older, and presumably more advanced, subjects show higher scores than younger, presumably less advanced subjects. Longitudinal studies over 10 years that followed the same subjects and re-tested them also showed increasingly higher scores. The average DIT P-score for junior high school students was 21.9, for senior high students 31.8 and average college students 42.3.

All Perry's (1999) initial work on ethical development was done on university students and the scheme thus describes development in late adolescence and young adulthood. For Perry's scheme there are less accurate and widely tested data on the developmental stages of university students, but according to the interview data students arriving at university are predominantly at Perry types A and B, while they developed during their university years into being Bs and Cs. The same developmental trend along the Perry scheme has been identified by Gray (1997) and Katung *et al.* (1999).

The evidence for moral development continuing in adulthood provides strong support to the claim that it is not too late to teach ethics in higher education. Students' moral skills are still developing and they are cognitively and socially apt to develop morally, possibly even more so than they were a few years earlier in high school. This claim is further supported by research findings on the gains of moral education programmes, where some of the largest gains have been achieved with participants in their 20s and 30s rather than with participants of school age (Rest, 1988). Meta-analyses by Thoma (1984) compared the effects of ethical education programmes in four different age groups: 13-14 year olds (junior high school), 15-16 year olds (senior high school), 18-23 year olds (university), and over 24 year olds - and found that educational efforts were most successful with the over 24 group, followed by university students and high school students, in that order. Thoma also pointed out that the difference might be partly explained not by more adaptable age, but because

1. most of the adult groups were exposed to Kohlberg's theory during the ethics teaching, or

    2. most adult subjects were volunteers, which might suggest a higher

       motivation to participate in the programme.

But Thoma also suggested that the difference might be due to adults being able to draw from a wider spectrum of experiences and thus attach greater personal meaning to topics which would lead to more accelerated moral development through role-taking.

Knowledge of the moral, cognitive and social developmental stage of students is also an important element in successful course design. Individuals have a tendency to accept moral reasoning of one level higher than the one they are on themselves and reject the moral reasoning of the levels below their own. This has lead to the theory that exposure to higher level moral reasoning will encourage moral development (Kohlberg, 1976). But, importantly, students can only perceive and be attracted to moral reasoning immediately above their own level. Exposure to levels higher than immediately above one's own is not comprehensible to the student and thus does not encourage development, but only frustration or boredom (Finster, 1991).

Moral education may also be hindered if it is introduced at a time when students are involved with other major developmental projects. For example, it may not be most beneficial to introduce ethics at the start of first year in university, when students are going through the major transition from pupil to student, with new freedom and responsibilities over personal learning, from dependent child in a family to an independent young adult, and socially gaining new ground with new friends and a new environment. Developmental capacity is limited and if it is taken up by other developmental processes, moral development may not be encouraged by ethics education.

6.3.3 Duration

The length of the moral education programme and the number of contact/independent study hours have a direct influence on the gains of the

programme. In their meta-analysis, Schlaefli *et al.* (1985) divided the studies using DIT as a measure of moral development into three groups according to their duration: short duration of 0-3 weeks, medium duration of 4-12 weeks, and long duration of 13-28 weeks or very intensive courses of 8 weeks and longer. Programmes of short duration show the smallest effect, long term ones the second largest and medium duration programmes the largest effect. Thus there was no linear relationship between effect and duration of the educational programme - increased duration does not guarantee increased benefits in moral development.

Three possible explanations can be given for the reduced impact of longer duration ethical education programmes (Schlaefli *et al.*, 1985):

1. The longer lasting educational programmes have sacrificed quality for quantity and thus produce weaker impact per contact hour.

2. Students grow tired of moral education past 12 weeks and the impact of intervention falls off - this could possibly be counter-balanced by changing the educational approach dramatically after 10-12 weeks.

3. Artificial stimulation of moral decision-making is effective only for a limited period after which students need time to rest and consolidate the moral development achieved thus far.

6.3.4 Methods

All moral education programmes have unique features that depend on the circumstances, chosen aims, pedagogical approach, teacher characteristics, and how students interact with the each other, the teacher and the material. In their meta-analysis of moral education programmes Schlaefli *et al.* (1985) found three major core elements that capture the different types of moral education programmes:

1. Dilemma-based programmes that emphasise peer discussion of controversial moral problems where the teacher has a role as facilitator. This approach can provide concentrated practice in moral problem solving stimulated by interaction with peers - challenging one anothers'

thinking, re-examining assumptions, learning about different points of view and responding to counter-arguments.

2. Personality development programmes emphasise personal psychological development in general with an understanding that moral development is a major strand within it. Activities, which often involve direct interaction with others, have an objective of promoting reflection about the self and self in relation to others. Students can also be given formal instruction on developmental theories.

3. Academic courses emphasise the learning of the basic tenets of the academic discipline that has ethical components, which are discussed during the course.

In their analysis Schlaefli *et al.* found the dilemma-centred approach to produce the highest effect. Personality development programmes also produce significant moral development, while academic programmes have very small effect on moral development. The basic course design in this research is identifiable with the dilemma discussion programme approach described above.

The dilemma discussion approach to moral education has three important elements:

1. The students are active participants in the discussion, they act as moral agents involved in the case as opposed to being passive recipients of information given by a teacher. In other words, dilemma discussions involve student-centred learning. The personal involvement allows students to reflect on their own views and study the view points of others participating in the discussion. The experience of moral agency can be enhanced by encouraging role-taking - actively adopting a view point other than one's own. Kohlberg (1976) emphasises role taking as one of the core elements in moral development.

2. In group discussions students can be exposed to (and the facilitator of the discussion should make sure they are!) moral reasoning more advanced than their own. This provides the students with an alternative model of moral decision-making and according to Kohlberg and Rest,

people are most drawn to decision-making that is immediately above their own level of ability.

3. Students face a moral problem that does not have a simple straightforward answer and they are thus forced to consider several elements of the problem, evaluate the importance of different shareholders and often face cognitive dissonance, where simple methods, possibly adequate in other situations, no longer apply. Dilemma discussions can provide an opportunity to consider viewpoints and roles not otherwise accessible to students: the issues discussed can be such that students would not otherwise take the time to think about them and the information associated with the dilemma can be new and exciting which encourages students to get more involved in thinking of the subject in their own time and make adjustment to how they view the world. When a new experience cannot be assimilated into existing categories of experience or when expectations are violated, people attempt to revise their categories and expectations so that experience will once again make sense and be predictable. The experience of cognitive disequilibrium can therefore provide a motivation for acceptance of new moral ideas (Rholes *et al.*, 1982). Experiences of cognitive conflict can occur either through exposure to decision-making situations that arouse internal contradictions in one's moral reasoning structure, or through exposure to the moral reasoning of significant others which is discrepant in content or structure from one's own reasoning (Kohlberg 1976). Such re-organisations of the mental map involve considerable psychic energy in the labours of unlearning and relearning. The inner urge seems to be a compound of many motives: 1) sheer curiosity, 2) striving for competence in understanding the relations of the environment, 3) an urge to make order out of incongruities, dissonance and anomalies of experience; 4) a wish for authenticity in personal relationships; and 5) a wish to develop and affirm an identity. The counter-forces appear to consist of such tendencies as the wish to retain earlier satisfactions or securities, to wish to maintain community in family or hometown

values and ways of thinking, the reluctance to admit one has been in error, the doubt of one's competence to take on new uncertainties and responsibilities, and most importantly, the wish to maintain a self one has felt oneself to be. Thus, while there is an urge to develop and change, there is also an urge to conserve (Perry, 1999).

6.3.5 Teaching skills

Based on the research findings above, an ideal setting for ethics education would be the following: the course would last between 4-12 weeks with weekly meetings which would be student-led and problem-centred. The facilitator would be an expert in ethical matters as well as highly skilled in leading a discussion without controlling it. The students would have high motivation to participate in the discussion, they would have the preliminary logical and social abilities to develop morally, and they would have an interest in the moral matters. This ideal setting does not specify the actual pedagogical method used during the group sessions. Several alternatives prevail, while any pedagogical method should appreciate the three levels of moral decision-making described in section 6.4.

One alternative is to adopt a direct approach to moral education. This means that the methods of teaching involve direct and transparent teaching of moral decision-making methods as opposed to dilemma discussions alone where skills are nurtured without direct attention to their development. The direct approach can include dilemma discussions, but it would have a strong additional component in direct skill training. Wilson (1990) describes one form of direct ethics education, while Penn (1990) provides some research results to support the hypothesis that direct teaching of moral decision-making methods is beneficial to students' moral development.

Penn builds his approach to ethics education on the premise that students can best develop skills in principled moral reasoning if those skills are directly modelled and applied to specific moral issues. Penn (1990) draws an analogy between learning mathematics and ethics:

"Just as students are not likely to develop skills in higher mathematical or scientific reasoning without direct teaching and modelling, it is unlikely that students will develop skills in higher moral reasoning without direct teaching and modelling (p.126)."

Penn's teaching method is unique in two features:
1. It targets directly and in tandem the fundamental moral reasoning skills (logic, role-taking, and concepts of justice).
2. Students are given direct instruction and modelling of higher level cognitive skills including tuition on the stage typology of moral development and classic philosophical texts.

Penn's statement on direct instruction should not be taken as being against student-centred problem-solving. He rather argues that students need not invent everything themselves. Students personal thinking can greatly benefit from didactically taught basic logical and philosophical elements before embarking on discussions of moral controversies (Rest and Narváez, 1994). This view is at odds with most moral education courses that consist of discussions around moral problems without any formal tuition or attention to the structured skills of moral decision-making.

Penn (1990) supports his approach by results. He divided students into five groups - one group being taught all four elements (below) and the other four lacking one or more of these elements, so that all groups received teaching at the very least in element four:
1. Study of formal logic - an ability to distinguish premises and conclusions of an argument, to understand the distinction between truth and validity of an argument, to uncover premises of an argument, to test hypotheses systematically, and differentiate between valid and invalid forms of argument.
2. Developmental theory and stage typology according to the moral development theory of Kohlberg.

3. Philosophical methods of ethical analysis and their application to controversial social issues - reading included Socrates, Aquinas, Declaration of Independence (US), Rawls and Martin Luther King Jr.

4. Application of generalised methods or theories to the analysis of social issues.

Assessment of all five groups was carried out by using the DIT. The largest P-score change between pre- and post-tests occurred in the group that had received all four elements of direct teaching (mean P-score change 15.16), the second highest score was with the group receiving no formal logic teaching, but all other elements (mean P-score change 12.74 for post-graduates and 11.91 for undergraduates), the third highest score was in the group that received teaching in developmental theory (2) and application of methods to social issues (4) (mean P-score change 3.88) and the lowest score changes occurred with the group that had only received teaching on application of generalised methods to social issues (mean P-score change 3.02). The P-score gain of 11-15 points is a significant change, which normally represents four to six years of formal education.

The results showed two consistent trends:

1. Students exposed to philosophical methods of ethical analysis (3) scored significantly higher in DIT than students not exposed to these methods, while inclusion of logic had a lesser positive effect.

2. Exposure to moral stage typology had a consistent, but moderate effect on DIT scores.

The latter of the trends in more controversial. Moral development theorists have not been able to agree why the exposure to Kohlberg's stage typology increased DIT scores. This debate started well before Penn's research and Rest (1986b) suggests two possible explanations:

1. Reading the stage descriptions in effect instructs the subject how to perform in DIT and thus produces 'contaminated' results that no longer represent true moral development stages.

2. Exposure to theory is a powerful educational tool for actually changing moral thinking.

Penn (1990) asserts that exposure to Kohlbergian stage typology effects the subjects' thinking profoundly, rather than equipping students with the skills to fake DIT. He gives four reasons:

1. To believe that learning about Kohlbergian typology leads to contaminated results, requires one also to believe that students seek linguistic items in DIT that sound Stage Five or Six rather than choose items that correspond to their own point of view. This goes against general experience that students rather hold onto their own moral opinions.

2. The great effect-size in Penn's research group exposed to all four elements of moral reasoning, may not have been due to exposure to Kohlberg's theory alone, but to a combination of elements.

3. To assume that students learn to fake DIT after being exposed to Kohlbergian stage typology underestimates greatly the difficulty undergraduates have in understanding the Kohlbergian stages.

4. If teaching Kohlbergian stages is focused on teaching mnemonic devices to identify stage structures, rather than on the analysis of the formal organising structure, then it is possible to be wary of contamination. But if the emphasis is on the latter, it should not be surprising that students exposed to this teaching show significant increases in the sophistication of their moral reasoning skills.

The introduction of a formal skills approach to ethics education within this project will be further discussed in chapter 8.

6.4 Moral Decision-Making Model

The way in which we perceive moral decision-making has influenced ethics teaching directly. If we believe that moral decision-making is mainly learning a logical problem-solving syllogism which in then applied to each case, we would

teach the structure and usage of that syllogism and how to find the necessary facts to apply the syllogism to. On the other hand, if we believed in rules to guide us through moral problems, we would teach a list of rules. Therefore before proceeding to study different methods of teaching ethics, I will discuss a moral decision-making framework I have developed. This framework is not intended to provide answers to detailed questions on the process of moral decision-making, but to provide a general framework on which to build an approach to teaching ethics.

This model is also an approach to moral decision-making that will allow description without identification with a particular ethical theory or set of moral values. This is not a value-free model, just as no model is, but the values it is based on - rational thinking and freedom of moral thought - can be associated with several ethical theories. They are structural values rather than values of content. One will find content elements in this model that resemble utilitarianism (the importance of contextual and consequential aspects), virtue theory (appreciation of personal/character ideals), or deontological theories (the recognition and appreciation of values). These elements are essential parts of this model, but none of them is presented as an overarching moral approach which would sub-ordinate the others. The moral decision-making model is applicable to several types of ethical theory and moral values, though a purely deontological approach is inappropriate because it excludes contextual elements as valid factors in moral decision-making.

Moral decision-making, as I discussed in Part I, consists of four components - 1) recognition and analysis of the moral facts, 2) cognitive solution of the moral problem, 3) priority given to moral issues, and 4) personal abilities to implement the decision. Each of these components forms an essential part of moral decision-making. The first two components describe the decision-making process and the latter two its implementation. The moral decision-making model I have developed describes the relationships between the first two components.

The decision-making components can also be described in terms of the process as three distinct elements: 1) ability to recognise and appreciate values, 2) ability to analyse and appropriately classify case particulars, and 3) ability to combine basic values and case particulars into a solution. These form three elements of moral decision-making:

1. Recognition and appreciation of basic values

2. Analysis of case particulars

3. Finding a solution that appreciates both case particulars and basic values.

The process of moral decision-making works as an interactive network - the analysis of case particulars and the solution to the problem are both underpinned onto the basic personal values. The basic values influence, often unconsciously, the analysis of the case and the solution that is evaluated against them. If the solution contradicts the basic values, there are two options available: 1) to re-adjust the importance of the basic values to conform with the solution, or 2) work through steps 2 and 3 again to find an alternative solution.

The basic network of moral decision-making is therefore:
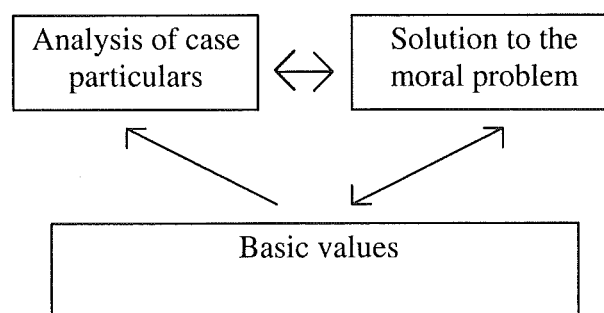
Figure 3: Moral Decision-Making Model

Moral decision-making is often carried out without incorporating all the elements; one can either move directly from basic values to the solutions (rule-based approach) or one can reduce the basic values to their bare minimum and work exclusively from case analysis to solution (this method is often identified with

utilitarian decision-making where the only basic value acknowledged is the maximisation of benefit). I strongly believe that moral decision-making is impaired if one of these elements or any of the dynamic links are ignored. Ethics teaching should thus work to create skills required in all parts of the decision-making network.

Moral decision-making is not an isolated form of decision-making, and therefore I wish to emphasise that moral decision-making needs to incorporate not only the specific elements of moral decision-making described above, but also to have a bed rock in basic decision-making rules. From Wilson (1990), they are:

1. Adherence to laws of logic.
2. Correct use of language.
3. Attendance to facts.

I will describe each of the moral decision-making elements in more detail and illustrate the process with the following moral problem:

The Draize Test

The Draize eye test is used in cosmetics research to test the eye irritancy of new products - cosmetics, shampoos and so on. The substance to be tested is injected into the eyes of rabbits - 0.1. mg is injected into the conjuctival sac of one eye in each of six rabbits with the other eye serving as a control. The lids are held together for one second and then the animal is released. The eyes are examined at 24, 48, and 72 hours to see if there is corneal damage. No anaesthetics are used and the eyes are not washed. The large doses often result in permanent eye damage, but are used to provide a large margin of safety in extrapolating for human response. Should companies test their new products in this way or not?

6.4.1 Element 1 - basic values

Recognition of basic values is a fundamental project, because basic values are the core of a personal view of life. It is an analysis of what is important to oneself and why. The process of recognising basic values is not a one-off procedure, which, once completed, would provide a solid and static basis for future moral decision-making. On the contrary, the recognition of basic values, and learning to

appreciate their effect, may evolve during every time moral decisions need to be made. I am prepared to state an even stronger claim; a static set of moral values indicates ceased moral development which is at risk of degenerating to more unsophisticated moral decision-making methods. So I cannot over-emphasise the importance of keeping the interpretation of basic values open to re-evaluation.

In solving moral problems or choosing between different alternatives, one needs to consider which sub-set of basic personal values is meaningful to the question in hand. Not all basic values are equally important in all decision-making situations, and the decision-making process is simplified if one is able identify which values are significant in each case. Also in most cases it is essential to be able to recognise which values are more fundamental than others, as all moral dilemmas, by definition, involve a conflict of two or more core values.

At any point in time one can perform an inner search of basic values and attempt to rank them in order of importance - 'if I need to sacrifice one value, which one will I sacrifice first, second, third etc.'. The evaluation carried out without pressures of making a moral choice can produce an 'ideal' set of values. One can imagine, for the sake of an exercise only, that there are fundamentally right and wrong answers, some values which are always superior to others, and imagine a hierarchy that causes no problems. Appreciation and construction of imaginary value hierarchy will be beneficial in learning about one's own value system and providing a base-line to evaluate solutions against. Naturally it is important not to become a prisoner of one's own value appreciation, because the world is not ideal and rigid systems allow for a very limited ability to provide acceptable solutions to real moral problems.

Using the Draize-test as an example, I will outline one possible approach to working through the moral decision-making model. Over the next several sections, text in italics is a possible working through of the moral decision-making process using the Draize-test as an example. The views expressed are not intended to represent a particular person's actual views: rather to give one possible and reasonable response to the case.

Without any particular order, some of my basic values are:

- *non-maleficence (not harming others)*
- *protecting the environment*
- *equality*
- *protecting the innocent and weak*
- *protecting my personal autonomy*
- *having an opportunity to do what I enjoy (personal autonomy)*
- *safety (financial, mental and physical)*

In a case of conflict, I believe that the last thing I would be ready to compromise is personal autonomy. I believe very strongly that everyone should be allowed to do what they want, make their own decisions without anyone else being allowed to interfere. After autonomy the next most important value is protecting the environment, because if we do not protect nature, everything else will be pointless. The next most important value is safety, because I believe that living in fear would make life intolerable. Then is non-maleficence - I cannot stand the idea of harming someone. These are followed by protecting the innocent and weak, having an opportunity to enjoy what I do, and equality. So the hierarchical order of my basic values is:

1. *Personal autonomy*
2. *protecting the environment*
3. *safety*
4. *non-maleficence*
5. *protecting the innocent and weak*
6. *having an opportunity to do what I enjoy*
7. *equality*

In forming an opinion about the Draize-test I envisage that the most influential values are non-maleficence, safety and protecting the innocent and weak (underlined above).

6.4.2 Element 2 - analysing the case

Before any actual moral problem can be solved, it must first be analysed. The analysis includes a recognition and separation of facts from values, placing the problem in a wider context, and evaluating facts and values according to their importance in the particular case.

The key to any moral decision-making is the recognition of moral facts, for without recognition and proper separation of values and facts one cannot perceive moral problems, and without perception there cannot be solutions. The initial recognition of moral aspects in a situation is followed by placing both the 'hard' facts and moral facts within the wider context and evaluating their importance to the particular case. It is not enough to be able to recognise and analyse the values embedded in a particular situation alone, but one needs to become aware of the context within which the case exists. Moral philosophers may work with abstract problems and even argue that solutions to such problems are universal, but I believe such detachment to be a detrimental approach to applied ethics problems. Inability to appreciate the social context from which the moral problem arises, is equal to omitting contextual facts in a legal deliberation. Without contextual analysis, we are not able to evaluate which facts and values should take precedence in case of conflict. It is important also to realise that one will never be able to consider all the facts and that one's observations are inevitably influenced by personally held core values.

Awareness of personal values can assist decision-makers in reducing the influence of personal point of view in a case analysis, but that influence can never be completely removed. An appreciation of the impossibility of being able to consider all the facts guides decision-makers in evaluating their decisions with an appropriate lack of supremacy.

The contextual analysis moves the moral decision-maker away from the basic set of values of element 1 into an appreciation of the factual and moral context in order to be able to move onto making a decision with the best possible tools to

provide the most appropriate and acceptable solution to the moral problem in question.

> *At first sight the Draize-test constitutes the following factors:*
>
> - *the test set-up (time, animals used, testing on eyes, dosage)*
> - *pain experienced by the animals*
> - *damage to the animals*
> - *the type of products tested*
> - *safety of human users*
> - *economical/financial situation of the cosmetic industry*

> *Most of these factors contain both moral and factual components. The test set-up as a description of the Draize-test is purely factual. Pain experienced by animals is a scientific observation, though not an easy one, which can be also viewed as a moral problem - is it acceptable to inflict pain on animals? The type of products tested can also be perceived as a factual list of the products actually tested with the Draize-test. But the type of products can also be given a moral evaluation - does it make a difference if the test is carried out for cosmetics which are not 'necessary'? And how do we define what is 'necessary'? Safety for human users is factual in a sense, that we can provide data on the risk factors or the type of risks involved, but the actual risk assessment and judgement based on it all requires a value decision - what size of risks are we ready to take? What are we ready to pay for reducing the risks? Risk assessment is a trade-off between these two factors. Damage to the animals is factual as we can provide evidence of the damage cause by the Draize-test for each substance. The moral aspect of damage is the acceptance of damage and further, what type of damage is acceptable.*

> *These are aspects of the actual case. The further contextual analysis includes thinking about the legislation regarding cosmetic products - what are the requirements for the cosmetics companies regarding product safety? Further, one is easily led to consider other types of animal tests in*

the pharmaceutical industry: are they acceptable? Also, one can view the
use of animals in other contexts - for food and clothing and the treatment
of animals in those situations. One should also look into alternatives to the
Draize-test for product safety - Are there less painful methods available?
And one can consider the dynamics of the cosmetics industry in general -
whether they have a meaningful role to play in society that would support
their research aims and if some further social issues are embedded in the
functions of that industry, e.g. the status of women.

6.4.3 Element 3 - finding a solution

In the final step the moral decision-maker needs to combine the basic values with
the contextual evaluation. In most cases forming the synthesis is not
straightforward. One may need to look for further facts, balance several
uncertainties, and compromise some of the basic values, when trying to find the
best possible solution to the problem. Finding a solution is in many cases striking
a balancing between benefits and harms. This does not indicate a utilitarian
approach to solving moral problems, however, because one should include in the
cost analysis both the cost of compromising basic values and the indirect costs to
social structures.

Not all decision-making methods are equally good. Good methods are grounded
in reason and supported by rational arguments and contextual facts. So the first
element of good reasoning is an ability to include and appreciate both the basic
values and the contextual facts. Good moral reasoning is not possible without
good groundwork - the proper analysis of both basic values and contextual facts.
In analysing personal core values and contextual elements of the case one needs to
become aware of these elements, while the decision-making process requires
evaluating and combining these elements into one moral solution.

Second, one must follow logic in presenting solutions, arguments or agreements
in any issue. It is important to follow logical steps from premises to facts to
conclusions. A typical deductive argument goes as follows:

Premise	if A, then B (If Spotty is a terrier, then he has four legs)
Fact/Evidence. Supporting statements	A (Spotty is a terrier)
Conclusion	Therefore B (Therefore Spotty has four legs)

Deductive arguments have truth-preserving quality which means that if the premises are sound/true and the argument is valid, then the conclusion must be true (Olen and Barry, 1996).

We can approach moral problems with deductive logic, by starting with a moral premise and working logically from it to a solution. In purely abstract moral problems, we do not have supporting evidence, while the number of premises is often greater than one. For example:

Premise$_1$	It is wrong to harm.
Premise$_2$	Abortion harms the fetus.
Conclusion	Therefore abortion is wrong.

But because there is rarely a unanimous agreement on the moral premises, we can form an equally valid logical argument with entirely different conclusions by changing one premise and/or including an additional one into the logical sequence, for example:

Premise$_1$	It is wrong to harm a person.
Premise$_2$	Abortion harms the fetus.
Premise$_3$	The fetus is not a person.
Conclusion	Therefore it is not wrong to have an abortion.

The reasoning process is valid in both examples, but the conclusions are contradictory. Even though logic is important in solving a moral problem, these examples should highlight the point that logic alone will not be sufficient. One needs to scrutinise the premises of arguments.

In the abortion examples, the problematic premises are:

1. What are the grounds for stating that the fetus is not a person - what are the criteria for personhood?

2. Is it *always* wrong to harm - can we think of instances when some level of harming can be accepted for producing a greater good, for example self-defence?

3. Does abortion harm the fetus - what qualifies as harming?

A lot of moral premises are intuitive and thus one cannot refer to facts to prove them to be true or false. They are often convictions based on core values, not on rational deliberation based on observable facts. Sometimes, they are a result of careful consideration of relevant facts while at times a conviction just strikes us as right, but all the same, in the end we cannot, in a logical sense, prove that our premises are true. Logic only provides a tool to test the argument, not the premises. So we often have to accept disagreement about premises, which inevitably results in disagreement about the conclusions as well. This acceptance does not free anyone from the importance of trying to explain one's moral intuitions, to re-assess them in the light of new evidence or experiences, or to evaluate any counter-arguments that may be presented against them. But in the end, people do disagree on moral premises.

The deductive logic explained here is void of case particulars. It works on the premise that there are case-independent values, which can be used in these types of deductive syllogisms and a right answer is provided at the end. I hope that the examples above are enough to cast doubt on the sufficiency of approaching real moral problems in this manner. Besides logic, moral decision-making requires an ability to understand the relationships between facts and values, an ability to predict consequences of decisions, and a skill to weigh different elements against each other in order to arrive at the best possible solutions at that point in time.

The description of the decision-making process is unavoidably fuzzy. There is no available formula into which to feed the basic values and situational information

and out pops the 'right' answer. Moral decision-making is a balancing act of competing values based on rational and logical reasoning. It cannot be learnt once and for all as it is a process that continues for a life time. Practice in making moral decisions gives one an ability to understand problems and solve them with skill and confidence.

For example, returning to the Draize-test:

> *The most important fact in this case to me is that animals are suffering without good cause. I believe strongly in safety (no 4 of my basic values), but I do not believe that this is an acceptable way of achieving product safety. I would like to see cosmetics tested in a way that would not cause suffering, which rules out the Draize-test. If I am asked why I object to the suffering of these animals, I would respond that I do not believe in causing harm. And I believe in that very strongly. I can imagine there could be cases where causing some harm is acceptable, maybe some types of drug tests, but I would always want to make sure that the suffering is limited to its minimum and with the Draize-test, nothing is done to reduce the suffering of these animals.*
>
> *Also, because cosmetics are luxury items and it is an industry that makes high profits and is wealthy, I believe it would be fair and just to require companies to make investments in developing alternative testing methods, which would not cause any suffering. Testing on tissue cultures could be an alternative.*
>
> *I realise that this case limits the rights of the cosmetic companies to do what they want, their autonomy as companies. But I also believe that curbing their freedom is acceptable because no one should have the right to harm others without exceptional justification. So the freedom of choice should always be a limited one, freedom to do what does not harm others. Therefore my conclusion is that cosmetics companies should not perform Draize-tests and that society should use legal powers to prevent these tests because they cause unnecessary pain.*

6.4.4 Dynamic links

None of these Moral Decision-Making elements exists in isolation and the links between them are not limited to one direction only. There are three dynamic links in the moral decision-making model:

1. The basic values influence the evaluation of contextual facts.

2. Previous solutions guide the analysis of contextual facts.

3. Most importantly, after the solution has been reached it is important to do a re-check on how that solution supports the originally held basic values. In case of serious discrepancy, the decision-maker has two options, either to re-evaluate the basic values or to work through the decision-making process again in order to find a solution that does support the basic values. In many cases both are necessary: re-tuning the understanding of basic values and adjusting the solution to the problem.

In my solution I find that personal freedom is not unlimited and that I feel very strongly about not harming animals unnecessarily. Looking at my personal values as I viewed them before I was presented with this problem, I now feel that I need to make the following adjustments:

1. Personal autonomy needs to be supplemented with a statement that one has the freedom to do and choose as long the decisions do not cause harm to others.

2. I am even more serious about my commitment to the value of non-maleficence; I may not be ready to make it my first and most important value above personal autonomy, but I do find now that it is more important to me than safety.

6.4.5 Sequence of elements in teaching

It may not always be most productive to try to work through the moral decision-making process starting with basic values, then moving onto contextual facts and finally reaching a solution. Starting with basic values may lack motivational

elements due to its quite abstract nature. Also, asking students to start with in-depth analysis of their own values can prove detrimental as the students may feel inadequate and lost when asked to analyse values and iterate how their values reflect on reality, in particular in front of their peers.

I believe that the most useful place for an ethics curriculum to begin is with analysis of contextual facts i.e. raising moral sensitivity. Analysing moral problems provides an opportunity to introduce students to forms of moral thinking without a need to provide any material on complex moral theories or to ask the students to make instant evaluations of their own value systems. Moral sensitivity exercises should present students with a problem situation and guide them through the analysis, supporting their own ability to recognise moral aspects in the situation and to provide them with further insight when their ability to reach further has been exhausted.

Moral sensitivity exercises can be followed by two types of exercises. One approach is to deepen the moral analysis downwards into the basic values. This can be done by distilling from the moral problems found in the moral sensitivity exercise the basic values present in each moral aspect. An alternative approach is to move onto moral decision-making and seek a solution to the moral problem. After a tentative solution has been expressed, the process can work down to the basic values and examine what type of values the solution supports and whether these are the values that should be supported.

6.5 Summary

The premise for this chapter, and the entire research described in this thesis, is a view that moral development involves improvement in skills and abilities to recognise and solve moral problems. It follows from this that ethics education which aims to encourage moral development has to adopt a skill-based approach instead of a rule-based one. When moral education is approached as the learning of moral decision-making skills and practising moral recognition abilities, there should be no reason to believe ethics should be fundamentally different, from an

educational point of view, to other subjects and disciplines. Moral skills can be taught alongside and equally well as scientific skills.

But moral skills are not on a par with scientific skills in all aspects. One of the important differences is that moral development occurs without any educational intervention, while scientific development normally requires a conscious effort from both the teachers and learners. Naturally one can learn science without a teacher, but all the same, a conscious learning effort is required. In moral development, on the other hand, learning is more often 'hidden' within everyday experiences. And even more importantly, unlike science, moral development takes place more often than not outside the sphere of formal education. Therefore moral education cannot claim any monopoly in encouraging moral development. At the time of young adulthood and higher education research has found that general intellectual stimulation, and personal involvement in studies and leisure activities, predict a general rise in moral development scores without any overt educational intervention.

But ethics education has an ability to accelerate the natural progress of moral development. Research has shown that an ideal arrangement for a moral education programme targets young adults who have developed their basic logical skills. The ideal programme would last between 4-12 weeks and have a student-centred structure that encourages students to get involved as independent moral agents.

The Moral Decision-Making (MDM) model described has three elements: 1) Recognition and appreciation of core personal values, 2) skills in analysing the case and/or situation, and 3) skills in forming a solution that appreciates the two previous elements adequately. The three elements form dynamic links, where the core personal values influence the analysis and recognition of the case and situational particulars and where the decision is judged against the core values. This model provides a framework for understanding the different elements ethics teaching should incorporate. This incorporation does not take one possible route alone, but allows for great variation in teaching styles and methods. It is suggested

here that a best place to start is with the case/situation, from which the discussion can move to either personal core values or decision-making.

7. Teaching approaches

Teaching methods are at the core of moral education programmes, as was discovered by Schlaefli *et al.*'s (1985) meta-analysis discussed in section 6.3. The key elements described by Schlaefli *et al.* related to increased support of student involvement and encouragement of participation as an independent moral agent. This chapter will look at the teaching methods that provide a setting for student involvement as moral agents.

The chapter starts with a discussion on student-centred learning (section 7.1). The emphasis is on a review of what it means to say that an educational programme is student-centred and how such programmes differ from teacher-centred ones. Adoption of a student-centred approach has a major impact on the teacher's role, and how the student-centred approach in practice influences the role of the teacher is discussed in section 7.1.1.

The last two sections in this chapter outline two pedagogical approaches which allow uncomplicated assimilation of the MDM into the teaching of ethics. These two approaches are Problem-Based Learning (PBL) and Structured Discussion Groups. Both methods have been used extensively in other subjects and they require only minor changes to become suitable methods for teaching ethics. The elements in these approaches will be discussed in sections 7.2 and 7.3 respectively. The chapter ends in a short summary.

7.1 Student centred learning

The advances in moral development as a result of moral education discussed in chapter 6.3 indicate that the logical approach to moral education is student-centred instead of a teacher-centred. Student- and teacher-centred approaches are two ends of a continuum, and a course can include elements from both extremes or work with methods that would be best located between the two extremes. In teacher-centred learning the teacher is solely responsible for what the student is

expected to learn - what information, by which methods, in what sequence, and at what pace. Traditional teacher-centred learning is in the form of lectures and structured laboratory work where teachers/lecturers dispense information, assign readings, choose learning criteria and test methods, and structure the learning process. Self-study modules can also be part of teacher-centred learning, as long as the teacher determines the resources to be studied, the sequence of study and the learning that is to be mastered and the only elements the student is allowed to influence are the pace and learning methods.

Teacher-centred learning can be an ideal format from the teacher's point of view - they can readily dispense their specialist knowledge in a compact form based on their previous teaching experience. They can formulate learning of complex issues into more easily digestible capsules, given the control they can exercise over the structure of the learning contact. The teacher can also be certain that the students will have been exposed to all the knowledge and concepts considered necessary for the level of learning in question. The teacher-centred approach likewise saves students from the agony, frustration, and time it would have taken them to work through the subject alone and gives them confidence of knowing what they are supposed to learn in each particular course.

One down-side of the teacher-centred approach is that it does not take into consideration the heterogeneous student backgrounds. Students come to the course with different levels of knowledge, cognitive structures, and learning styles and abilities. In a teacher-centred learning situation all students meet the same information in an identical form when their learning might have been better supported by different methods or a different cognitive starting point. Even more importantly, students have a passive role as an acceptor of knowledge instead of an active processor, an element which has been shown to hinder learning leading to moral development (Hartwell 1995).

Alternatively, education can adopt a student-centred learning approach. The key element in student-centred learning is that students learn to determine what they need to know. The students will eventually take full responsibility for their own

learning in respect of what, how, and why. This may require considerable input from teachers/facilitators at the beginning, but even at the early stages, the aim is to increase student responsibility and decrease teacher dominance.

Since the learning is self-determined and acquired through the students' own study, the student becomes an active participant in the learning process, an autonomous agent. This responsibility can foster motivation as the rewards are internal - the desire to learn for personal or professional growth, not for teacher-dispensed rewards (Albanese and Mitchell, 1993; Norman and Schmidt, 1992).

According to Race (1998) successful learning is underpinned by five factors which influence motivation:

1. wanting to learn (intrinsic motivation)
2. needing to learn (extrinsic motivation)
3. learning-by-doing
4. learning from feed-back
5. making sense of what has been learnt.

Wanting to learn is the most satisfactory state for students to be in - their motivation is not dependent on external changes and the rewards are equally independent. Student-centred learning, in particular in small groups, provides wonderful opportunities for the last three factors to support student learning. It is more difficult to design lectures to support these learning factors. Designing motivating ethics lectures for science students would be further hindered by the low levels of intrinsic motivation of science students to study ethics: they may not have thought that ethics should be part of their science training, though this view is not necessarily justified.

Student-centred learning is not only a process of learning facts, but a process of learning to learn, which is an important asset in meeting a lifetime need to adapt to new knowledge, challenges and problems. This form of learning requires self-discipline and an ability to organise one's own learning to fit personal learning styles and cognitive structures. At the early stages of student-centred programmes

learning to learn might even seem to dominate over learning facts and skills, but this should change with experience. In student-centred learning the value of the learning process becomes almost as equally valued as the content (Katz 1995).

A student-centred learning approach is most suitable for an ethics education which has autonomy and skill-based aims, as in this project. The core of the aims is in the personal development of the students. Ethics education is successful if the students gain skills that they can effectively use in their personal lives. As the aim is to encourage skills applicable to personal life situations, it is logical to approach teaching those skills from a student-centred perspective. The students are then given an opportunity to define the scope and nature of ethical problems relevant to their personal lives and to practice the moral decision-making skills at the level they find comfortable and productive. Also, ethical skills are for personal benefit more than anything, and the motivation to learn them is most naturally nurtured within students rather than exposed from the outside.

7.1.1 Role of the teacher/facilitator

Adoption of student-centred learning should not lead anyone to believe that the teacher's role is reduced to that of a peripheral consultant who is only marginally necessary. Quite the contrary, effective facilitation is essential to successful student-centred learning. Teaching ethics is no exception to the rule. Even further, a good teacher is likely to be essential in sparking students' interest in ethics. Teachers have the power to nurture and even create enthusiasm towards a subject. This power of directing and coaching students' motivation and interest can also be used negatively, to kill inborn enthusiasm and dampen students' desire to learn. I believe that the role of the teacher in an ethics course is particularly important because ethics is a very personal subject which invites students to share their intimate beliefs about the world with the class. They are unlikely to do this if the teacher/facilitator has not managed to create a good learning environment and build a positive and encouraging relationship with the students. The success of an ethics course, in other words, depends heavily on the skills of the teacher/facilitator.

The essence of facilitation is to create an environment for learning - for participants to define and advance their own learning goals. It is a process of helping students to learn to trust their own decision-making and problem-solving skills, and to foster the students' ability to think critically. The aim is thus not to transmit empirical data, but to assist in the integration of theory and practice - to support students to develop the required enquiry skills to identify what they need to know and how they may proceed to find the required knowledge (Katz, 1995).

Facilitators need not necessarily be experts in the subject the group is working on. Research has shown advantages in both the facilitator being an expert and a non-expert. Groups with expert facilitators have been found to be less likely to engage in student-directed discussions and collaborative learning (Silver and Wilkerson 1991). A similar effect was found by Davis *et al.* (1992), whose research showed a trend for expert facilitators to move towards using directive behaviour such as lecturing or giving advice that increased their influence over the group. But while expert facilitators may use less orthodox facilitation methods, the students from expert-led groups tend to score higher in the course final examination. Albanese and Mitchell (1993) suggest that while expert facilitators may be less facilitative, they appear to be better able to help students to identify relevant learning issues and correct gaps in knowledge and errors in processing.

Based on these results and my personal experience as a facilitator, I support the use of expert facilitators in ethics teaching as long as they have the motivation to adhere to the facilitator role and skills to be able to support three elements of student-centred learning: 1) prepare good material, 2) facilitate appropriate cognitive learning, and 3) encourage group skills. But I hasten to add that simply being an expert does not make anyone a good facilitator. Just as expert knowledge alone does not turn anyone into a brilliant lecturer, experts are not necessarily good facilitators either.

Ethicists do not claim to be morally wiser than the rest of us, but they have expert skills in the analysis of moral problems: to examine the consistency and

coherence of moral reasons by employing a more disinterested reflection on a wider range of experience (Smith and Boyd, 1991). These skills are beneficial in ethics teaching. One of the advantage of expert facilitators is their ability to recognise all the moral problems involved in a case, and thus be able to guide students in their own recognition process. Non-expert facilitators may overlook some ethical consideration and thus fail to encourage the students to look at all the relevant issues. Also, expert facilitators have a firmer grasp of ethical theory and moral decision-making procedures, which gives them an ability to clarify moral issues to the students if they were to find themselves faced with a moral dead-end. Further, expert facilitators in ethics may be more equipped to isolate their own moral judgements from the discussion and thus allow more room for the students to establish their own. An expert facilitator in a subject like bioethics, which contains elements from more than one academic discipline, would benefit from expertise in both subjects. But in my view, in the case of bioethics, the more important expertise in is in ethics, not in bio-science.

My personal experience as a facilitator of bioethics courses with science students supports this view. The students often arrive at the discussion with only an inkling of the ethics issues relevant to the topic for discussion, despite the fact that they may have completed the preliminary reading. It was thus my task as a facilitator to both initiate recognition of the main ethical themes and systematically analyse with the students the views they held of the subject. I felt it was essential that I could draw from a strong knowledge-base in philosophy and not be caught in hesitation or confusion born out of lack of competence. From my experience I can only imagine how difficult good ethics facilitation would be if the facilitator could not rely on a background knowledge in ethical theories and have experience in dealing with bioethical issues.

There is no single template for an ideal small group facilitator, expert or non-expert, while there seems to be certain characteristics that most capable small-group facilitators share (Westberg and Jason, 1996):

1. *Enthusiasm*: not only enthusiasm for the subject matter, but also for the processes of learning and teaching. Enthusiastic facilitators enjoy teaching and contact with students.

2. *Caring for learners as people*: effective educators care about their students as people and are able to convey these feelings - they enjoy nurturing others and watching them grow.

3. *Capacity to put oneself in the learners' shoes*: it is vital that the facilitator has an ability to understand the students' reactions from the inside and to be sensitive and aware of the ways the students perceive their learning.

4. *Trust in the group*: effective small-group facilitators respect the intelligence and experience of their students and convey this faith in their teaching.

5. *Comfort with empowering learners*: this is an ability to let go and retreat to the side-lines when the group learns to lead itself.

6. *Commitment to personal learning*: a desire to maintain a personal learning curve, eagerness to improve and active pursuing of new challenges and knowledge.

7. *Flexibility, adaptability and inventiveness*: effective educators adapt to, and even welcome, the surprises of small-group teaching - they can make moment-to-moment adaptations and continuously reshape the instructions.

7.1.1.1 Preparation

Just as in traditional teacher-centred teaching, preparation by the facilitator is essential to successful student-centred learning of ethics. One of the elementary preparatory tasks is to become aware of one's group: how large is it going to be, how diverse will it be in age, cultural backgrounds and educational experiences, and what they have learnt of 'your' subject before. Familiarity with these and some basic resource limitations (time, room, equipment) will allow the facilitator to approach the planning with realistic aims and methods in mind.

An important part of the preparation is to find out how familiar the students are with the chosen teaching approach. If the group is new to the student-centred teaching approach, time needs to be allocated to introduce the concept and the first tasks should cater for the need to learn the process as well as the cognitive content. Poor understanding of students familiarity with the teaching approach can be detrimental to learning as poor preparation can easily lead to confused and frustrated students. This might be particularly true with teaching ethics to science students, when the students are not only introduced to a new academic discipline, which requires them to develop new modes of thinking, but also to a new form of teaching. I will discuss in more detail the skills involved in coaching students to use the new teaching approach effectively in section 7.1.1.3.

Facilitators in bioethics groups are always leaders in the sense that they choose the topic and the time scale for the exercise, while facilitating the actual exercise lends itself to group leadership and one of the essential tasks of facilitators is to encourage students to share the leader role. The facilitator is responsible for fostering and encouraging leadership skills and behaviours among the students.

But before effective group-leadership can be supported, students require good learning material. The quality of learning is dependent on the quality of materials used, and it is the facilitator's task to choose materials that are deep, interesting, and worthy of discussion and further research. No matter how well the group dynamics work and how motivated the students are, poor material will hinder the group in reaching its learning potential. The facilitator needs also to be clear before the group sets to work about what alternative resources may be needed, and to sort out any possible access difficulties to these resources (Barrows, 1992).

For this reason, if the facilitators themselves are not experts in the academic subject the students are investigating, they should seek expert support in the preparation of the material. In ethics, having personal opinions in moral matters does not qualify one as expert in preparing material for an ethics course, whether it is PBL or a discussion group. To teach ethics effectively, the material needs to provide a platform to discuss ethical issues in a structured manner and to prepare

material to provide that platform requires knowledge of those structures, i.e. ethical expertise.

The facilitator should also be very clear on what sorts of activities the learning group will be involved with. I will examine below two alternative methods for ethics teaching: PBL (Problem-Based Learning), and learning through structured discussion. The choice of activities will lead to decisions to clarify the schedule to ensure that all activities can be completed within the time scale with sufficient time for independent study and breaks.

Teachers should prepare, similarly to teacher-centred learning, what they feel are the appropriate learning objectives, learning resources and evaluation material based on their own particular experience and knowledge. In student-centred learning, these objectives and resources serve as guidelines to the students, to be adapted and used as they feel appropriate. Facilitators are better equipped to assist the group in framing questions and focusing their thoughts on the essentials, rather than getting fixed on minor details, when they have a clear picture of the learning objectives and methods of achieving them. Without assistance, groups may become overwhelmed by the task of selecting questions which will assist their understanding of the subject and forward their learning. Again, discipline-specific expertise will guide the facilitator to choose the aims and objectives, which in turn will provide the students with a better starting point in accepting or modifying them. In setting the learning goals, facilitators, experts or not, need to ask themselves the following questions (Westberg and Jason, 1996):

1. Are the goals sufficiently specific?
2. Are the goals clear and understandable?
3. Are the goals appropriate to the learners' stages of development?
4. Can the goals be achieved with the available resources and the time allotted?
5. Are the goals worthwhile?
6. Are the goals consistent with the overall goals of the school or programme?
7. Should the goals and objectives be in writing?

7.1.1.2 Facilitating cognitive learning

In advancing the cognitive learning process in student-centred learning, the facilitator's role is to guide students through the enquiry and decision-making process, question the rationale of their judgements, and challenge their assumptions. In other words, facilitators have a task to probe students' knowledge deeply.

This role is important because students are often not aware that they are making assumptions when they speculate on factors they are not familiar with. The facilitator's role is not generally to give answers, but to ask the right questions, to ignite new thoughts and ways of looking at things the students would otherwise have ignored or disregarded. The facilitator advances the learning process by presenting alternative interpretations of what has been analysed and by relating the particular learning content to the wider social and political world. This is also where expert facilitators have the advantage as they know what is important and can thus encourage learning in appropriate directions. Teaching ethics to science students may require a lot of facilitator guidance and questioning because: 1) ethics is a subject open to multiple interpretations, 2) defining terms is an important part of an ethical inquiry, 3) ethical issues are inter-linked with social, political, and economic policy, and 4) students are relatively unfamiliar with the methods of ethical inquiry.

To do this, according to Barrows (1992), the facilitator must constantly ask 'why?', 'what do you mean?', 'why do you think it is true?'. The facilitator needs to ask these questions again and again until the students have got down to the depth of understanding and knowledge expected of them, and until they have explored all they know. Facilitators must never let ideas, terms, and explanations go undefined. Only by fulfilling this role can a facilitator have any guarantees that the depth of learning in small groups is adequate.

Barrows (1992) also emphasises that it is the facilitator's responsibility to continually monitor the educational process of each student in the group.

Facilitators need to be alert to detect any difficulties in reasoning, understanding or finding information so that these problems can be brought to the group's attention for early intervention and help. As facilitators monitor students' learning, they also have a responsibility to make sure that students do not embark on research within the exercise that is too unwieldy.

7.1.1.3 Facilitating group skills

Even though students should have control over their own learning process, in particular at the early stages of introducing students to this type of learning, the role of the facilitator is important in guiding the process and providing support and materials when needed. Learners tend to need direction when their competence and confidence are - or feel - low. At the same time, the facilitator should not always interfere when problems occur or students feel insecure, as learning benefits can be derived from learners struggling to figure things out for themselves. It is often most beneficial to make decisions regarding facilitator participation together with the students. Facilitation is not a one-off process, but an on-going one. Thus the decision regarding facilitator participation needs to be re-considered from time to time (Westberg and Jason, 1996).

In a sense, facilitators are group-trainers. They will train students in the process of working in groups, whether that is around a PBL exercise or a discussion group. At first, students are unfamiliar with the process and it is the facilitator's task to train them in the process skills and support them during the early practising of these new skills.

The facilitator may also assume the role of a chairperson by reminding students to listen to one another and to provide suggestions impartially and objectively. There is a fine line to tread in controlling the learning situation and providing assistance where it is needed and when it *can* assist the students in their learning. One sign of successful facilitation of group processes is that the facilitator's role as a chairperson will increasingly be taken up by other members in the group.

Learning inter-personal and communication skills are significant learning tasks in group work - to learn to give and receive positive feed-back, active listening, understanding and awareness of body language, the dynamics of communication and confidence in presenting one's point of view. The level of group members' inter-personal skills will significantly affect the facilitator's role in the group process - every group situation needs to be assessed independently and facilitators need to adjust their role to suit the group and the situation it is in. In some situations facilitators are quite clearly leaders, while in others less definitely so and yet in others definitely not.

7.1.2 Learning in groups

In groups, knowledge is created by expressing ideas and perspectives, by trying to make sense of information, by engaging in collective inquiry, by hearing the perspectives of others, by reflecting on new information and challenges, and by constructively challenging one's own and each other's assumptions and thought processes. The learners give and take - they teach and learn - simultaneously. It is also hypothesised (De Volder *et al.* 1985) that students might learn better from their peers than from faculty members because of the greater congruence between the semantic network representations between peers than between students and faculty members. The group members, in other words, engage actively in doing what is needed for meaningful and lasting learning (Westberg and Jason, 1996).

Ideally a learning group should comprise 5-15 students. Katz (1995) suggests 10-15 for PBL groups, while Westberg and Jason (1996) recommend 5-8 for other types of small groups. Personal experience contradicts Katz's recommendation to some extent as groups of 15 students seem very large for PBL and the ideal number seems to be closer to ten. Larger groups are less successful because the more verbal or assertive members tend to dominate the discussion and groups smaller than 6 have a problem with lack of diversity, viewpoints, and ideas, despite the increased potential for interaction. The actual group size is naturally

influenced by external factors of class size, teaching facilities, and resources, but it should also reflect the task at hand.

Learning in groups is based on interactions between the group members and in the independent study completed before, during, or after the group work. Group interaction can either encourage or hinder learning. The personalities of group members are given and cannot be changed by facilitation styles or by other group preparatory measures. But dysfunctional personality traits are not necessarily a recipe for a dysfunctional group. Even students that come across as difficult, have personal properties that might be useful for the group and it is the task of the facilitator and other group members to learn about each other in order to use the expertise of each group member in the best possible way. In Appendix XI - Group Personalities, I have briefly looked at group personality theories.

7.2 Problem-based learning

The idea of problem-based learning (PBL) is nothing new, but rather a very old approach to learning. It has a starting point in a problem, a query or a puzzle, which the learner wishes to solve (Boud, 1985). It is an instructional method characterised by the use of problems as a context for students to learn problem-solving skills and acquire knowledge (Albanese and Mitchell, 1993). As in the real world, the problems can be complex and multi-faceted and not confined to the often artificial boundaries of a single discipline. This is particularly true for science ethics.

PBL exercises have been most commonly used in medical education, where clinical situations are presented to students who then proceed to make a diagnosis and suggest treatment. PBL as a method in medical education originates from McMaster University in Canada, where it was established over 40 years ago. In medical schools PBL based curricula were adopted due to dissatisfaction with traditional curricula. PBL has been considered a way to provide an exciting and motivating way for students to learn (Barrows 1980).

PBL can be used to help learners at all levels and in all fields of professional education while PBL can take many forms depending on the subject and of the particular goals of the programme. Many universities have chosen to convert some subjects completely to PBL, e.g. the University of Glasgow medical faculty, but it is also possible to introduce PBL into a curriculum which maintains elements of traditional didactic teaching.

Typically, PBL exercises consist of six basic elements (Barrows and Tamblyn, 1980):

The problem:

1. The problem is encountered first in the learning sequence, before any preparation or study has occurred.

2. The problem situation is presented to the students in the way it would present itself in reality.

The challenge and learning needs

3. The students work with the problem in a manner that will permit their ability to reason and apply knowledge to be challenged and evaluated.

4. Learning needs are identified in the process of work with the problem and as a guide to individualised study.

Learning results and application

5. The skills and knowledge acquired by study are applied back to the problem, to evaluate effectiveness of learning and to reinforce it.

6. The learning that has occurred working with the problem is summarised and integrated into the students' 'knowledge and skills'.

7.2.1 What characterises a good PBL case?

In a well-designed PBL the ways the learners approach and solve problems, and the ways they acquire and organise knowledge, parallel what they will need to do as professionals in their discipline. These activities give them practice in approaching and solving problems, being self-directed, and participating in the process of collaborative learning (Norman and Schmidt, 1992).

The problems are primary and all learning is directed at solving the problem, reaching an agreement, or presenting a recommendation. In some professional studies it is easy to provide students with realistic problem-situations they are likely to confront in their professional lives - in medical schools they can be presented with simulated patients manifesting particular symptoms, in engineering a problem requiring students to design a structure to fulfil a given purpose, and in law, ask students to advise a hypothetical client. In all cases the students should have access to all the information a real practitioner would have access to - students can ask for results of certain medical tests, the performance requirements of the engineering device, or copies of the legal contracts. In science ethics PBL exercises, the problem can similarly represent a problem the students are likely to meet in their professional lives, but the problems are often less concrete and direct and may require students to adopt the hypothetical roles of decision-makers. The problems can range from a situation of personal judgement on the use of animals in research to deciding on the ethical acceptability of certain corporate activities.

In designing a PBL scenario Prideaux and Farmer (1994) suggest that you remember to make your case:

1. Relevant: the problem should illustrate a problem, which is realistic to the students, with which they can identify.

2. Multifaceted: the problem should provide more than a narrow view of a situation, but provide a holistic view of the entire field.

3. Integrated: the case should link with other studies the students are working on in other parts of their course, to increase its relevance and usefulness.

4. Consistent: the case should be consistent with the general course-requirements and other forms of teaching to create the best possible learning outcome and significance to the students.

5. Motivating: the problem should attract the students to find a solution because they see its relevance, find it intriguing intellectually, and see a chance of success at the end.

The degree of difficulty of a PBL exercise needs to be carefully weighed. The problem should be challenging enough to provide an opportunity to extend students' thinking and levels of knowledge, but not too challenging for the students' ability to complete it successfully. PBL exercises providing insurmountable challenges are less likely to be positive learning experiences: even though learning may occur, it is gained via negative emotions and feelings of inadequacy.

A good PBL case, according to Barrows (1994), presents the students with a problem:

1. In which more information is needed for understanding the problem than is immediately available.
2. There is no one right way to get this information.
3. The understanding of the problem evolves and changes as new information is acquired.
4. The problem solver cannot be sure that the analysis of the problem, or its solution, is definitely the correct one.

Moral questions are naturally of this kind, as they are rarely simple, information is often difficult to acquire, the view of the problem frequently changes during investigation and there is hardly ever guarantees of having arrived at the one and only correct answer to the problem. In many ethics PBL exercises, students are able to choose their own level of approach and tackle the problem according to their capabilities - all ethical problem can be presented in a way that students can comprehend and get started on, while at the same time no ethical problem is too trivial or easy when enough depth and rigour is applied to its analysis. Because of this extendible quality of ethical problems and the multitude of possible approaches of both kind and depth available, the presentation of the problem and the role of the facilitator are important for the success of the exercise.

7.2.2 Steps in PBL

From the students' perspective PBL consists of seven steps:

1. Clarify terms and concepts not readily comprehensible

2. Define the problem

3. Analyse the problem

4. Draw systematic inventory of the explanations inferred from step 3

5. Formulate learning goals

6. Collect additional information outside the group

7. Synthesise and test newly acquired information (Matthew, 1999).

Bransford and Stein (1993) compressed these into five steps with the acronym IDEAL:

1. Identify

2. Define

3. Examine

4. Act

5. Look

Even though Bransford and Stein approach problem-solving from a very practical and application/invention view point, their model of problem-based decision-making is a suitable format for ethics PBLs. The description of the IDEAL components here differ in many places from the description given by Bransford and Stein, as I have adapted the original structure to suit moral problem-solving.

The first activity in relation to any problem is the clarification of the terms and concepts not understood at first sight. This identification step involves finding the standard meanings of the terms involved in the problem scenario, either by discussing them with the group or consulting a dictionary. For example, if the PBL exercise starts with the following problem formulation: "You are a team working on vaccines. You have two options for future research: either the development of a vaccine for Meningitis or Leprosy. Present your choice with supportive arguments." - the students may wish to check what is meant by vaccine (dictionary definition: any agent used for inoculation and immunisation), meningitis (inflammation of the of the three membranes that envelop the brain and spinal cord owing to an infection) and leprosy (a chronic contagious bacterial disease, usually characterised by the formation of tubercles or of painful inflamed

nodules beneath the skin, thickening of the skin, loss of feeling, and ulceration and necrosis of the affected parts.) Other problems may include value-terms like 'good' or 'fair', which require conceptual analysis in order to agree on a working definition for the following group discussion.

The second step is to produce an exact definition of the problem. The key element is to decide, as a group, which interrelated phenomena should be explained. Many PBL exercises do not present difficulties in this respect, but in ethics PBL exercises this is possibly the most important aspect of the process. Ethical issues are more often than not hidden, and this requires students to investigate the problem deeper than the surface level before the ethical issues can be discovered. Ethics PBL exercises can be phrased to emphasise one particular ethical problem, but in most cases that is not the most realistic presentation of an ethical problem and more importantly, if the ethical problem is very apparent, students will not gain the learning experience of recognising the moral issues, which was defined in Part I section 3.2.1 as one of the core elements of moral development. The role of the facilitator in this stage is very delicate. On the one hand, facilitators should restrict their input to the minimum, allowing the students to determine the recognition process, while at the same time the facilitators need to be able to ask appropriate probing questions if the students either find it difficult to get started or if they have ignored a fundamental element in the definition of the problem.

The definition of the problem is probably most successfully carried out as a brain-storming session - either in the whole group together or in the group divided into smaller sub-groups whose results are then collected for the entire group. The problem definition substantially consists of recapitulation of group members' opinions, actual knowledge and ideas about the underlying processes, mechanisms, and value structures. Analysis can be a free association-round where students verbalise both their previous knowledge (I've read somewhere that...') and try to formulate relevant hypotheses by reasoning ('Could it work like this...').

After the definition of the problem students need to examine various explanations to their problem. In ethics PBL exercises this means looking at possible elements in the ethical analysis and how to investigate them further and possibly also a discussion on the decision-making methodology the group is going to apply once they feel they have learnt enough of the different issues. An ethical problem consists mainly of the two elements, understanding the value-related elements of the problem and choosing a method of deciding between competing values in cases where any solution is at odds with one or more values.

Actions taken by students involve studying the material available to them on the problem, discussions with each other on progress and deciding on the division of labour, use of methods, and time-scale. At the end of the action phase the students need to agree on a solution to the given problem. In many cases this is the crunch time for the students - they need to pull together all the information they have collected and compare that with their basic values and agree on a decision as a group. The task is multi-faceted and demanding.

The last and important step of an ethics PBL is to return to the problem, look back to the analysis and make a judgement on how well the group have answered the questions, how satisfied they are with their work and its results.

7.2.3 Facilitating PBL

Through engagement with the problem students are expected to identify the kinds of knowledge and information they require in order to find a solution. The facilitator can be helpful in this process both in assisting the students to ask the right questions and as a source of reading material or access to other information sources.

Asking the right questions is of paramount importance in ethics PBLs. The analysis of the problem is more crucial and difficult in many ethics PBL exercises in comparison to more concrete PBL exercises and the success of the exercise will largely depend on a well conducted initial analysis of the problem, agreed usage

of terminology and shared understanding of key concepts and the shareholders. An expert facilitator is best able to judge when the group has sufficiently analysed the case and should be encouraged to move forward.

How extensively the students should be expected to find their own resources is highly dependable on their previous skills and on the time constraints of the exercise. In a pure PBL model, students are responsible for all resource acquisition and the facilitator is there to help with questions, but not to directly offer material. When time is more restricted the facilitator can act as a material source more directly, either providing detailed guidance on how to find it or providing necessary information upon request. In some cases it might also be necessary to provide students with a resource base (a collection of books, articles or other material), which should be adequate to solve the problem. The students' task is then to use this material to their best ability, but they are not required (though allowed) to request further information. In general, all the facilitator skills described in section 7.1.1 are applicable to facilitating ethics PBL exercises successfully.

In the process of a PBL exercise students, while working in groups, may also get involved in a considerable amount of individual study. In many cases this may not differ significantly from the study a student might undertake in any course, except in so far as it is focused on finding answers to particular questions the student regards as significant to the problem.

7.2.4 Assessing PBL

Assessment of students performance in PBL exercises should measure student ability to *apply* the knowledge and skills learned through group and individual study in the problem situation. At first developing the skills of goal-setting and self-assessment can be very demanding, but should require less effort the more PBL exercises the students participate in. The assessment should reflect the evolving learning tasks for the students. Applying traditional exams to PBL exercises may be detrimental, as students would be given contradictory signals on

how, and for what to study. The core idea of PBL is for students to choose their own learning goals, while traditional exams propose those goals and measure their attainment according to external criteria, not against the students own learning goals and starting points (Barrows 1980).

Dr Bob Matthew suggested in a private discussion that a suitable assessment for PBL would be a PBL-exam, where the group is asked to solve a similar problem within a limited time, where both their skills in finding information and working as a group would be measured alongside factual information.

Another alternative would a learning log, where students would be asked to record their progress during the PBL exercise. This diary would be handed in at the end of the exercise and checked by the facilitator to make sure all group members participated and learnt something new. The log would be evidence of students' thinking and work. I heard of this approach in an informal meeting at the Teaching and Learning Service, University of Glasgow, and developed it further to be used as assessment in this project.

7.3 Structured group discussions

Learning through discussion should be a student-centred learning activity. The students are to learn from interaction with each other - and to learn not only facts about the subject, but also the skills of being a group member, how to communicate, how to lead and how to differentiate essential from noise.

Too often discussion groups are a source of frustration and provide poor learning environments because they do not cover the material intended, they spend too much time on one aspect to the detriment of all other, get side-tracked to other topics, topics arise haphazardly, contributions are dominated by a few members, and in general lack direction and worth. But it does not have to be so: discussions can be wonderful learning tools in all areas of learning. Ethics in particular is conducive to learning through discussion as the aim of ethics teaching is to

encourage recognition and analysis of moral problems and where to do it better than in discussion with your peers.

Not all groups talking together form a structured discussion group in the educational sense. Ross (1994) suggests a continuum of discussion types that ranges from raw debate to polite discussion, to skilful discussion and finally to dialogue. The further the group moves towards dialogue the more attention is paid to revealing and examining the assumptions and values behind the words that are spoken. In Ross's analysis the primary difference between skilful discussion and dialogue is intention - in skilful discussion the group intends to reach some sort of closure (make a decision, achieve agreement, identify priorities etc.) while in dialogue the group's intention is exploration, discovery, insight and a richer grasp of complex issues. Dialogue may result in an agreement, but that is not its primary purpose. The ethics discussions conducted in this research fall into both the categories of skilful discussion and dialogue. Sometimes it helps students to focus on a particular task, while in others, it is possible to hold their interest with pure dialogues. Exercises devised for this research involved both skilful discussions and dialogues. In many cases the same group will be encouraged to do both, either in the same session or in consecutive sessions. Therefore for the sake of simplicity I will below use the term discussion to cover both skilful discussion and dialogue.

A good ethics discussion topic is similar to a good PBL one:
1. it provides a cognitive challenge to the students,
2. there are clearly more than one possible answer to the ethical dilemma,
3. the topic can be adequately approached using different ethical theories, and
4. it has relevance to students' own experiences.

Bioethical topics, like the use of animals in research, or development of genetically modified (GM) organisms for commercial use, provide excellent topics for successful ethics discussions. They are challenging both cognitively and ethically: the scientific issues are complex and the ethical problems relating to these issues are debated in public and general agreement does not prevail. These

bioethical dilemmas can also be 'solved' using different ethical theories: for example, we can consider the animal testing issue from a utilitarian point of view (minimising harm and maximising benefit) or from a deontological angle (what rights do animals have?). Bioscience students also have immediate personal experiences with animal testing, and their science careers are intimately tied to difficult research questions, such as GM crops.

Discussions are usually most effective when they are provoked by external events (panel discussions, visits etc.) or internal events (reading, videos, role plays etc.). The use of narratives has provided a good starting point for many discussion groups as they can provide a safe access to very personal fears and emotions. Reading stories means participating imaginatively in other lives. This encourages readers to construct their own stories in relation to the reading and consequently the readers come to know themselves better (Weisberg and Duffin, 1995). These events generate interest, focus attention, and create a relevant context. In the discussion exercises created for this research the discussion was provoked by internal events, mainly reading before the session. Most reading was academic in order to maximise cognitive learning within a very limited framework. Fictitious material could have been included if more time had been available.

7.3.1 Structure of a discussion

Rabow *et al.* (1994) provide a procedural tool that outlines an orderly sequence a group can follow in order to learn from discussion. The approach is based on a group having done reading prior to the discussion. The procedure has eight steps:

1. Checking in   (2-4 minutes)
2. Vocabulary   (3-4 minutes)
3. General statements of author's message   (5-6 minutes)
4. Identification and discussion of major themes   (10-12 minutes)
5. Application of material to other works   (15-16 minutes)
6. Application of material to self   (10-12 minutes)
7. Evaluation of author's presentation   (3-4 minutes)
8. Evaluation of group and individual performance   (7-8 minutes)

Checking in. Checking in is not a feature of the first meeting alone, when all members get to know each other, but a feature of all discussion sessions. It allows students to greet one another and express feelings related to the group discussion. This will help to establish a good working environment by showing concern and interest in what everyone thinks and has to say. This should be kept brief though!

Vocabulary. The act of defining terms is not only helpful for the further discussion, but is also learning in itself. In philosophical texts the vocabulary check is not only clarifying the dictionary meaning of an unfamiliar word, but it should also include discussion on how certain words are going to be used in the discussion. Philosophers can spend endless hours on deciding what is meant by basic moral concepts like fairness, or goodness and while students are not encouraged to engage in such in-depth analysis of terminology, they should be encouraged to think of alternative ways of using the value-terms and to agree on some basic ground rules for their discussion regarding word usage. An often heard question in philosophical discussion is 'what to do mean by benefit in this context?' or 'Define useful' - this is essential and vocabulary issues should be revisited later in the discussion if (and when) confusion arises.

General statement of author's message. The purpose of this step is to form an overall understanding of the assigned reading. In some cases this is very easy as the author might have provided a purpose statement or the title of the reading will clarify the issue beyond much doubt. But even when the task is simple, students should attempt to state the meaning in their own words, and when this is well done it launches the group into the next step. It may be beneficial to ask all students to prepare a short, one paragraph, statement on the author's intent before the session in order to get the maximum benefit from this element.

Identification and discussion of major themes. Reading material can be broken down into several important themes and sub-topics. For the sake of time management, a discussion group should not identify more than three or four topics. If more than four themes emerge, the group may have to decide which are

most significant or decide to continue the discussion in the next session when possible. In philosophical texts, identification of themes is of major importance as themes often represent different ethical problems that need to be discussed in order to gain full understanding of the ethical elements in a particular dilemma. Identification of themes as a group is likely to create more diverse understanding of ethical problems and encourage students to learn from each other on the type of issues that can be recognised in an ethical problem. Therefore this step is essential in encouraging better recognition of ethical elements, which was chosen as one of the key aims of ethics teaching. Again, students should be prompted to consider these elements in advance.

The learning through discussion method should place greater emphasis on determining what the author has to say on a particular topic than on the opinions of the students. To a large extent that is important for ethics discussions as well, for learning the skills of recognition and analysis is often best achieved by analysing the text first before launching into discussions on personal views and opinions. Students often have very strong ideas of moral problems while they have read very little on the subject. If personal views are discussed before the author's, less learning is likely to take place as the discussion may become more a platform to strengthen personal views than an opportunity to learn something new which may or may not influence personal opinions. Also, groups starting the discussion with personal opinion may never get to discuss the author's message. So identification of major themes should not concentrate on the students' personal views, while time needs to be allocated for that as well in order to give the students an opportunity to exercise commitment to values.

Application of Material to Other Works. To counteract an often prominent feature of learning - its fragmentation with isolated facts out of context - discussion groups put aside time to make a conscious effort to relate the reading to previous learning situations. This allows the group to be able to evaluate the relevance of what is being said. The purpose of an application is to take the arguments of one author and either refute or support them by cross-referencing them with another expert point of view.

In ethics discussions, science students are not likely to have many reference points on philosophical theory to draw from and to compare different expert opinions. What they have, on the other hand, is plentiful experience - either direct or indirect - of similar moral issues and of different ways of approaching and solving them. This would be the time to let students explore the level of identification between different moral issues they are aware of and to create a mental map of where the discussed problem touches on other similar problems and how these issues are possibly linked.

Application of the material to the self. Self-application encourages students to make the discussion personal. This is likely to increase the significance of the learning as long as the topic has some value to the students: it can deepen and challenge their understanding of themselves, their relationships, and the collectives to which they belong. The aim is not only to state one's personal opinions on moral issues, but to seek a contact point of the moral issues to one's personal life. If, for example, the issue is animal welfare in research, the students can reflect back to laboratory work they have carried out involving animals.

Evaluation of the author's presentation. This is an opportunity for the group to express their reaction to the reading - whether they felt frustrated, excited, bored or bewildered. The purpose of this part of the discussion is not fully served unless the students provide reasons for their feelings - 'I felt frustrated because the text had no structure, it was a lot easier after I wrote down the main points and then tried to read it again'. Emotional unloading is important, but it is even more important to rise above that and provide a well-considered appraisal of the theory, the logic, the method, and the conclusions of the author.

Evaluation of group and individual performances. This step is essential, but possibly the one that meets most resistance - the task of evaluating one's own performance and that of others. The facilitator of the group can assist the evaluation by asking questions (modified from Rabow *et al.*, 1994, p. 21):

1. How well do you feel we covered the subject?

2. What areas did we as a group agree upon? What did we disagree upon ?
   Do you feel everyone was heard and understood?

3. Were there questions that would have needed further clarification?

4. What areas of disagreement did we find that we could not agree upon in
   the end?

5. Who contributed greatly to the discussion? If you feel you did not
   contribute that well today, why do you think that is?

These questions should help students to express who and what helped them, who
and what was constructive and who and what inhibited the discussion. It will be
hard to confront non-functional and disruptive behaviour in the group, but that
must be done during the evaluation in order for the group to improve their ability
to work together. Likewise, supportive and co-operative behaviour should be
applauded and encouraged. One of the ground rules I set for the evaluation was
that students could directly criticise only themselves or me (the facilitator). This
was to pre-empt any possibility of personal attacks that would have been
detrimental to the group's future work.

According to Rabow *et al.* (1994) the times given for each step are better adhered
to quite strictly in the early stages of a discussion group. If this is done, all steps
will become part of the group members' expectations of the discussion. Later,
when all members have internalised the structure, the group will have more
freedom to alter the structure to cater for specific texts or the needs of the group.

7.4 Summary

Supporting student participation as independent moral agents requires a student-
centred approach to both teaching and learning. This means that the students have
an increasing responsibility for their learning together with increasing rights to
make decisions regarding their learning aims, methods and schedule. Student-
centred learning does not mean teacher-free learning, but it does change the
teacher's role from a fact-giver to a facilitator of student learning. Good

facilitators are not necessarily specialists in the subject, but they need to have enthusiasm and skill to ask questions that encourage student learning, know-how on resources and a desire to promote learning. Being a specialist gives the facilitator additional skills, because an expert facilitator is more able to ask the right questions and make sure that no important element has been dismissed by the students.

Two approaches to teaching ethics are discussed in this chapter; problem based learning (PBL), and structured discussion groups. A successful PBL exercise has a problem that catches students' interest, offers an opportunity to discuss and study all the issues relating to the case and is not be too broad to overwhelm or discourage the students. PBL offers a excellent approach to studying ethics.

Structured discussion groups can be organised around a problem similar to PBLs, but they also offer an opportunity for a more focused study of particular moral issues, for example animal rights. Discussion groups are not meant to be free-flowing exchanges of opinions, but structured assessment and analysis of the issues, usually based on reading completed before the session.

The teaching approaches considered in this chapter indicate that there are several interactive methods for teaching ethics, which should fall into the category of successful teaching approaches by allowing a strong student-centred approach and involvement of students as moral agents.

8. Teaching strategy

There are several different ways of introducing ethics into the science curriculum and the one detailed in this chapter is just one of many. The University Survey (see Appendix I - University Survey) carried out at the start of this research described three main approaches:

1. Module/course devoted to ethics, either as an option or a compulsory course.

2. Incorporation of ethics material into the subject matter of a science course i.e. ethics as integral rather than 'other'.

3. Informal discussions (tutorials).

The last of these options relies on the staff members' interest in ethics and usually no attempt is made to monitor the effectiveness of ethics discussions nor is ethics assessed. The second option has more structure and provides an opportunity to assess ethics as part of exam questions. The benefit of this approach is the close interaction between science and ethics which can highlight the role of ethics in good scientific practice. The difficulty of this approach is lack of ethical training of scientific staff and possible down-playing of ethics in comparison to the scientific content of the courses. The first option, an ethics course, provides an opportunity to approach ethics in a more structured fashion, but may not be as effective in highlighting the interaction of ethics and science for students. Ethics can remain as an isolated and disconnected element in the curriculum.

In IBLS, students are introduced to ethical thinking in several different ways. In level 1, all biological sciences students participate in two exercises which require them to consider scientific practices from the point of view of ethics: The Alien Squirrel exercise and the Cloning debate. The Alien Squirrel exercise has been found to have a significant impact on students values in environmental decision-making and to improve their self-reported group skills (Clarkeburn et al., 2000). In level 2, one of the options is a module in 'Science Communication', which covers in detail areas of scientific misconduct and integrity. Ethics is also discussed in tutorials.

This chapter outlines the research arrangement for teaching ethics to science undergraduates at the University of Glasgow during the 1999-2000 academic session. It provides a description of the practical arrangements for integration of ethics into the science curriculum. Three elements are considered in detail: 1) how time restrictions influence the choice of teaching styles, 2) how the ethics courses could be included in the general assessment, and why in most cases they weren't, and 3) how topics were chosen. This is followed by a description of the course contents. The chapter ends with a summary.

8.1 Time constraints and ethics in IBLS

In reality external factors limit the opportunities to create an ideal teaching approach to encourage moral development. Time is the greatest limiting factor in the design of an integrated undergraduate science ethics course. The time within the existing curriculum has been allocated before the introduction of ethics, and the introduction of ethics requires something else to be removed in order to make room for the new teaching. Very few elements in the current curriculum can be judged irrelevant or unessential (they have been included based on their importance) and the teaching staff is understandably reluctant to cut down on the current contact hours dedicated to their courses. Ideally then, ethics should be integrated into science curricula simultaneously with other major changes. General restructuring of the curriculum would allow ethics to be integrated without the laborious negotiations and compromises required when it is introduced into an already existing curriculum.

In the University of Glasgow's Institute of Biomedical and Life Sciences (IBLS), where this research took place, a decision was made to introduce ethics into the third year curriculum (L3). The reasons for choosing L3 as the research population were dominantly negative i.e. other levels were unsuitable for the following reasons:

1. L1 biology modules have a student population of 800+ and the smallest group teaching takes place in laboratories groups of 50 students.

Further the course structure was very rigid and provided only very limited opportunities to introduce even a minimal ethics component.

2. L2 contains a selection of core modules and elective modules, which form a structure where it would have been impracticable to maintain control and test groups.

3. In L4 students are partly involved with their research projects which do not provide access to them as a group, and partly with highly structured option courses where control and test groups could not have been established.

L3 students are divided between 14 different honours courses[1] studying more or less independently of each other. The choice was then made to treat each honours course as an independent component and design the ethics intervention for each course individually. Four of the L3 courses - Anatomy, Biochemistry and Medical Biochemistry, Neuroscience, and Biotechnology - were unable or unwilling to take part in the research.

Out of the remaining 10 courses 8 had ethics introduced as part of their existing tutorial scheme, while the remaining two (Parasitology and Microbiology, and Sport Science) were in a process of major curriculum changes, which allowed an introduction of a more extensive ethics course using the PBL format.

The PBL ethics courses consisted of 14 (Sport Science) to 30 (Parasitology and Microbiology) hours of study, out of which 4 and 10 respectively were contact hours. The tutorial scheme ethics teaching had an average of 5-10 hours of study out of which 3-6 were contact time. If more time had been available across the courses each course would have had either an ethics PBL or a more extensive discussion programme with a minimum of 5 meetings each including at least an hour of independent study. Unfortunately this was untenable with the restricted opportunities to arrange contact hours with the students and due to limitations set

---

[1] These course are: Anatomy, Aquatic Bioscience, Biochemistry and Medical Biochemistry, Biomedical Sciences, Biotechnology, Botany, Genetics and Molecular Biology, Immunology, Neuroscience, Parasitology and Microbiology, Pharmacology, Physiology, Sport Science, and Zoology.

by the course co-ordinators on the amount of hours the students could be expected to study ethical issues independently. Even though ethics was considered important by all course co-ordinators, it was still in most cases considered a threat to the time students spent on studying science. This attitude was in part because the new ethics course was at this stage experimental and unlikely in most cases to form part of the assessment. Thus only limited time was made available for ethics teaching.

Alternative points for contact of ethics were significantly reduced by the decision to have control groups within each course, as ethics could not be introduced in small snippets into existing laboratory work, which could have otherwise offered an optional entry for ethics into the science curriculum. This might be a preferable introduction method if ethics is introduced without additional research requirements.

The time available for the ethics teaching was less than what was aimed for, based on previous research which indicates that in order to produce significant acceleration and gain in moral development, the teaching programme should be no shorter than 4 weeks and no longer than 12. The teaching within the tutorial schemes falls short of these guidelines, while the PBL courses were within the medium length programme limits which was found to produce the highest moral development gains (Schlaefli *et al.*, 1985).

The lack of contact time in the tutorial scheme was partly compensated by giving students reading before every contact session, which increases the time they spend on actively considering ethical issues and extended the time period when ethical issues are part of their study programme. For similar purposes the tutorial sessions were held on average two weeks apart, in some cases even in different terms to increase the time period when students are expected to keep ethical issues in their minds.

8.2 Programme design in IBLS

The three session tutorial discussions were organised so that the discussion in the first session concentrated on moral sensitivity issues by presenting the students with a situation containing plentiful moral aspects and encouraging them to identify these. The second session concentrated on moral decision-making tools and the third session on personal moral values and how they affect the entire decision-making process. The details of time allocation and themes in each L3 course are given in Appendix XII - L3 Structured Discussion Programme'.

In the PBL programmes, students were given problems that have several ethical elements, though these elements were not apparent in the problem outline. The problems were designed so that successful solution cannot be found without ethical considerations. The PBL exercises therefore required students to exercise their moral sensitivity skills as well as reflect on their personal values. See Appendix XIII - PBL for a detailed account on these exercises.

One of the problems in motivating students to take an interest in ethics is the lack of opportunities to assess the introduced ethics components. Based on anecdotal evidence, students seem to have a tendency to take an interest only in the course elements that directly contribute towards their final marks. All other elements are considered less important and removed from the priority list first. The lack of assessment limits the external motivation for students to study ethics. This can be viewed as an opportunity as well as a problem. When students discuss ethics without assessment pressures, the interest thus sparked is internal and possibly better retained in the future. The problem remains that possibly only small numbers of students gain a personal interest in the subject. The lack of assessment also has further negative implications. Teaching without assessment fails to highlight the importance of ethics as an important part of the curriculum and of being a scientist. Without assessment ethics is easily viewed as an additional, not integral, part of the curriculum, something that is not really important.

Assessment could not be introduced for most of the L3 ethics sessions, at least in this research, for several reasons. The three main ones were:

1. Lack of time to carry out internal quality assurance procedures for assessing a new piece of teaching.

2. Problem with equality if different assessment methods were created for students participating in the ethics teaching and those who didn't. This might have caused unacceptable inequality between students.

3. Tutorial schemes are a non-assessed part of the curriculum, at least in some degree courses.

As a result, Immunology was the only course with a discussion programme which included assessment, in the form of an optional question in the term 2 class test. The PBL based ethics courses on the other hand were included in the assessment, which was purely skill-based so that the same assessment could apply to both the research and test groups (Details in Appendix XIII - PBL).

8.3 Facilitator

In both PBL and structured discussions the facilitator can influence the learning process greatly - both positively and negatively. For research purposes all ethics teaching for this research was carried out by a single facilitator (myself) with co-facilitators for assisting with large groups. This resulted in a very substantial work load, but also a wonderful opportunity to experience the learning process first hand.

8.4 Choice of topics

One of the important elements in ethics teaching is to choose the study material so that it provides the students with the most relevant moral problems, which are focused and structured to provide a good opportunity for learning and discussion. One of the core assumptions in this research is that ethics should be integrated as far as possible with the existing science curriculum so that ethics does not stand

out as a separate issue from science, but links with it at as many points as possible. These links are assumed to provide a useful parallel for future recognition of moral problems in science and thus increase the usefulness of the ethics teaching, as well as providing motivation for students to study them in the first place.

The topics for each L3 course were chosen in collaboration with the course co-ordinators. As a starting point, the co-ordinators were given a list of three possible categories of moral problems within which the most relevant teaching was assumed to fall. These categories were:

1. Research and Professional Ethics

   - Applied science - distribution of benefits, control and responsibility
   - Values in science - are there issues we should not study? Is science really neutral?
   - Fraud and scientific integrity.
   - Animals in research
   - Research design and testing hypotheses - risk assessment, autonomy of research subjects and benefits of research

2. Environmental Ethics

   - Importance of diversity/natural environment
   - Ecology vs. Economy - whose interests?

3. Genetic Ethics

   - Genetic engineering - nature vs. nurture, risk assessment, distribution of benefits
   - genetic testing/screening - what, who and when?
   - changes in the gene-pool - whose duties and rights?

Based on this list I met with each course co-ordinator to discuss both the possibility of introducing ethics to their L3 course and to find out what were the topics they believed to be the most important in their field and which they subsequently believed the students to both have interest in and to benefit most from. In these discussions, all course co-ordinators suggested moral issues

relating to their scientific discipline from more than one category, while all suggested at least one moral theme from the first category of research and professional ethics.

The themes were then further discussed with the L3 students in spring 1999. The issue of animal use in research rose above all others in student discussions. This interest and other student suggestions were included in the list of topics. The students interests were very similar to those of their course co-ordinators and the student discussions gave further support that the chosen topics would provide the learning base they were designed to provide - genuine interest and scientific relevance for the L3 students.

The course material development that followed involved two main tasks: finding/writing suitable reading material to accompany the sessions and design of the course outline. Using pre-existing material from a variety of sources was preferred to writing material specifically for the ethics courses in order to maintain as wide a perspective on the ethical issues as possible, and to highlight the interdependent links between ethics and science. The preliminary materials were then discussed again with the course co-ordinators and some of them were tried out with volunteer student groups. Student hand-outs were developed in the last stage of course development.

8.5 Summary

The restrictions placed on ethics teaching by both structural and organisational matters reduced the ethics teaching to less than it ideally should be. To compensate for this lack of contact time the exercises had independent study time included. This independent study increased the active time ethical issues were considered and with spacing the contact and independent study periods appropriately the ethics teaching extended to match the lower limits of successful ethics programme design. The organisation and structural restrictions also dictated the choice of teaching approaches. With limited hours for both contact and independent study PBLs were not even an option for most L3 courses.

The guideline for choosing discussion and PBL topics for the ethics teaching was their relevance and ethical complexity. During the choosing process students and staff members were widely consulted.

PART IV - Results

The previous Parts and chapters have described the theory behind and
preparation for the ethics intervention at Glasgow University in 1999-2000. This
Part details the outcomes and results of that intervention.

There are three main types of results:

1. Students' self-reported learning gains and benefits of ethics teaching.
2. Analyses of learning logs.
3. Results from the Moral Development Questionnaire.

It is important to look at all of these outcome types in context. Whether or not we
can objectively say that students have 'learnt' something during the ethics
teaching, it is important for two reasons to find out what they themselves believe
they have learnt. First, the Moral Development Questionnaire may not detect all
types of learning and benefits students gain from ethics teaching. The students'
self-reported learning can thus provide an important insight into the learning
experience. Second, student experience is an important factor in motivation,
which in turn influences learning. If the students find ethics teaching exciting,
challenging, and rewarding, they are more likely to develop a growing interest in
ethics, which facilitates their moral development in the longer term.

The first chapter in Part IV describes the use of the teaching material developed
for the ethics courses and analyses the student responses collected in conjunction
with the teaching. Chapter 10 details the analyses carried out on students'
learning logs.

The results from the Moral Development Questionnaire can stand to support the
teaching approach chosen. They can provide objective data about students' moral
development. The Moral Development Questionnaire has three distinct parts,
which are in chapter 11 analysed separately first and then comparisons between
the parts are made.

The Moral Development Questionnaire as a measure may not be able to detect the smallest of changes in moral development due to two factors:

1. The Questionnaire (in particular DIT) is developed to measure long-term development and the small changes resulting from a minimal ethics intervention may not be detectable, even though they may function as a starting point for further development.

2. The changes to be expected are relatively small, and they may be masked by the general noise in the results.

Chapter 11 describes the Moral Development Questionnaire results in the same order as the different elements appeared in the questionnaire itself:

1. Moral Sensitivity
2. DIT
3. Perry
4. Comparisons between parts

In this research, even though it was not possible to introduce ethics into the Level 1 biology curriculum, L1 students filled in the Moral Development Questionnaire at the start of the academic year. These data provide us with an opportunity to study natural moral development during the early academic years. These results are discussed in chapter 12.

9. Course material and student responses

The application of the Moral Development Questionnaire gives out of context appraisal on the impact of ethics exercises. It measures how students' thinking in general is affected by this new teaching element. Another way to look at the success of ethics exercises is to ask the students themselves what they have learnt and how they see the benefits of ethics teaching.

In order to collect this information, I asked students to write answers to three questions at each ethics exercise session. The first one was to gauge their views - partially at least - based on the pre-exercise reading. An additional benefit was to focus student attention on the subject of each exercise at the start of the session. This was done before any discussion had taken place. Then, right at the end of each session students were asked to write what they thought was the main learning gain of that exercise, and what was still unclear. These answers could then be compared with the set learning objectives for each exercise and the inherent components of each ethical theme discussed.

These written responses were collected anonymously. This was done to encourage honesty in the student responses. When asked, the students did say that having to put their name on the paper most possibly would have changed what they wrote. The down-side of anonymous responses is the inability to compare individual students' responses to different exercises, and whether enthusiasm towards the exercises is reflected in the moral development scores. So honesty was the trade-off for some research benefits. Was I to do this again, I would probably encourage students to write their names on the forms, though emphasising that they should not do so, if they would answer any differently because of it.

The student hand-outs can be found in the appendices (XIV-XX). All the reading material was photocopied for the students and the material was in most cases given approximately a week in advance. All that students knew about the

sessions was written on the introductory page of their hand-out. They were thus not given a list of objectives for each course, but rather a set of general aims.

This chapter consists of seven sections, each describing one discussion exercise. In each section, the motivation for the exercise is discussed first, and that is followed by a brief outline of how the discussions progressed and were facilitated. The second half of each section is an analysis of the written responses. Each question is looked at separately and a short appraisal of the results is given. Each section has got its own short summary, and the material and student responses are in general summarised at the end of this chapter.

9.1 Animals in scientific research

18 groups of 8-14 L3 students participated in the 2-hour 'Animals in scientific research' exercise during first and second term in the academic year of 1999-2000. This was the first session in all ethics programmes, with the exception of Botany. I chose to start with this theme for two reasons:

1. The ethical issues are apparent.

2. Most students have views on the subject and it has direct implications on their life as students and their career choices in life sciences.

For these reasons the 'Animals in scientific research' exercise has direct relevance to the students and getting started with the discussion is not dependent on highly developed moral sensitivity.

Prior to the session students were given two short readings; an introductory text on moral theories (see appendix 'Animals in scientific research') and either an extract from Katz's (1997) book on 'Nature as subject' which gives a controversial view on animal testing or a short descriptive paper on the issues for and against animal testing which I prepared specifically for this purpose. I also prepared the introductory text mainly as reference material for students reading Katz's paper, which for most of them was their first attempt to understand philosophical writing. The purpose of this paper later evolved to be a short introduction to different decision-making methods students have at their disposal. Though the paper provides insight into two quite complicated philosophical decision-making methods (deontological and utilitarian), the students reacted to it positively. My impression was that they had only in very limited terms ever previously considered the methods they use in moral decision-making and the explicit explanation of two alternative, and often conflicting, methods functioned as a catalyst for exploring further alternatives. The more experience I had with this exercise the more thoroughly I tried to challenge the students to consider the usefulness of these theories and to explore alternative ones.

Katz's text was not only controversial (suggesting that animals are human artefacts and thus we do not need to care about the pain we cause them) but also demanding in its use of philosophical terminology and writing style. It was chosen because it covered all the elements in the debate and provided a good starting point for a discussion by not shying away from making a point. But students found it sometimes too demanding and discouraging to read. I believe in challenging the students, but I realise that the message is wasted if the students as a result do not complete the reading. So as a response to student criticism I wrote a simpler, shorter and less controversial text covering the same issues. Students were less frustrated with the modified text, but also less enthused to discuss the issues. Were I to start again now, I would probably either keep the Katz paper or write a new one which includes the controversial element of Katz's paper, but one addressing the issues in a less complicated format. Despite the change of reading material, the session outline remained the same, though more was required of the facilitator when the material did not in itself provoke student reaction.

Each session started with students writing a brief (maximum 2 minutes) answer to the question: '*What do you think is the most important issue of animal rights and animal welfare?*'. There were two reasons for this: 1) to focus student attention on the subject at the start of the session, and 2) to collect data on their priorities regarding animal rights and welfare. Later in the session, when different points were considered, students were encouraged to share their viewpoints with the group.

Short introductions followed the written task. Students were asked to give their name and say how they felt about the reading. I encouraged honesty in stating their feelings, and honesty I got. Students were straightforward with their frustration, boredom and/or interest in the text. Katz's paper unsurprisingly provoked more frustration in both his style and the views he put forward, but most students had still found it interesting. The tailor-made paper received a lot less emotional response, though the level of interest was maintained.

The introductions were an important ice-breaker in the first of the ethics sessions. They set an informal and open atmosphere, where students' views were appreciated and participation encouraged. At least that is my interpretation of their impact.

The vocabulary clarifications which followed were an opportunity to bring up some of the philosophical issues. For Katz's paper, students wanted clarifications of the following terms: normative ethics, intrinsic, empirical, *reductio ad absurdum*, holism, utilitarianism, and deontological. In most groups, students themselves asked for further clarification for deontological theory and utilitarianism, and if they didn't I asked them to explain in their words how they had understood them. More often than I would have liked, I found myself lecturing at this point, students shying away from the discussion. Not all groups joined the discussion, but with the groups that actively contributed to the vocabulary clarification, the discussion resulted in very fruitful exchanges and student enthusiasm was palpable.

Vocabulary out of the way, the group listed all the main issues they had found in the text: speciesism, pain and animal welfare, and when can we justify pain. Speciesism involves a complex argument about our inability to differentiate humans and non-human animals. The core of the argument is that if we cannot define humanity in any other terms than our genetic make-up, we are making moral judgements based on irrelevant characteristics not dissimilar to racism and sexism. First the students were asked to suggest ways to describe humans in order to distinguish humans from other animals. Their answers were along the lines of intelligence, consciousness, and communications skills. Then we explored whether other animals would have these skills, and in most cases the agreement was that some of them do, at least to a degree. The line was blurred further when the discussion turned to find out whether all humans possess these qualities and the agreement arose that not all do. The students had a strong intuitive feeling that people are more important than animals, but struggled to find reasons why. We then looked at the issue from another perspective by asking ourselves what would happen if DNA was allowed to be a decisive

characteristic when making moral judgements. Most groups quickly identified problems of sexism and racism to follow from this, though on many an occasion the consequences of accepting this type of reasoning needed to be outlined to them. At this point many groups discussed the importance of consistency of moral rules. The agreement was that this is a strong argument against use of animals in testing for safety and efficacy of new compounds, though it went against their intuitive judgement for the higher importance of humans. During several sessions that strength of the intuitive feeling for the superiority of humans persisted despite an inability to support it. The discussion that followed was in many groups partly undermined with the inner-feeling 'that humans are always more important'.

The groups then discussed the role and importance of pain. Most groups quickly established that animals can feel pain and that not all pain is bad. Good pain was described as something that protects the individual and/or teaches something valuable. Almost as quickly they agreed that pain endured by a laboratory animal is rarely, if ever, good pain from the point of view of *that* animal, though the pain might benefit other individuals. All groups also agreed that pain experienced by animals counts for 'something'. At this point I often encouraged the students to try to establish a cut-off point for which animals are included in their moral considerations. To illustrate my point I would ask how they would react if someone pulled the tail of a cat really hard for no apparent reason. Most thought it would be wrong and they would try to stop it. Then I suggested that they'd see someone poke a mouse with a hot iron, again for no apparent reason. Same response. Further, what if someone pulled the wings off a lady bug, and again students objected. Further I suggested putting needles into a tapeworm and students agreed that that should not be done for fun and entertainment. This then established that there is a default schema of not hurting and that we need special justification for performing painful tests on animals.

The natural progression from here was to discuss in what circumstances we could justify animal experiments. Most groups came up with the division between cosmetics and medical research first. Use of animals in cosmetic

research was condemned because make-up is not essential or necessary. To further define 'essential' or 'necessary' I encouraged the students to consider whether it would make a difference if 1) the number of people that would benefit would be 2 or 10 000, 2) whether the condition was fatal and a cure could be found, or whether successful research would 'only' improve quality of life, and 3) whether the research would improve quality of life or survival chances by 1% or 50%. In all groups students needed to be explicitly encouraged to consider different sources of criteria, but once the issues were on the table the discussion flowed freely and enthusiastically. The common trend was to start with a very limited definition of 'necessary' which then was broadened as students explored their feelings by applying the situation to themselves, 'what if that one person was my mum?' or 'what if it was my quality of life?'.

At this point I introduced utilitarian calculus again, and asked them to think of this decision-making as an exercise in finding a positive balance between the level of harm and the level of benefit a particular animal experiment would produce. At this point the group was clear on the side of benefits and needed to think of what would need to be considered on the side of harm. Students discussed whether it mattered what animals were used: often they thought it did - it was harder to imagine an acceptable test with primates than with flies. Number of animals was included on the side of harm together with the level of pain the animals would have to endure. The latter point often raised questions on our ability to know in how much pain the animals are. Further, many groups included the harm that would result if we did *not* carry out the experiment.

In evaluating the strength of the elements in the harm and benefit calculus I often asked the students to consider how they would determine their confidence in the chances of the research and how that would influence their evaluation of the elements. First they listed things that would give them confidence: previous research and previous experiments. Then they considered if different confidence levels (e.g. 20% or 80%) that an experiment would be successful (would result in a cure for example) would influence their decision how 'necessary' particular research is. Again most students were more stringent first but the more they

thought about it, the less important high confidence levels became. Unsurprisingly none of the groups came up with a definite formula they would follow, but that was not the aim of the exercise. Regardless of the confidence levels or the balance of benefits and harms, most groups agreed that research should always be designed to involve the least amount of animals possible; that animals should be well cared for, and that animals should not be used unless that was the only way to get the results.

To tie the theoretical discussion together with reality, students were then asked to offer their personal views on animal testing, how they had felt about the use of animals during their course, how much these issues would influence their future career choices, and whether they consider animal research issues when making consumer choices. Discussions on the use of animals in teaching were in most instances lively, with some students rejecting them strongly and others defending their inclusion for learning benefits. Also students had strong views on whether they would be ready to do animal research themselves even if they accepted it and saw the possible hypocrisy embedded in their conclusions.

The sessions were wrapped up by asking all students to briefly describe what they had learnt during the session, if anything, and how they would improve the sessions in the future. Only a negligible number of students said they had not learnt anything or that they found the sessions uninteresting. Improvement suggestions were of two kinds; practical and self-critical. Practical suggestions often included a round table to sit around, a warmer room or having the reading earlier. Self-critical comments pointed out the importance of reading the material next time, or to participate more/less. In cases when it was possible to change the session according to students' comments, it was always done. Before leaving, the students were asked to write down answers to two further questions: 'What is the big point you learned from the discussion today?' and 'What is the main unanswered question you leave the discussion with today?'. Further, students were asked to give themselves, the group and the facilitator a score out of five.

I was very happy with the session. It allowed students to discover and discuss important ethical issues in context. Still, I can think of at least two ways of improving the session.

1. It would be helpful to have some visual aids at your disposal during the sessions, in particular when encouraging the students to define what animals are included in their realm of moral consideration and how to balance benefit and harm. Black/white board or OHP would suffice.
2. With quiet groups short case studies that could be quickly completed in pairs or small groups could provide a method of overcoming shyness.

9.1.1 Analysing student responses to 'Animals in scientific research exercise' - qualitative material

Q1 What do you think is the most important issue of animal rights and animal welfare?

There were three themes in the students' answers: pain and suffering, moral status of animals, and testing protocols. In all three themes, a further three sub-themes can be found: opinions/statements, questions, and deliberations. To some extent these three sub-themes can be placed in a developmental order where the lowest level would be opinions, then questions, and deliberations as the highest level. This is similar to the qualitative analysis structure of the moral development questionnaire (see section 5.3.2.2).

A breakdown of student responses is shown in Table 13.

	Opinions/Statements	Questions	Deliberation
Suffering	1 Unnecessary pain and suffering 43%	2 Whether animal is in pain due to the experiments? 16%	3 Is the suffering of the animal justified? 19%
Moral status of animals	4 Animals have rights 3%	5 What moral rights should we give to animals? 8%	6 The problem of defining the difference between animals and humans? Speciesism? 6%
Testing protocols	7 The exploitation of animals for profit 9%	8 Are there alternatives to animal testing? 12%	9 What is necessary, where do we draw the line? 5%

Table 13: Animals in scientific research, Question One

The percentages in the table are based on the percentage of students identifying each element and as some students identified more than one element, the sum of the figures is more than 100. The analysis is based on 146 student responses. In addition to these, there were three responses that did not fit into the qualitative key, all of which stated personal views unrelated to animal testing and these were removed from the analysis.

Pain was most commonly cited as the most important element with 78% of students including considerations of suffering in their response. I interpret this as a positive sign of student sensitivity to animal pain, which in turn can be interpreted as evidence of empathy. The 17% of students referring to the status of animals is best understood as evidence that students have read the material and it has given them food for thought. Statements and questions regarding testing

protocols are the least cohesive group in the analysis, though again the student responses show clear awareness of relevant ethical issues.

The shaded boxes indicate sophisticated ethical questions. Having 38% of students asking these questions before the sessions is a good start to an ethical discussion. It can also be taken as some evidence that the reading facilitates student thinking.

Q2 What is the big point you learned from the discussion today?

In the student responses I could identify four distinct categories: elaborations on personal relationship with the session content, interest in and benefits from the group discussions, views on moral answers and animal testing, and factual statements of different kinds. All four categories could be roughly divided into three different sub-groups: stating an experience, evaluating the learning or asking questions, and projecting the learning further and often taking a strong stance. This was then worked into the 4x3 table below (Table 14).

	Experience/statements	Evaluation/Questions	Projection
My view	1 Realisation of the complexity 35%	2 can/can't justify 5%	3 Need to think more, to have an open mind. New methods. Sense of ability. 20%
Interest and involvement	4 Good to talk. Hearing views of others 20%	5 Awareness promotion, learnt something from views of others 11%	6 We need to be able to justify our views. 4%
Morals of animal testing	7 no moral right or wrong: a) strong; b) weak a) 9% b) 13%	8 What gives us a right to choose? Alternatives must be sought. 6%	9 is does not equal 'ought'. 7%
Factual	10 My view 3%	11 When can we justify pain? 2%	12 How do you know that an animal is in pain? 2%

Table 14: Animals in scientific research, Question Two

Again the percentages in the table represent the percentage of students reporting a particular element and as some students listed more than one learning element, the percentages do not add up to one hundred. Some students had to leave the session before completing the form for other commitments and some left the question unanswered, and thus the number of responses was only 123.

My objectives for this session were:

1. To increase students' understanding in and sensitivity to the ethical problems in using animals in research.

2. To encourage them to develop and use their moral decision-making skills in general and in particular in relation to the animal testing issue.

The lighter grey squares (1 and 5) represent student responses where their self-stated learning fulfils my first objective and darker grey squares (2, 3, 6, 8, 9, and 11) represent my second learning objective. These data are very encouraging as 46% of students have learnt what was stated in the first objective and 36% of students fitted their responses with the second objective. Altogether, 101 students (82%) reported learning in at least one of the objectives. Only one student reported to have learnt nothing. This is a highly satisfactory outcome and a promising sign that the session design suits the subject matter and supports students' moral development. A further 20% of the students explicitly report their satisfaction in the discussion-style of the session (response 4).

Problems with relativism recurred time after time in the sessions. Either students felt frustrated with the lack of clear-cut answers or they felt that it justified their lack of care and interest in the subject. Therefore it was not surprising to have 22% of students stating relativism in their response forms. The weak form ("Perhaps there is no right or wrong.") can be identified with the realisation of reality in this situation, while the strong relativistic stance ("You cannot moralise using animals for research at all.") is better understood as a protective shield against the need to make decisions or have involvement. The strong form of relativism is in accordance with the Perry theory and the results found in the Perry questionnaire. Some students are still in Perry B level (see Appendix III for Perry scheme and chapter 12.3 for the results), while most have progressed beyond escapist relativism.

Q3 What is the main unanswered question you leave the discussion with today?

The motivation to ask this question was to collect data on the students' view on the main unanswered questions, which would reflect both whether the session had omitted some crucial element and what thoughts the students left the sessions with. Three distinct categories of responses could be found: 1) students

queried the rights to use animals in research, 2) they were still unsure how to define the status of animals, and 3) they elaborated on the decision-making issues relating to animal testing. These three categories could then be divided further into three sub-groups: 1) questions, often in a form 'should we...', 2) search for justification of actions, and 3) projecting their thoughts into future actions.

	Questions, should we?	Justification	Projection, who should
What are our rights?	1 Should we use animals? 1b personal account a) 20% b) 6%	2 Can we justify our use of animals? 2b personal account a) 16% b) 5%	3 Who should decide what we can do? 5%
Status of animals?	4 What is the status of animals? Marginal humans? Where do we draw the line? 23%	5 How can we define the status of animals? 6%	6
Our actions	7 Who decides? Law? 4%	8 How people decide? 11%	9 We can't solve this! Are there any fundamental rights or wrongs? 6%

Table 15: Animals in scientific research, Question Three

138 students responded to this last question and again the percentages represent the number of students raising each element and as some students raised more than one element, the percentages do not add up to 100. The motivation for this question was to find out what sorts of questions students are left with. The result is interpreted as positive if the students are left asking central and complex questions about the discussion theme. In the response matrix I consider questions

1-5 to fulfil this criterion. This means that 80% of the students left the discussion with an important question to consider. This is promising as it indicates that student learning may not end together with the session.

Students also gave a score for themselves, the group and the facilitator on how well they thought each had done during the session. The students' self-assessment average was 2.9, their group appreciation 3.8 and the facilitator average of 4.3; all out of five.

9.1.2 Summary

The 'Animals in scientific research' session was successful. The sessions provided students with a relevant discussion topic that invited both philosophical and personal considerations. The qualitative analysis of the results gives confidence that the students are learning the issues the exercise was designed to teach and that their learning is not confined to the session alone, but that they leave the session with important and challenging questions.

9.2 Ecological decision-making - DDT/Malaria

This teaching unit was prepared with Zoology and Aquatic Bioscience students in mind. The Zoology and Aquatic Bioscience degree courses not only concentrate on animal morphology and physiology, but also on the interaction between animals, humans and the environment. Conservation issues are naturally included in the curriculum and ethical issues are inherent in many of these concerns. The attempt was to design an ethics exercise around a topic that would link directly with these conservation questions addressed in other parts of the course.

The use of DDT in malaria control provided a good case. DDT has been found harmful to the environment (or has it, as the paper suggests), but it is successfully used in warmer climates as an affordable malaria control. In making their decision on whether DDT should in fact be banned in 2003, as suggested by the UN, the students need to consider several ethical and scientific issues: What is adequate scientific evidence, what is 'safe' decision-making concerning the environment and human welfare, what risks are we ready to accept and for what returns; who has a right to make decisions, and what is our responsibility to help other people? The issues are fundamental, complex, and interesting.

In term 43 (four groups) Zoology and Aquatic Bioscience students took part in the exercise. In my view, this was the most difficult exercise in the course package I developed. None of the problems seemed apparent to the students. They failed to see the ethical concerns in this issue and thus they were not forthcoming in the discussion. After the first laborious group session, I was better prepared for the last three. The discussions improved significantly as long as I was able to guide students to recognise the issues. The description of the session I present here is the one I devised after the first session and thus incorporates my learning experience from the less successful first attempt.

The pre-session reading was Curtis (1994) - a scientific paper with a significant section of scientific data, but also clear considerations on alternatives and the

social/ethical issues involved in DDT use as a malaria control. The session started with a short written comment on the paper and then a quick round with all students having an opportunity to express their feelings and thoughts about the paper. The article had been well received: many students had struggled with the statistical information, but all in all, they seemed pleased as the paper had been informative and it tied in with other knowledge they already had. During the vocabulary clarification the students mainly queried scientific terms: anophelines, lachrymation, paraesthesia, latency, corroboration, excito-repellency, and exophilic. The scientific dictionary was put to good use! I then asked what they thought was 'safe' as one of the key points in the paper was the consideration of safety in DDT use. Students suggested the extent of damage that results if things went wrong, how permanent that damage would be, what is the risk of damage taking place, how prone is it to human error, etc. The group then applied these criteria to controlled DDT use in malaria areas. The evidence they could find was not convincing either way. I then emphasised that in very few instances will scientific data provide decision-makers with data that beyond-doubt would state that something was 'safe', but in fact most safety related decisions are social/ethical judgements on the acceptable levels of risk.

The group then moved to discuss alternatives to DDT, with a hope of finding a suitable and more safe one to replace DDT which has a doubtful safety record. The only viable one they agreed upon, after considering anti-malarial drugs, vaccines, improved sanitation, and air-conditioning, was the use of pyrethroid impregnated bed-nets. The worry with pyrethroids was evidence of resistance development in malaria vectors and high starting cost with the purchase of bed-nets. Students also stressed the importance of research for a malarial vaccine, but realised that it could not solve the immediate problem of malaria control. Costs of both control-methods and of research was thus introduced to the discussion.

Students discussed in length their role as scientists in the future and whether they could have power to influence decision-making. Quite sadly, most groups ended up by concluding that they had no power. They also discussed how much we could expect people in Britain to devote public money in search for a cure for a

disease that does not affect the local population. This also included personal reflection on how much *they* would be willing to pay to see malaria control improved.

The discussion then progressed to consider who are the people who should (in an ideal situation) make decisions on banning DDT, or on the use of resources to find or fund alternatives. One of the crucial questions was whether Western countries have a right to ban DDT use as a malarial control, if its use cannot be proven to have a negative influence and it has a positive impact on the health of the people in developing countries. An analogy was drawn that if we have a right to do that, then we might have to grant the developing countries a right to curb our consumer choices that have global effects.

We finished the session with a vote on how many students would, knowing what they know now, be ready to ban all use of DDT in 2003. For the two groups that I collected the data, 14 voted for no ban, 1 voted to ban and 6 were non-committal.

This session was not easy, neither for the facilitator nor for the students. The issues were demanding, requiring the facilitator to take a more definite role in leading the discussion and poor preparation resulted in a very poor session. It would be important for the facilitator to start the sessions with a clear discussion format in mind, one that could be abandoned if students found an alternative issue of interest. I would be ready even to suggest that the discussion questions were given to the students more explicitly before-hand. The student hand-out (Appendix XV - Ecological decision-making - DDT/Malaria) does propose some questions, but they are probably too broad to truly guide students to consider the deeper issues when reading the article. I am still happy with the use of the scientific article, as it provides an additional benefit for students finding themselves able to form opinions about scientific data independently and discussing them in groups.

The positive experiences with the later groups actively participating in the discussions and most clearly finding them challenging gives me confidence to suggest that this exercise in its current form was a successful one.

9.2.1 Student responses to 'DDT/Malaria' exercise - qualitative material

Before the group had discussed anything, they were asked to write a short answer to the following question:

Q1 What do you think is the most difficult issue in Malaria control?

There were two reasons for this question. First, to help the students focus on the subject right from the start. Many of them took the reading material out and had a look back before answering the question. Second, I was interested to find out what the students had picked up from the reading, to get an idea how much it provoked their thinking.

Two basic categories emerged from the students' answers: damage control and how to balance harm and benefit. Both of these categories could be further divided into three sub-categories: questions, problems with data, and deliberations on the data. The deliberations where clearly more sophisticated than the other two sub-category types. The percentages in the table below represent the percentage of students (n=43) who raised a particular issue, and because some students contributed to more than one category, the percentages do not add up to one hundred. There were no responses (n=44) that could not be fitted into the scoring key.

	How and what?	Problems with Data	Deliberation
Damage control	1 How to kill vectors with no damage to environment? 44%	2 The lack of data to make a good decision. 16%	3 Should we cause damage to environment when trying to control malaria? 12%
Balancing risks and harm	4 Weighing up cost and potential harm OR balancing cost and benefit. 16%	5 Do we know enough to make a decision? 2%	6 Are the side-effects worth facing? What is more important? 12%

Table 16: DDT/Malaria, Question One

Already the fact that all students raised one of these points as their most difficult issue indicates that they actually perceive some of the problem. This is an important indication that the reading does address the issues it was chosen to address and that students do recognise them while reading the text. Naturally it is also possible that students had formed their problem-web on the DDT/malaria case before this exercise. Judging from the anecdotal evidence from students, most students were in fact new to the issue and the reading had been a revelation to them. It is thus surprising that during the discussions they seemed so oblivious to these ethical concerns. One interpretation of this discrepancy is that the students recognised one or another ethical element, but this did not provoke them to truly investigate the big picture, but that they quickly moved on to make their own judgement on the matter, based on the information they had. This is supported by the discussion experience, where students from early on offered their solution to the problem (ban/no ban) and explicitly indicated that they had really only considered the issue from one point of view. Thus reading seems to offer a good starting point to the discussion, but on its own, it would not provide sufficient provocation for students to consider the issue from several angles.

After the session was over, students wrote short answers to two questions. The first one was:

Q2 What was the main point you learned from the discussion today?

The main reason for asking this question was to get student feed-back on the learning gains from the session. I had one main objective in mind when designing the exercise: to raise student awareness on the complexities in ecological decision-making.

The student responses could be divided into three main categories:

1. Straight-forward decision-making considerations.
2. Reflections on the discussion.
3. Factual issues and questions about DDT.

These could be further divided into three sub-categories:

1. Describing the current state of affairs, or asking questions.
2. Reflecting on the complexity of the issue.
3. Explicit accounts of learning.

In the table below the percentages are again percentages of students reporting any of the learning gains and as some students (n=43) reported more than one learning element, the percentages do not add up to one hundred. All responses (n= 48) could be analysed by using this key. To clarify the issues, the table includes short sample answers for each category.

	Current state of affairs	Complexity	Learning
Decision-making	1 Elements in decision-making - 'It is important to balance pros and cons when making a decision.' 21%	2 How complex the issue is -'decisions about a topic like this are extremely hard to make because everything has a knock on effect.' 40%	3 Decisions need to be made - 'moral judgements need to be based on current knowledge, they cannot be deferred.' 12%
Discussion and Opinions	4 Describing reality - 'sometimes we are more concerned with our environment rather than people dying, who we don't know and never will.' 14%	5 Describing the nature of discussion - 'to be aware of other people's point of view'. 9%	6 Reflecting on the importance of discussion -'my knowledge was poor before and now I have actually changed my mind.' 2%
DDT	7 Should DDT be banned? 2%	8 How good are the alternatives? - 'complete ban would not be ideal. unless suitable alternatives were available'. 9%	9 Factual learning - 'evidence suggests that DDT has not caused as much damage as originally thought'. 2%

Table 17: DDT/malaria Question Two

The highlighted squares indicate a match between the set objectives and students' self-reported learning gains. Altogether 26 students out of 43 (60%) indicated at least one of these as their main learning gain. The students reporting learning in relation to the discussion (categories 5 and 6) is also interpreted positively because they fulfil part of the overall aims of the curriculum: to provide an opportunity for discussion and learning that will support moral development.

I interpret this result as encouraging, particularly when keeping in mind how difficult the students found the topic during the discussion. The directly reported student-learning gains may be less than in some other exercises (animals in scientific research for example), but they are still high. Also, it is worth noting

that no student stated a learning gain that would have been irrelevant to the discussion or its objectives, though not all of them corresponded as accurately with the stated objectives.

The final question answered by the students was:

Q3 What is the main unanswered question you leave the discussion with today?

Most complex moral issues do not have singular right answers, and further, the group could not find any one particular answer that they would have been confident with. Asking students what is the question in their mind as they leave the session is therefore a way to gauge their ability to identify the key element that could not be answered during the session and which possibly could not be truly answered at all. Students asking important and fundamental questions would be a positive indication that the session has set their minds in motion.

Two main themes, both with three sub-themes, emerged from the student responses. The main themes were:
1. Decision-making - who and how.
2. Questions (often factual) about DDT and its usage.

The table below summarises the student responses. Once again the percentages reflect students (n=43) asking a particular question, and as some students asked more than one, the percentages do not add up to 100. All responses (n=47) fit the designed key.

	Action	Complexity	Methods
Decision-making	1 Who should make the decisions? 28%	2 How to deal with conflicting information? 9%	3 What are the appropriate decision-making methods? 23%
DDT	4 Should it be banned? 16%	5 Are the alternatives really viable? 25%	6 Factual and scientific questions. 7%

Table 18: DDT/Malaria, Question Three

The questions about decision-making are without a doubt important and fundamental in conservation. 26 (60%) students indicated at least one of these as their main unanswered question. Question 4, whether DDT should be banned, seems like an obvious question, as most students were unsure at the end of the session about fundamental issues in deciding on the matter. This is further reflected in question 5 about the alternatives. This question indicates that students are not happy to ban DDT unless alternatives are available, i.e. they are concerned about malaria control, possibly even more than the risks of DDT polluting the environment.

This is a positive result with 60% of the students explicitly recognising one of the key ethical issues in the discussion. Also students asking questions 4 & 5 have elements in their questions that relate to the fundamental ones.

Students also gave a score for themselves, the group and the facilitator on how well they thought each had done during the session. The students' self-assessment average was 2.9, their group appreciation 4.1 and the facilitator average of 4.3; all out of five.

9.2.2 Summary

Ecological decision making - DDT and Malaria is not an easy exercise. It can be successful, but it requires the facilitator to have a clear idea of the core issues and skill to lead the discussion so that the students are encouraged to recognise the underlying moral issues. The demanding scientific reading is hypothesised to have encouraged students to think, but only on a quite limited scale in relation to the complex problem. 60% of the students reporting learning gains corresponding with the set objectives is a good result, in particular when a further 21% of the students report learning gains corresponding to the general learning aims for the course. A majority of the students left the session with a clear understanding of what are the important, but still unanswered, questions.

9.3 Moral implications of genetic screening

This research started with an intention to teach genetic ethics as it was reflected in the original title 'Coping with genetic knowledge'. The theme was later greatly extended, but some elements of the original theme survived, this exercise being one of them. The idea had been to encourage students to think of their personal roles as both generators and consumers of genetic knowledge; to prompt them to make decisions on what information they wish to have of themselves or their future children, and in what sort of projects they want to participate professionally. The expansion of genetic knowledge is philosophically intriguing and complex, and the desire was to give students a window into an alternative way of considering genetic knowledge.

An exercise on the moral implications of genetic screening was generated to fulfil these aims. The most natural target-group of students for genetic ethics was Genetics and Molecular Cell Biology (MCB) students. A decision was made to introduce ethics to the Genetics curriculum and to use MCB students as a control. The genetics students were then given a choice on whether they wished to do this topic or embryo research (see section 9.4) during their second ethics session. Two groups (15 students) chose this topic, where genetic screening is examined through a case example of cancer with additional consideration on whether screening could be used to eliminate disease. The remaining two groups chose embryo research.

The topic being complex made facilitation crucial. Without a strict focus on the issue, the discussion easily branches out to other issues and the overall benefits of the session are less obvious and concrete. I had two different experiences. With the first group I let the discussion progress without enough guidance and the session was enthusiastic, but unfocused to a point that I don't believe students gained the maximum benefits from it. With the second group I kept more stringently to the guideline notes I had prepared and the session was superb. The students were active, the discussion progressed logically and the students seemed very pleased with themselves and the group at the end of the

session. Students were asked to give marks to themselves, the group and the facilitator and in all cases the second group gave marks which averaged 0.5 higher than the first group (myself 2.7/3.2; group 3.9/4.3, and facilitator 4.1/4.7). The session progress I describe here represents the session I had with the second group, with whom I followed my pre-prepared guidelines.

The pre-session reading was Ponder (1997), who discusses the use of genetic testing for cancer susceptibility from scientific, social and ethical viewpoints. The session started with a short written answer to a question on the topic. The aim was, as in all sessions, to focus student attention on the subject and to collect information on how students had perceived the given reading. After this we had an informal check-in, where students were given a chance to give their personal feelings and views on the reading. This was to encourage honesty and friendliness within the group. This time there were no major problems with the paper, students had found it interesting and challenging in a good way. This was followed by a vocabulary check, and the students queried a few medical terms like endoscopy. We then proceeded to identify the main message in the text. Students raised three types of central concerns: how to interpret the results, how the information can be misused, and whether screening is truly cost-effective. They also recognised two possible benefits that they wished to add to the central message: genetic screening could reduce other types of medical screening and negative results could reduce anxiety. The group then spent 15 minutes going through the views found in the paper on these issues in order to create a shared understanding on them.

After this the students were encouraged to widen their perspective from cancer screening to screening in general. They were prompted to think of screening possibilities that they would consider to be without problems and ones that they would think should not be allowed. The white-board was used to draw a continuum between absolute good ones and absolutely bad ones. Later the same continuum was used when the students attempted to draw a line between those they believed the health care system should adopt and those to reject. Among the definite rejects were screening for external features like eye colour and among

the accepts were pre-symptomatic testing for preventable genetic conditions. The group debated long and hard on the middle-ground possibilities, including in their considerations the scientific problems, resource allocation issues, and emotional impact of genetic knowledge. Quite naturally, no decision was reached, but neither were the students frustrated by this.

The group was then asked their personal judgement on how much these considerations would influence their choices in future employment. Two polar answers emerged, greatly or not at all. Those saying not at all supported their view by the need of general advancement of science, which was considered inherently good, and on the division between knowledge and its application. For a moment the discussion was very lively as people holding different views actively challenged that of the others. The session ended in short comments from all students on how happy they were with the session and what they would do to improve it next time. Most students in both groups were happy with the session and the main improvement suggestions were personal: to prepare better and talk less/more.

In my view the session was a success, at least as long as it was properly facilitated. The issues seemed highly relevant to the students, they enjoyed the discussion and on many instances I could observe real learning taking place with students having revelations on how things are connected and what influences what.

9.3.1 Student responses to 'Genetic screening' exercise - qualitative material

The short written question at the start of the session was:

Q1 *What in your view is the most influential issue when deciding on genetic testing?*

All 15 students answered this question and all responses (n=16) were analysable using the key developed. There were only two main kinds of consideration: how beneficial is it to know genetic information, and how reliable are the tests. The first concern about benefits could be divided into three sub-categories: is it worth it?, do we really understand, and the importance of counselling. One student wrote down two issues (2 and 3), so the percentages don't add up to one hundred, but represent the percentage of students (n=15) recognising each issue.

Knowing	1 benefits - is it worth it? 67%	2 Do we understand the info/ramifications 27%	3 Counselling? 7%
Test	4 Are they reliable? 7%		

Table 19: Genetic Screening, Question One

The student responses are very positive as the majority of students are asking one of the core ethical questions - is it worth it? This is a good start to the discussion and a positive indication that the text chosen has supported the students' perceptions of the important issues.

At the end of the session, students answered two short questions in writing. The first one was:

Q2 What is the main point you learned from the discussion today?

This question serves as a check on the main learning gains the students themselves perceive. My set objectives for this session were to encourage the students to consider their personal role as both generators and consumers of

genetic knowledge, to gain understanding on the different aspects of genetic screening and to allow them an opportunity to practice their skills in making difficult decisions.

The student responses all belonged to one category; the complexity of the issue. This could be further divided into three sub-categories: 1) recognition, 2) statements and views, and 3) deliberations. All student responses (n=15) fitted into this key and as three students gave an answer (n=18) that had more than one element, the percentages do not add up to one hundred. The quotes are from the student responses and represent typical answers in each category.

	Recognition	Statements	Deliberations
Conflict and complexity	1 "These issues are really complex ." 60%	2 "(Nothing is as clear as it first seems, element 1), lots and lots of research may help, but all situations need to be considered individually." 13%	3 "We need to establish guidelines - what is a disease? Where next? Counselling is essential." 47%

Table 20: Genetic Screening, Question Two

All the responses match the set learning objectives. I am particularly pleased with the number of students who wrote down sophisticated deliberations rather than just short answers. This corresponds with my experience during the sessions; students were keen, they discussed the issues enthusiastically and the learning gains were substantial.

The final question students answered at the end of the session was:

Q3 What issue do you find most confusing after this session?

Answers to the question can provide two types of information; have the students recognised the dilemmas in this subject, and whether the session design left some simpler issues unclear.

The responses could be divided into three categories: questions about decision-making, questions about testing itself, and elements in opinion. These could be further divided two sub-categories: straight questions and methodological questions. One student answered 'nothing', and was given a category all on its own. Thirteen students answered this question and all stated only one element, so this time the percentages add up to 100 and represent both the percentage of students asking a particular question and the proportional popularity of each response.

	Straight questions	Methodological questions
Decision-making	1 Where to draw the line? 39%	2 Who should make decisions and how? 15%
Testing	3 Is it worth it? 8%	4 How should we use it? 23%
Elements in opinion	5 Unsure of my opinion 8%	6 Nothing 8%

Table 21: Genetic Screening, Question Three

These responses represent very good recognition of the central issues, with the core questions (shaded boxes) asked by all but two of the students. Also the student questioning his/her own opinion seems to have been prompted to ask very crucial questions. The student who stated nothing is less encouraging, as the session discussed issues for which answers are at this stage very tentative.

9.3.2 Summary

The exercise on genetic screening has potential to be a wonderful one, but only when facilitation is carefully planned and quite strictly adhered to. The given reading has supported students' recognition of the main issues and the students report having learnt what the exercise was designed to teach. Also, the vast majority of students left the discussion with important questions in mind. The small student number to some extent reduces the confidence in the conclusions, but nevertheless, the results are encouraging.

9.4 Embryo research

In the original plan, the discussion programme with Genetics students was to consist of five sessions, and 'Embryo Research' was designed together with 'Genetic Screening' to highlight both methodological and theoretical issues in genetic research. With the final programme of only three sessions, each of the four groups was given an opportunity to choose between these two exercises, and two groups (15 students) chose 'Embryo Research'(understood as human embryos).

The use of embryos invites one to consider when the methods of science may become unjustifiable, regardless of the potential benefits. Use of animals in research raises similar issues. Embryo research is a high-tech method used for very specific research purposes. The core concerns are the potential devaluation of human life, and opportunities to create 'super humans'. Religious appreciation of embryos is part of the problem. During this exercise I made a conscious choice not to discuss religious judgements *per se*. If a student brought religious aspects to the discussion, they would have been considered, but in general I tried to create a discussion where the issue of embryo use was considered on as neutral and scientific grounds as possible. Both groups had Catholic and Protestant students (they volunteered the information at the start of the sessions), but religious considerations never spontaneously entered the discussion.

The pre-session reading was Eisenberg and Schenker (1997), which describes the potential benefits from embryo research, the sources of embryos, and the ethical considerations. The authors are scientists and the coverage of the issues is concise and from a philosophical point of view sometimes even simplistic.

The session started with a short written answer to a question about embryo research to focus student attention on the topic and than a quick check-in to get a feeling for what people thought about the issue. It was obvious from the beginning that the topic had sparked the students to think and the sessions were lively, sometimes almost too lively, with students eager to contribute. The

vocabulary of the text was not hard and the words queried were mainly scientific: e.g. ICSI (Intra-cytoplasmic Sperm Injection), and Lesch-Nyhan syndrome. The recognition of the main themes in the text was likewise quick, as the text was clearly indexed and titled.

The group first got down to discuss potential benefits of embryo research, mainly improved fertility treatment, but also better understanding on basic biology, human disease, and possible future gene therapy. Both groups ruled improvement of infertility treatment as a non-essential scientific goal. They went even further by stating that the use of infertility treatment was in fact immoral. When asked why, they first reflected on over-population. The students said that the world is already so over-populated that we should not assist anyone to have any more children. The counter-question I asked was; 'why limit the reproductive rights of just infertile couples and not all couples?'. To answer this question they provided an evolutionary reason: if you are infertile, you are not meant to reproduce and you would be passing on 'bad genes'. I questioned this by asking, would they apply the same criteria to people who are otherwise 'imperfect'; should they be allowed to die, as they are not in an evolutionary sense 'meant to thrive', or should their reproductive rights be limited, as they also would be passing on 'bad genes'. To some extent they saw the fallibility of their argument, but they still believed there was an inherent difference that could be captured. The third suggestion they made was to say that wanting to reproduce is such a 'want-thing' and people just cannot get all they want. My counter-question was: why is this want the one you cannot have, and who makes the decision which wants are 'allowed' and which are not? Again, the students could follow the argument, but were not convinced that it effectively discredited their view that fertility treatments are immoral. They did not bring into the question whether infertility treatment should be provided by the public health care system, or that they were worried of consequential issues of donated gametes, for example. Even when unable to find a good argument why infertility treatment is immoral, they persisted that it was so.

The students' reaction to infertility treatment truly puzzled me. The only sensible interpretation I could think of was that they are still very young (approx. 20) and that they have not ever seriously considered having children of their own, and thus are incapable of perceiving the issues involved. This alone does not explain their view though. It would, if they had protested against public money being used for infertility treatment, but not when they point-blank classified it as immoral.

They further discussed the use of embryo research to gain greater understanding of human development, which they agreed was a 'good' aim. They were more concerned about the increasing opportunities to manipulate the human genome, branding that undesirable, at least as long as the aims were not related to alleviating pain and disease. The group then briefly tried to agree where to draw the line of what pain relief is; whether, for example, compromised appearance or intelligence could be considered as a 'source of pain'.

The group then proceeded to discuss the status of embryos and whether it makes a difference how research embryos are collected. Three alternatives were considered: embryos are persons from conception, or they are a collection of cells until point X when they become human, or they are potential humans and thus require respect, but not full rights. The intuitive feeling in both groups was that an embryo cannot be a person from conception. The groups then tried to define when an embryo would gain enough human-qualities to gain at least some human rights. The white board was used to mark potential cut-off points; primitive streak, brain-activity, viability, and birth. The groups could not agree how to make the decision or where to draw the line and reluctantly agreed that conception seems like the most logical point of difference. Still, they wanted to give embryo research an opportunity and not draw equation marks between embryos and humans, thus granting embryos full human rights and effectively making embryo research impossible.

When prompted to think, the students did wish to use some sort of harm/benefit calculus to decide when to use embryos for research. The underlying assumption

was that embryos are something 'special' so we cannot use them as any other collection of cells, while at the same time, they are more disposable than fetuses, infants, and/or adults.

The students were then invited to consider their personal view; would they be happy to work with embryos? The groups were divided, with approximately 2/3 believing they would have no problem with it and 1/3 not wishing to participate.

The session was finished with the traditional quick questions; did you learn anything today and how would you improve for next time. The students were enthusiastic about their learning and enjoyment during the session and the improvement suggestions were once again very limited, mainly suggesting that they themselves talked less or more next time.

I was pleased with the exercise. It encouraged student participation, provided good material to consider important issues, and allowed students to discover new areas. The facilitation of this session was rather to control student enthusiasm, than to try to create it. It was important to direct the discussion quite strongly, because the emotional elements were so apparent, that the group needed to be kept on the topic by external guidance.

9.4.1 Student responses to 'Embryo research' exercise - qualitative material

At the start of the session, before any issues were discussed, the students were asked to write a short answer to the following question:

Q1 What do you think is the most difficult issue in deciding whether to do embryo research?

By asking this question I hoped to focus student attention on the subject from the start of the session, and I was also interested in the influence of the article and the students' views on the subject in general. Three main issues emerged from

student responses: whether an embryo is human or just tissue, can we justify it by benefits, and personal views. Two students (n=15) wrote a response that had elements from more than one category, so percentages that represent the percentage of students stating each element and add up to more than one hundred. All responses fitted the designed key.

1	2	3
Is an embryo a human, where do we draw the line?	Benefits and harms of research - can we justify it?	That it is potential life.
73%	27%	13%

Table 22: Embryo Research, Question One

Most of the students raised important points and used ethical language, which indicates that the reading was provocative and influential. This was a good start to the discussion.

As a last thing during the sessions, students were asked to answer two questions. The first one was:

Q2 What is the main point you learned from the discussion today?

The set objectives for the exercise were: to create understanding of all the elements that influence our choices to use embryos in research, to encourage consideration of the status of human embryos, and to support decision-making that uses a harm/benefit approach. This question was designed to collect data on how well these objectives were met.

Practically all the responses reflected on the difficulty of making decisions on this issue. Students either state the fact that it is a complex issue, that it is possible/impossible to agree, that science must/always does progress, and that they now have to reconsider their own view on the subject.

There was one response that did not fit the key: 'Embryo research - great argument starter.' The student actually came forward afterwards and apologised for his comment (after receiving the e-mail summary), as he agreed that the session actually proved that discussions about embryo research do not need to turn into arguments. In this case all students stated only one learning gain, so the percentages represent both the percentage of students choosing a particular category, and the proportional representation of each category.

1	2	3	4	5
These issues are really complex	It is hard (if not impossible) to agree on this issue.	Consensus is possible.	You cannot stop science.	I realised I cannot justify my view.
21%	7%	21%	21%	29%

Table 23: Embryo Research, Question Two

Having 29% of the students explicitly state that they need to re-consider their view is promising as a sign that the exercise has provided something new and fundamental about the subject. The students in category one (21%), stating that the issue is more complex than they had thought could be similarly read as a positive indication on the impact of the exercise on students' appreciation of the topic. From elements 2 and 3 it is obvious that the difficulty of decision-making on these issues is appreciated by the students, some being more positive than others on the future success of our attempts to make them. Element 4 is the most worrying one from an ethical viewpoint: students explicitly stating that science cannot be stopped. It is possible to interpret this as the students saying that we have no power to make decisions about the methods and aims of science. If this is the case, it could be a sign of passive acceptance of all decisions in science or believing in actively pursuing all opportunities in science as there are no grounds for limiting scientific pursuits. It could also be interpreted as a reaction to the discussion which explicitly focused on the needs to limit science, and thus rather as a view that science *should* not be limited. This reflects a common attitude that

says: if anything is possible in science, it will be done by someone, so imposing limits is worthless.

The second question students answered at the end was:

Q3 What is the main unanswered question you leave the discussion with today?

Again, the question was asked in order to get an understanding of what questions the students leave the discussion with, which indicates the issues they will spend more time thinking about.

The student responses were of two main types: queries and views on decision-making, and straight questions about the acceptability of embryo research. All students who answered this question (n=13) identified only one unanswered question, so the percentages reflect both students asking each question and the proportional popularity of each question asked. All responses fitted the key.

	1 Where to draw the line? Where should we stop? 53%	2 Can we choose to stop and who should decide? 20%	3 Can we ever agree? 7%
Decision-making			
Questions	4 Is it right or wrong? 7%	5 What is the balance of harm and benefit? 13%	

Table 24: Embryo Research, Question Three

73% of the students asked important and fundamental questions (elements 1 & 2). Decision-making is crucial in scientific methods and it is encouraging that the majority of the students have recognised these questions and highlight their

importance. Element 3 reflects the doubts already apparent in the above question. Also, the question asked, is it right or wrong, and balancing harms and benefits are important questions. Students asking them are forced to consider some of the fundamental issues in the subject, which is positive.

Students also gave a score for themselves, the group and the facilitator on how well they thought each had done during the session. The students' self-assessment average was 3.2, their group appreciation 4.2 and the facilitator average of 4.2; all out of five.

9.4.2 Summary

I am moderately happy with the session. The learning gains were not as good as in some other courses and the session was not easy to facilitate. The strong view the students put across about the immorality of infertility treatment was also unexpected. Developmental biology students, for example, have not expressed similar strong attitudes (personal correspondence with Dr J.R. Downie). More research might be interesting to study the popularity of this view and the possible changes during the next few years in university. Regardless of the complexities in the exercise, it did create enthusiasm and the students left with different important questions from what they arrived with, which can be interpreted as a sign of learning during the session.

9.5 *Genetically modified organisms*

This is the second of the original exercises dedicated to explore recent advances in genetic research. The aim of the session was to introduce students to the social and ethical aspects of genetic modification, in particular in food crops. The session was designed to allow students an opportunity to practice their own decision-making skills while they were learning more about the issues.

In the end only Botany students (n=12) participated in this exercise. It was considered important and interesting by other course-co-ordinators as well, but when the sessions in many instances were limited to three, other subjects took precedence.

With the Botany students this exercise was divided into two parts; the first looking into the basic ethical principles relevant to the evaluation of GM technology and the second into issues of justice and consumer choice. The sessions were held three weeks apart. I will discuss here both sessions together. I have also carried out the qualitative analysis of student responses for both sessions together.

The pre-session reading was the introductory chapter of The Nuffield Council on Bioethics Report on GM Food (1999) and three case studies (see Appendix XVIII - GM crops). The reading was chosen to give a concise presentation of all the relevant ethical issues in GM crops. The material was demanding, but well-written and clear. Both sessions started with a short written question to focus attention on the issues to be discussed. Then the group was given a chance to state their views and feelings about the reading. The students found the text interesting, but challenging, because of its unfamiliar content and philosophical vocabulary. The vocabulary clarification that was carried out after the introductions included the following: definition of science, intrinsically wrong, welfare of citizens, hubris, and Pareto-optimality. Thus in this case, the group spent time discussing more than just dictionary descriptions of words - what they really mean and how they are used in the text and in general language. The group

was very quiet and shy during these discussions and views had to be directly asked before the students participated in the discussion.

The group was then guided to identify the main issues in the dedicated section of the reading. For the first session, these were the first two of the five ethical principles discussed in the text: general welfare, and rights of people. The group went through them one at a time, clarifying their meaning and discussing their importance. These seemed completely plausible and sensible ethical principles to the students and they found no difficulty in discussing them in relation to GM crops. During the second session the main themes were justice, ethical status of nature, and rights of the environment. The meaning of justice was discussed first. The group considered utilitarian approach to securing justice and used Pareto-optimality as a special tool. The other two issues, rights of nature and what is unnatural, were more difficult to elucidate. The concepts seemed foreign to the students and the discussion dried up very soon if the students were not prompted with further questions or alternative ways to look at the issue.

After the basic issues, in both sessions the students applied the concepts to the case-studies provided with the preliminary reading (see Appendix XIII - GM Crops). In the first session the students first considered issues of safety as a variant to the principle of general welfare. It was emphasised that safety is rarely a scientific result, but a judgement we make based on our evaluation of risks and benefits. The students were more participatory when they had a chance to look at a real case with real data and when given an opportunity to apply their own judgement to it. The second case during the first session was on labelling and it was designed to highlight problems with rights and GM crops - do consumers have a right to know, or producers a right not to tell, and who pays for the extra cost of labelling? Again discussion flowed freely and students were keen to exchange views on the issue. During the second session, the case study was on international justice and the development of GM crops. Once again the students were more forthcoming in the discussion when a clear starting point was provided by the case study.

When students were asked their personal view on GM crops at the end of each session, the responses were ones of hesitation. This was somewhat surprising considering they were Botany students and very likely to be working on GM related projects in the future. On the other hand, the hesitation can also be interpreted as a positive sign that the students do not have a strong pre-chosen view on the subject and are thus more open to new information to influence their choices.

Both sessions ended in short responses from all students to re-cap what they had learnt and how to improve for the next time. During the first session, the main improvement suggestion was a better location - we had been squeezed into a narrow space between lab-benches, which definitely was not ideal. After the second session, most students criticised themselves for either poor preparation or participation in the discussion. As a last thing the students were asked to write short answers to two questions.

I was happy with the exercise. The sessions were quite different. The first one was quiet and a lot of facilitation was required in order to encourage student participation, while the second one was wonderful and the students truly took part and came up with really sophisticated and complex thoughts. At least two things might have contributed to the difference between sessions: 1) the first session was held in a lab, where students could not see each other well and the discussion between students was thus relayed via the facilitator, 2) the first session was during week 2 in term 1 and the students did not know each other at that point. The material worked well by providing a suitable challenge and good coverage of the area.

9.5.1 Student responses to 'GM Crops' exercise - qualitative material

Once again, before anything else, the students wrote a short answer to the following questions in sessions 1 and 2:

Session 1: *Q1 What do you think is the most important issue of rights and general welfare in relation to GM crops?*

Session 2: *Q1 What do you think is the most important issue of justice in relation to GMOs?*

In both sessions issues of consumer rights and division of benefits could be identified as a theme. In the first session responses of safety could be identified as a further theme. All student responses (n=8) could be categorised within these themes and as all students raised only one issue, the percentages represent both the proportion of students raising a particular point and the frequency of each theme raised.

	1 Consumer autonomy	2 Safety	3 Who benefits?
Session 1	37.5%	37.5%	25%
	1 Consumer autonomy	2 Who benefits?	
Session 2	12.5%	87.5%	

Table 25: GM crops, Question One

The student responses indicate that the reading raised very specific points and that students have taken these to be important. Also all the issues raised are relevant to the ethical analysis of GM technology, which prepared the students well for the discussion.

At the end of the session students first answered the question:

Q2 What is the big point you learned from the discussion today?

The aim of the exercise had been to give students better understanding about the complexity and issues involved in the development of GM crops and to prepare an opportunity for them to practice decision-making on the issue.

The student responses from both sessions spread relatively equally between three categories:

1. Conflict and complexity
2. Insufficient data
3. Views and learning skills

One student stated more than one learning gain, so the percentages here represent the percentage of students (n=9 in both sessions) recognising each issue rather than the proportional popularity of each theme. All responses (n=19) could be analysed according to this key.

	1	2
Conflict and complexity	"These issues are really complex" 22%	"Need to consider moral issues first." 40%
	3	4
Insufficient data	"Who is really going benefits - big companies?" 11%	"Risks associated with this need to be weighed up against the benefits of doing it for the population of the world.". 11%
	5	6
Views and learning skills	"There seems to be no one solution that everyone will agree with." 11%	"Better prepared next time, to think more for myself and rely less on others' opinions." 11%

Table 26: GM Crops, Question Two

The highlighted themes correspond directly with the set learning aims for this exercise. This means that 69% (13) of students report having gained the designed learning . This is very encouraging. Further, the stated learning gains in 4 and 6 are in accordance with the general aims of ethics teaching, though not directly corresponding with the session objectives. Category 5, on the other hand, corresponds only minimally with the chosen objectives.

The second question at the end of the session was:

Q3 What is the main unanswered question you leave the discussion with today?

Again, the hope was to gauge whether the students had recognised any of the genuinely difficult to answer questions and had left the session with those in mind for further consideration and thought.

The student responses were of two main types: questions about decision-making and straight questions about GM crops themselves. All responses (n=17) fitted into the key and all students presented just one question, so the percentages represent both the proportion of students asking a particular question, and the proportional popularity of each theme.

	1	2	
Decision-making	Do benefits outweigh the risks? 35%	No definite answers, insufficient data 24%	
	3	4	5
GMOs	Yes/no? I still don't know whether we should go ahead. 24%	Natural/unnatural 5%	Worry about environment 12%

Table 27: GM Crops, Question Three

Categories 1 and 4 are best interpreted as questions resulting from the discussion and reading material directly as they use language most students were introduced to during this exercise and one which they struggled to use at the beginning. This is a positive sign with 41% of students having recognised one of the key questions in the GM debate. Also category 2 and 5 can be interpreted to stem at least partially from the discussion on the definition of safety and how they are always issues of judgement rather than results of scientific research. Category 3 is a personal reflection on the issue, which as such is very positive showing that students are interested in forming their individual response to the situation. This will always require thinking, which is definitely to be encouraged.

Students also gave a score for themselves, the group and the facilitator on how well they thought each had done during the session. The students' self-assessment average was 2.6, their group appreciation 3.8 and the facilitator average of 4.4; all out of five.

9.5.2 Summary

The exercise on GM crops proved to be an interesting and stimulating one with students participating actively. The learning gains of this exercise have been positive and a majority of the students left the exercise with important questions in mind. The length and the complexity of the exercise supported its implementation in two separate sessions, which could both also be used as stand-alone sessions.

9.6 Clinical trials in developing countries

Pharmaceutical research is often ethically complex, in particular when we consider the practicalities for setting up clinical trials. The medical professionals have for a long time recognised the importance of informed consent, the problems with placebo trials, and patient autonomy. More general problems with drug trials include how to decide when a trial has been a success. Pharmacology students are likely to be involved in clinical trials, though their role is not the same as medical professionals. This exercise was designed to give these students an opportunity to familiarise themselves with the ethical concerns of drug trials, in particular with those carried out in developing countries. The exercise was intended to raise their moral sensitivity towards these issues and encourage them to improve and develop their ability to make decisions in complex trial situations.

The exercise outline was strongly influenced by practical requirements. The Pharmacology students were to attend the exercise in a group of 20 students, which made a discussion approach unsuitable. As an alternative, the exercise was designed as a mini PBL where the student worked in groups of 6-8 on a problem (see "Appendix XIX - Clinical trials in developing countries"), which was then briefly discussed with the entire group. This approach was also suitable for the pharmacology students in particular, who, according to other members of staff, were reluctant to take part in general discussions.

The pre-session reading was Lurie and Wolfe (1997), which looks at the ethical problems in HIV drug trials in developing countries. The article describes basic ethical guidelines for developing-country trials, the role of placebo-trials and the adequacy of data analysis in this particular case. It was chosen as it describes ethical issues in context. Students were also given a copy of the World Medical Association Declaration of Helsinki on biomedical research.

The session started with students writing a short answer to a question on clinical trials. Many students referred to the reading before answering the question,

which improved its impact as a method of focusing student attention on the discussion issues. This was followed by a brief round on how students had perceived the reading. Most had not found it too difficult, though the science had been complex in places. The general view was that the reading had been interesting and relevant. The students, who were already sitting in groups (the room was set out so that they had chosen a seat around one of the tables as they entered the room), were than given the brief hand-out on a suggested trial protocol for a Malaria vaccine trial to be carried out in Botswana. They were given an explicit task to complete and thus set out to work. The small groups had difficulty getting started and it was important that the facilitator actively asked them questions and guided their attention to the problem. This needed to be continued through-out the short group-work session, as the groups tended to get distracted and discuss irrelevant issues of malaria (personal experiences) or other private matters. The groups were not particularly motivated, but in the end they did work through the required steps and the general discussion could begin.

It became apparent that in their own groups, the students had considered most issues from very limited view-points. The general discussion thus introduced several new issues into the discussion. This indicates that reading a text and discussing it with peers, may not be enough to stimulate moral sensitivity in students. Active and direct external encouragement seems to be necessary before new aspects are perceived and considered. In the end the students agreed on a revised protocol.

The session ended with a brief round which clarified what the students had learnt and how they would improve the sessions for the future. It was clear that they had enjoyed working in groups, as opposed to being asked to have a general discussion (method in the previous session). They also said that they had learnt a lot and that the topic had been interesting. This enthusiasm and learning was less apparent to the facilitator, so some level of doubt exists whether the students genuinely meant what they said, or said it because they believed it was the 'right thing' to say.

The experience with this exercise was not the most encouraging. I believe the design to be good and the reading relevant, but this particular group of students was not easy to motivate and thus the session was laborious from the facilitator's point of view. The group was also reluctant to discuss issues during the other two sessions, which indicates that there wasn't necessarily anything wrong in the session as such, but rather the students were less active than most of their peers.

9.6.1 Student responses to 'Clinical trials' exercise- qualitative material

At the start of the session, students were asked to write a short answer to the following question:

Q1 What do you think is the main ethical concern when testing drugs in developing countries?

The reason for this question was to highlight the theme of the session and to collect data on the impact of the reading on recognition of ethical themes.

The student responses all related to the treatment of subjects and could be classified within three main categories: informed consent, equality between research subjects in developing and developed countries, and whether poverty is used against participants' interests. All 16 students responded with one concern, so the percentages below represent both the students who chose each issue and the proportional popularity of each theme. All responses were included in the analysis.

1 Autonomy and informed consent. 25%	2 Equality between research subjects in developing and developed countries. 62%	3 Whether poverty is used against their interests. 13%

Table 28: Clinical Trials, Question One

The use of philosophical language in category one is possibly a result of students adopting new elements from the reading. Most of them said that they had never before reading this paper considered ethical concerns and thus would have been unfamiliar with the vocabulary involved. The other two categories involve considerations of fairness, which are both important ethical considerations in clinical trials. These comments should have provided a good starting point for the discussion, as all students had recognised important ethical concerns.

At the end of the session, students were asked to write short answers to two questions. The first one was:

Q2 What was the big point you learnt from today's session?

The learning objectives for the session were to increase students' awareness of ethical issues in clinical trials, and to give them an opportunity to practice their own decision-making skills in a complex moral problem. The responses to this question would indicate how well the exercise meets the learning objectives.

The student responses belonged to two main categories: ethical dilemmas and considerations about third world countries and drug trials. These could then be further divided into three sub-categories: recognition, deliberation and questions. There were more identifiable responses (n=21) than students (n=18), and the percentages here represent the percentage of students stating a particular learning gain, rather than the proportional popularity of each category. All responses fit this key.

	Recognition	Deliberations	Questions
Ethical dilemma	1 Existence of many ethical dilemmas 17%	2 It is often difficult to balance science and ethics in research, ethics needs to be considered. 33%	3 Should we test in developing countries, if the benefits are to go elsewhere? 6%
Third world countries and drug trials	4 It is easier to exploit people in developing countries. 11%	5 Education is difficult but essential when desigining drug trials in developing countries. 17%	6 What is a successful trial design? 33%

Table 29: Clinical Trials, Question Two

Sub-categories 1 and 2 correspond directly with the chosen learning objectives for the exercise. This means that 9 students (50%) have learnt exactly what the exercise was designed to teach. The other learning gains reported are also important ethical questions embedded in the exercise structure, though not as such included in the exercise objectives. All the reported learning gains are essential ethical issues. Thus, even though the percentage of students reporting learning gains directly corresponding with the chosen objectives for the exercise is relatively low, the response is overall very positive.

The second question students answered at the end of the exercise was:

Q3 What is the main unanswered question you leave the discussion with today?

This question was asked in order to find out whether the session had left some essential issues unanswered and also whether to students had recognised what the important ethical issues in clinical trials are.

The student responses belonged to three main categories: what is a successful trial design, what is a successful outcome of a trial, and how to educate participants in clinical trials. Two students asked more than one question and the percentages below represent the percentage of students (n=12) asking a particular question rather than the proportional popularity of each question. For this reason the percentages add up to more than one hundred. All responses fitted this key.

1 What is a successful drug trial design - who to include, how to help those that get sick, use of placebos etc.?	2 What is a successful outcome - vaccine protects 50% or 100%? Who makes the decision?	3 How do you educate people to ensure informed consent?
42%	58%	17%

Table 30: Clinical Trials, Question Three

All students asked questions essential to a good clinical trial design. In the first two categories the questions were in most cases very broad and fundamental to clinical trials, while the third category represents a more detailed question within this framework. This is a positive outcome, with all students leaving the session with a clear idea about the problematic questions.

Students also gave a score for themselves, the group and the facilitator on how well they thought each had done during the session. The students' self-assessment average was 3.6, their group appreciation 4.4 and the facilitator average of 4.6; all out of five.

9.6.2 Summary

The exercise on clinical trials in developing countries was based on a short problem which students were encouraged to solve during the session with the information they had gained from the given reading. The session was only run once and with this group, enthusiasm was not high and facilitation needed to be active. The student responses to the questions at the end of the session indicate

that despite the lack of participation during the session, students had picked upon the important issues in the clinical trial questions. Thus the exercise has promise to be successful with other groups as well.

9.7 Scientific misconduct and integrity

Scientific misconduct is often what people have in mind when science ethics is discussed - the ethical implications of scientists fabricating/falsifying their results or plagiarising each others work. These are important elements in being a 'good scientist' and thus it would be difficult to justify a science ethics course without including these considerations. But being a good scientist is a lot more than just following the rules laid down in basic scientific methodology. It means also integrity in the choice of research aims and methods. While the rules against falsification, fabrication and plagiarism are clearly stated, integrity is based on an ethical appreciation and understanding of the meaning and purpose of science and relationships between researchers, human subjects, animals and the environment. Therefore questions of misconduct and integrity are different, even though they are both elements in being a good scientist.

In the original plan, these two elements were to be addressed in two separate exercises. When the maximum number of ethics exercises was set to two or three by all course-coordinators, they were merged into one exercise. All students who participated in the ethics discussion group programme took part in this exercise, which was always the last exercise in the series. The aim of the exercise was to clarify what is considered scientific misconduct, what scientific integrity is, and to invite students to reflect on their personal role as future scientists in relation to these two themes.

The pre-session reading was a short paper (see Appendix XX- Scientific misconduct and integrity') prepared specifically for this purpose. It had brief descriptions of misconduct and integrity plus three short case studies to highlight the issues. Students were asked to prepare responses to all case studies before the session. I chose to write the reading for this exercise after being unable to find a suitable short summary in literature.

This exercise was one of the best ones. Students found the issues relevant and the case-studies provided a good starting point for interaction in the group.

Each session started with a short written reply to a question. After learning that students' self-evaluations in the previous sessions had been a lot harder than their evaluation of others and the facilitator, I was interested to find out the criteria students used to evaluate themselves. With the later groups, the question queried when students had been tempted by misconduct themselves. When the question was about self-evaluation, the group then briefly discussed the given responses and tried to agree on evaluation criteria. Many students could never consider giving themselves 5 out of 5, because they did not think they would ever be perfect. Also, students indicated that modesty is a highly regarded value, and no one wanted to seem boastful. They were then encouraged to consider 'moving targets' i.e. that they could reach one goal and give themselves 5/5 and then give themselves a new target to thrive for. They seemed genuinely puzzled with the thought of changing criteria of 'perfection'. With the groups who had been asked to consider when they would be/ had been tempted by misconduct themselves, the session started with sharing those thoughts. Both questions focused students' attention on themselves and how they fit into a network of assessment or misconduct. This was important as this exercise hoped to promote personal reflection on professional standards.

The groups were then asked to briefly describe how they understood falsification, fabrication and plagiarism. Then the focus was narrowed to falsification and fabrication and the students were prompted to consider reasons why scientists would resort to these methods and whether they could ever justify such a choice. Reasons they came up with included pressure from superiors, desire to make money/meet deadlines, desire to be known for a great discovery, and laziness. No one could justify misconduct, especially after the group explicitly considered the repercussions of such actions.

Case 1 was designed to highlight falsification. Most groups (but not all) were unanimous that the students depicted in the case could not remove the out-lying observations from their graph because 'they do not know if they are the 'right' observations'. It was mostly suggested that the students should repeat the

experiment if possible to find out if their hypothesis about the temperature controls was correct. If they could not repeat, then they would have to explain their problems and as a result they should have less confidence in their results. The discussion got more interesting when the students were asked how they would punish this sort of behaviour if these students were in level 4 and this was an important part of their Honours project. The initial reactions were to either do nothing or to kick them out. Each group then considered why they should be punished (to learn not to do it again, possibly to protect people who would follow up their research) and what in the end would be a reasonable reprimand for their actions. Most commonly groups chose to reduce marks and make it public without mentioning names, so that other students would learn from their mistakes. They also said that none of them had ever discussed the rules of science explicitly (apart from those who had chosen to take the L2 course 'Science Communication') with members of staff and many even said that they had been openly encouraged to falsify and plagiarise during laboratory work. The emphasis was then placed on staff to lay down clear rules and punishment that would follow if rules were broken.

The second case-study concentrated on plagiarism and the mentor-student relationship. Plagiarism was first discussed in general terms. Everyone could imagine being tempted and many thought that they had at least unwittingly plagiarised as well. The case study asked students to make a judgement on who should receive credit for a scientific discovery. All groups started with a view that the professor/supervisor was the one who should have the credit for a student's work. There were two main reasons: she wrote the paper, and she had designed the machine. The groups were then invited to consider what it is in a scientific discovery that is special. Most groups needed specific examples before they arrived at the realisation that it is the intellectual input, not necessarily who has done most work, that counts as something special and worthy of recognition. Still most groups believed that the professor was the originator and thus should get the credit. It was then suggested that the professor had already got credit for the methodology and whether that would change the situation. For most groups it did, but not for all. The students were than asked to describe what the students in

the case example had done and they all came up with the student building a hypothesis and the professor only helping to confirm it, and thus the student would be the main contributor to the discovery. It was then discussed what the student could do in a situation where the professor had written up the student's work as her own. Students considered the pros and cons for both parties in the dilemma and in most cases they were ready to encourage the student in the case study to fight for her rights by first seeing the professor and if she would not agree, then going to the head of department or the university senate to gain the credit she deserved. This case-study was a great revelation to many students in the sense that they realised that they themselves could be originators of scientific discoveries even as under-graduates.

The groups were then asked to describe how they understood scientific integrity. They recognised the difference between misconduct and integrity after some prompting and then the discussion focused for a short while on the integrity of scientific methods. Students were invited to think of situations where the aims of a project would be acceptable to them, but the methods would not - i.e. when the ends do not justify the means. Animal research was always recognised as a potential problem, but other examples were hard to find. With quite explicit prompting they recognised possible problems with large-scale field trials of GM crops (aiming to find out the safety of GM crops), human embryo research, and the use of human subjects. Because animal research had been discussed in the first session, the discussion focus was then moved on to the integrity of choosing research aims.

Aims and integrity were approached through a third case study, which describes a research proposal for intelligence-gene research. The students were explicitly encouraged to first consider the pro/cons of this research by using a harm/benefit calculus. Most of them thought that benefits were very few (they could be achieved by other, less controversial methods) and the potential harm significant. They were than asked whether they thought this was enough of a reason to stop this type of research from taking place and whether they would be ready to call a piece of knowledge 'bad' in itself, rather than its application. Students did not

outright reject the idea that in theory there could be pieces of information for which no good application could be construed and for which harmful uses were plentiful, but could not come up with any examples. The genetic research in the case did not qualify for the criteria. They were then encouraged to consider the responsibilities scientists have over the application of knowledge they produce. It was considered ludicrous to hold scientists responsible for all applications (one student said that it would be like blaming Henry Ford for all car accidents - a very fitting analogy). The students disagreed over whether scientists have *any* responsibility to consider potential uses of the knowledge they are producing and whether they should not proceed with the research plan if they see the balance to be negative. There seemed to be a gender difference here, with female students more strongly requiring moral responsibility from scientists, but I collected no evidence to back this up. As a last point the students were encouraged to think how much it would matter to them what the research they were involved in was aimed to do. Again, there were clear differences between students, some considering this to be of great importance and others saying that all information is neutral and that they would rather go and do what paid well and was interesting, rather than considering the further implications of their research choice.

I truly enjoyed facilitating this session. It was easy to get started as students could see the immediate relevance of the topic, and the case studied provided good ways to focus the discussion. With some groups I tried short role-plays to bring the case-studies to life. These worked fantastically and this teaching method is something I would like to develop further in the future.

9.7.1 Student responses to 'Scientific misconduct and integrity' exercise - qualitative material

Two different questions were used at the start of the session to focus student attention on the subject and more specifically on their own personal relationship with the issues. The first nine groups answered the following question:

Q1 a) How would you have to participate to give yourself 5/5?

The student responses formed three distinct categories: having a view of one's own, abilities to listen and contribute, and preparation and projection of learning. In all of these, three sub-categories could be also identified. 58 students from three different degree courses (Genetics (n=26), Physiology (n=4), and Zoology/Aquatic Bioscience (n=28)) answered this question and most student responses had elements that belonged to more than one category increasing the number of identified responses to 123. There were three responses (all from genetics students) that could not be fitted into the key:

1. If everyone agreed with me unreservedly
2. Very well indeed
3. Dunno

The percentages presented represent the proportion of students giving each response and because most of them gave at least two, the percentages do not add up to 100. I have also given group specific percentages for a) Genetics and b) Zoology/Aquatic Bioscience in the second category as there were interesting differences in their response patterns.

	1	2	3
My view	Have an opinion on every (important) point. 16%	Voice my opinion and talk sense. 31%	Be able to back-up and justify my views. 24%
Listening and contributing	4 Listen with an open mind. 43%, a) 27%, b) 64%	5 Involvement in the discussion, contribute significantly. 50%, a) 42%, b) 61%	6 Understand everything or have in-depth knowledge about the subject. 24%, a) 35%, b) 14%
Preparation and projection	7 Read article thoroughly. 14%	8 Never. 2%	9 Re-consider my view/ have an open mind/ think about it. 3%

Table 31: Scientific Misconduct and Integrity, Question One A

(a= Genetics and b= Zoology/Aquatic Bioscience)

In general students had a good idea of what type of participation is desirable in a discussion group. I would personally agree strongly with the shaded categories which emphasise participation and preparation as well as listening skills. I would be less likely to endorse criteria that required students to hold an opinion on all issues - it should not be considered undesirable to be in a process of forming one's opinion. The students' self assessment at the end of these sessions did not change significantly from what they had given themselves after the previous sessions. The benefit of this exercise was thus an improved shared understanding on what is 'good' participation in a discussion group.

The last seven groups (64 students) were asked to write a short answer to the following question:

Q1 b) Think of a situation where you have been/could imagine being tempted with scientific misconduct?

The question was changed mainly because I believed that the new one would provide a better starting point to the discussion.

The student responses formed two distinct categories: falsification, and plagiarism. Most suggested situations included a reason for misconduct and the responses could be divided into sub-categories according to the reasons given. Some students gave more than one scenario (n=73) and the percentages in the table represent the percentage of students in each category rather than the proportional popularity of each scenario. All responses were included in the analysis.

	1 Change data to fit hypothesis or 'right results'. 44%	2 Falsify for time pressure. 25%	3 Falsify to get better marks/more money. 30%
Falsification	Change data to fit hypothesis or 'right results'. 44%	Falsify for time pressure. 25%	Falsify to get better marks/more money. 30%
Plagiarism	4 Copy lab reports -> better mark. 8%	5 Adopt theories without credit. 8%	

Table 32: Scientific Misconduct and Integrity, Question One B

The vast majority of students seem to be most familiar with falsification. When specific cases were discussed, most experiences were with lab reports and often it was pointed out that staff had encouraged them to copy someone else's work when theirs had not worked out. Maybe this is an issue that should be addressed more generally by staff. Otherwise I believe this question set the scene for most sessions very well, with students bringing in their personal experience and being more likely to discuss their own experiences during the session.

At the end of the session students were asked two questions. The first one was:

Q2 What is the main point you learnt from the discussion today?

The objectives for this exercise were to give a clear description of what constitutes scientific misconduct; describe how it impacts research, encourage students to reflect on their own practice; and highlight the complexity of choices and responsibilities in scientific integrity and misconduct.

The student responses to this question formed four distinct categories:
1. Reflections on the importance of scientific integrity and misconduct
2. Consequences of misconduct
3. Personal reflections
4. Nature of knowledge.

Within these 2 sub-categories could be identified: 1) statements, and 2) evaluations/questions.

There were four responses that did not fit into this framework:
1. Ethics
2. Those who attended the seminar had not contributed to gross scientific misconduct.
3. People have different views on intelligence.
4. Science is subject to good and bad points of human nature.

Altogether 107 students answered this question generating 120 responses. The percentages in the table represent the percentage of students stating each learning gain rather than proportional popularity of each category and thus they add up to more than 100.

	1	2	
Importance of misconduct and integrity	Critical to good science - otherwise pointless. 10%	Lots of grey areas, not always easy to know what is right/wrong. 20%	
	3	4	
Consequences	Consequences of misconduct can be great - n eed to be considered. 28%	You need a big picture and at the same time consider all situations at their own merit. 7%	
	5	6	7
Personal view	You do not have to cheat! I learnt what it means to be a good scientist/clarified the issues. 10%	Personal integrity is important - we have responsibility/ importance of ethical considerations. 17%	Made me think! 6%
	8	8	
Nature of knowledge	Knowledge can be used for good and bad/ importance of considering implications. 8%	Who should be credited. 3%	

Table 33: Scientific Misconduct and Integrity, Question Two

None of the categories represent learning gains that would have been irrelevant in the light of the exercise objectives, but the shaded categories represent direct correspondence with the set ones. Categories 5 and 6 are considered to be particularly positive as they include personal involvement with the learning. 67 students (56%) reported at least one of these elements in their response. This is a very positive result that supports the positive impact of the exercise.

The final question as the end of the session was:

Q3 How would you describe the benefits of ethics seminars to students who have not yet participated in them?

By answering this question it was hoped that students would re-cap for themselves the benefits of ethics seminars and give feed-back on their views to the developers and facilitators of the seminar series.

The student responses generated two distinct categories: reflections on thinking and communication skills, plus three small sub-categories without a unifying theme. The main categories could be divided into three sub-categories to highlight different elements within each consideration. There was one response that did not fit into this key: "Probably came too late, better if tutorials were done in 2nd year, as they would be before you make your degree choice".

Altogether 135 students responded to this question generating 269 responses. The percentages in the table represent the percentage of students supporting each view rather than the proportional popularity of each category.

	1	2	3 Makes you re-consider your views / makes you think / helps you to develop ways to justify your views.
Thinking	Clarifies issues 9%	Broadens your view 43%	54%
expressing views	4 Helping you to express your views / communication skills / gives you confidence 20%	5 Important to listen to others / learning from others / interesting to hear others' viewpoints. 40%	6 Changed my views. 2%
other	7 Singular positive: Interesting / beneficial 6%	8 Should be more 23%	9 nothing 1%

Table 34: Benefits of Ethics Seminars - Student responses

Responses in the first category fit directly into the aims of the ethics course, to encourage moral sensitivity and give opportunity to practice moral decision-making skills. 74% (100) of the students stated at least one of the elements belonging the first category in their response. Further the second category shows student learning from a communication view point. This was not an explicit aim of the course, but a benefit all the same. Only one student saying that he/she learnt nothing is also very positive. Also 23% of students explicitly asking for more sessions suggests that students have found these sessions beneficial to them.

At the end of the session, students also gave a score for themselves, the group and the facilitator on how well they thought each had done during the session. The students' self-assessment average was 3.6, their group appreciation 4.2 and the facilitator average of 4.5; all out of five.

9.7.2 Summary

The scientific misconduct and integrity exercise has proven to be interesting and successful. Students participated actively and often reflected on the impact of the discussed issues on themselves. All students reported learning gains relevant to the learning objectives of the exercise. The data collected during the exercise also gives confidence that students appreciate the ethics sessions and the majority recognised learning gains in the areas designed to support their development.

9.8 Summary of course material and student responses

The evidence supports the conclusion that the course material used in the ethics exercises studied here has supported the growth of moral sensitivity in students. The student responses at the start of each session reflected the themes of the assigned reading and the students adopted some of the philosophical language used in reading material. I believe it was important to have dedicated reading for each exercise: this provided a common ground for the discussion and focused student attention on the relevant issues of the problem. From these results I have gained confidence that the chosen material fulfilled this task.

Students' self-declared learning gains matched well or very well with the learning objectives for each exercise. In very few instances, students reported learning that was irrelevant to the course and in some instance the reported learning was relevant though it did not correspond with the explicit learning objectives. Most of the stated learning gains were in moral sensitivity - students reporting increased awareness and understanding of the complexity of the issues. There were also clear indications of students having learnt significant issues about moral decision-making procedures during the sessions.

Most students left the sessions with a clearer understanding of the difficult issues. The unanswered questions they identified are some of the core ethical concerns in each subject area. The hope is that the identification of these questions could prolong their involvement with the subject matter.

After the last exercise, students described the benefits they had gained from the ethics exercises in general. The vast majority identified increased moral sensitivity as the main benefit of the exercises. Students also identified the importance of discussion skills, and an opportunity to consider their personal values. This is supportive evidence that the students perceived the gains of the ethics exercises to be the ones motivating their inclusion in the degree courses.

10. PBL experience and learning logs

PBLs were used in ethics teaching at IBLS during the 1999-2000 academic year for the first time. No academic literature could be found to describe the use of PBL in ethics teaching, and thus this approach may have no precedent. The aim of the ethics PBLs was to support students' moral development: to stimulate their moral sensitivity and to allow them an opportunity to practise their moral decision-making skills.

In IBLS, PBL was chosen as the method of ethics teaching in two Honours courses (Microbiology, and Physiology and Sport Science). In the new joint degree of Microbiology and Parasitology, the Microbiology students participated in two ethics PBLs (one in each term), while Parasitology students worked at the same time on scientific PBLs. Both PBLs consisted of 5 contact hours and 10 hours of independent work.

The Ethics PBLs were designed to give students an opportunity to explore ethical issues in context with scientific decision-making. The two Microbiology PBL exercises covered two areas of vaccine development: 1) the choice of research topics (TB or Meningitis B), and 2) the design of an ethically-sound testing protocol involving both animals and humans (see Appendix XXI - Microbiology PBL).

In the Physiology and Sport Science degree the 180 L3 students were divided into two groups, one working on an ethics PBL and the other on a scientific one. The ethics groups (6) were involved in a problem dealing with drugs in sport: 1) investigating the reasons why they are banned, and 2) deciding on a suitable punishment in two hypothetical cases. The PBL consisted of 3 contact hours and 10 hours of independent work (see Appendix XXII - Sport Science PBL).

The problems were designed so that a good answer could not be achieved without the inclusion of ethical considerations. The PBLs in both Honours

courses were assessed by students completing learning logs (see Appendix XXIII
- Introduction to Learning Logs).

It should also be noted that the ethics PBL was the first experience of PBL for
these students. Thus, the students required possibly more support than an ideal
PBL design would suggest. For both courses, this meant a preliminary session on
how PBLs work and an explanation of the assessment method in detail. For both
groups a resource file was also collected which included relevant articles to solve
the problem and a web-site was created with a large list of links to relevant
Internet sites.

The Microbiology PBL in Term 1 started with a quick revision of the PBL
format and then a reminder that the facilitator was not there to lecture, but to
support their learning. After this the students were given their problem and
encouraged to start the work with a brainstorming session. All three groups
adopted this approach and much of the first 2-hour session was taken up by this
and a preliminary division of labour. Two groups out of three had quite severe
difficulties in agreeing on the division of labour and how to group tasks, which
suggests that PBLs offer a great opportunity for learning team work and how to
organise research.

The second session (1hr) was a week later and the groups had agreed to come to
this session with more information in order to have a clear idea what still needed
to be done. With all three groups, students had prepared for the session relatively
poorly and the groups were predominantly still in the starting position or found
themselves to have mainly looked at the same issues, and the benefit of team
working on a problem had not materialised. The session ended with a further
division of labour.

The last session (2hr) followed two weeks later. The group knew that they
needed to reach an agreement on a vaccine to develop. In most groups the
students came into the last session with a clear view on which vaccine to choose.
When questioned further, in all groups the decision-making process needed to be

re-traced and the students only then realised the magnitude of issues they could consider before making a decision. The importance of facilitation was thus highlighted.

In term 2 the same groups continued to work together in designing a testing protocol for the vaccine they had chosen in term 1 (all groups chose TB). The number and length of the sessions was the same as in Term 1.The first session started with a return of learning logs. Students were given samples of 'A' and 'D' entries and the group then discussed how to improve on their logs and on the group work for the second term. Groups 1 and 2 worked very well together, much the same way as in Term 1. The main change in the group work strategy was a decision that everyone would provide a hand-out at the start of each session in order to share their learning and to reduce duplicating each other's work. These hand-outs were of excellent quality and they improved the team-work aspect of the second PBL immensely. The decision-making in this second PBL was a very lively session for both groups 1 and 2, and the students got into discussing some complex and interesting ethical issues in a very mature manner. Group 3, on the other hand, had serious problems, primarily due to poor attendance - I never met more than 3 out of 8 students at one time. In the end, communication with the group members was carried out by e-mail and most students worked individually rather than as a group.

The Physiology and Sport Science students started their PBL in much the same way as the Microbiology students, though all of their sessions were 1hr in length. The first session was brain-storming and dividing labour. Most groups (6) divided work based on the different drugs mentioned in the problem, and the elements in the coach's statement. The groups were lively and well-organised. One of the groups also asked for a sample of a good learning log entry, which I wrote and then distributed to all groups; the constitution of a good entry was discussed with all groups during the second meeting. In the second and last sessions, the groups often jumped into conclusions without much discussion, and interventions by the facilitator were necessary in order to guide the students to consider all elements in the problem. During the last session each group handed

in a decision on the hypothetical cases provided. The decisions between groups varied widely, while all of them used much the same arguments to support their conclusions.

10.1 Learning logs

90 students handed in a learning log, 21 in microbiology (two logs each) and 69 in Physiology and Sport Science. Most logs contained 10 entries in total. The instruction was for the students to include four elements in each entry:

1. what I did (1-2 lines)
2. what I learnt (2-3 lines)
3. was it important and why (10-20 lines)
4. how this shapes my next learning unit (1-3 lines)

The assessment criteria were based on the ability to reflect on individual learning (element 3).

I included two (4th and second last) entries from each log in order to analyse the widest possible spread of the 900 entries I had collected.

Figure 4 contains a sample of three complete log entries from the Microbiology term 1 learning logs; one from a log given an 'A' and two from logs given a 'D'.

ENTRY FROM A LOG THAT GOT AN A:

"In this learning unit, I used mainly the Internet search engines (Netscape and Excite) to gather sociological information and statistics of Tuberculosis and Meningitis. To my surprise, much of the sociology does involve developing countries due to many complex reasons as I later found out. The statistics, which I obtained, implies that death rates are not most highest for developing countries for both diseases and much disease occurs in the poorest resourced areas of the world, such as, parts of Africa and Asia. It was no surprise when I found that 'lower' social classes across the UK were more likely to have disease than the 'upper class' because I did hypothesise this in my earlier learning unit.

Only when reading information from reliable sources on the Web (WHO and UNICEF) that I discovered that the social factors for any major disease is extremely complex especially in developing countries. Some are large-scaled sociological factors (e.g. poor economics) whilst others are smaller but can still have a significant impact (e.g. hygiene). Every country has different behavioural attitudes, religious beliefs and ways of

---

life and can complicate sociological factors further. Why is this an important issue? It is important because the structure for vaccine implementation will be different for different countries and the demands for vaccines will vary for each country depending on which disease is more prominent.

This learning unit was most meaningful and I learned a great deal more on sociological issues in developed and developing countries mainly because I had much more information to back up previous assumptions and I had facts to compare and contrast with. In addition, reliable web pages from the Internet are one of the best sources to get hold of up-to-date information. The Internet certainly helped me in this learning unit. However, I had hoped for more information on developed countries.

I managed to print off lots of information too and I would like to analyse them further for other potential sociological factors of importance for both developed and developing countries."

ENTRIES FROM LOGS THAT GOT A **D-**

"Use Internet to look into relevant sites. GlaxoWellcome and Beecham."

"Today we received the problem that we will be looking at over the next few weeks. After some discussion we divided the problem into different subsections. These subsections will then need to be researched to find out information."

Figure 4: Learning log sample entries

In analysing the entries I first collected the following data

- the length of the log (lines, most of them are typed)

- identify the following elements: description (elements 1 and 2), reflection (element 3), and projection (element 4) and their line lengths

- the mark the log received (the logs were marked by the facilitator Henriikka Clarkeburn). The mark was part of the students' continual assessment.

A more detailed analysis was carried out on the reflective elements. There were three main reflective categories:

1. decision-making contemplation; e.g. "I think this was important because in an hour, what we learnt gave a solid foundation to what we already knew, that we should go ahead and recommend the TB vaccine. The charity is not likely to want to sponsor a vaccine that is going to be made redundant in a few months/years."

2. learning skills and habits; e.g. "I thought TB would be the best vaccine to make and therefore my next piece of learning was influenced towards TB. This I think was wrong of me as I should have been more open-minded when I searched for information, as when people are influenced in one direction people tend to look for information on that side of the argument and not the other side of the argument".

3. ethical considerations; e.g. " When I was reading the book I realised that the number of people who were suffering from each disease was not as relevant in choosing a vaccine any more as each one of those numbers was a single person with a real life."

In order to gain more detailed understanding of the ethical reflections, I developed the three-level ethical scale below (the quotes in level 1 and 2 are from Microbiology student logs, and the level 3 quote is a sample written to demonstrate the content of level three entries):

Level 1: Basic recognition of ethical issues - e.g. "the number of people dying is an issue".

Level 2: Representation of basic ability to reflect on ethical issues - e.g. "the number of people who were suffering from each disease was not as relevant in choosing a vaccine any more as each one of those numbers was a single person with a real life".

Level 3: Mature reflection and personal account in relation to the ethical issues - e.g. "we can't look at just number of people we can save, because to us it is not the same moral issue whether we see one of our loved ones die or someone we don't know. We might have more demanding moral duties to save those who we are responsible for. At the same time it would be morally wrong to save only one human if you could save 10 more as well. Also it makes a difference if we are talking about identified existing people or hypothetical/unknown people. If we are talking about the latter kind, then I believe our duty is to try to save as many as possible."

10.2 Analysis of learning logs

The average length of each entry for hand written learning logs (n=30) were 17 lines and for typed logs (n=60) 16. Because the average lengths of both hand-written and typed entries were so similar, further analysis considered the whole group together, even though the lengths represent slightly different actual word lengths.
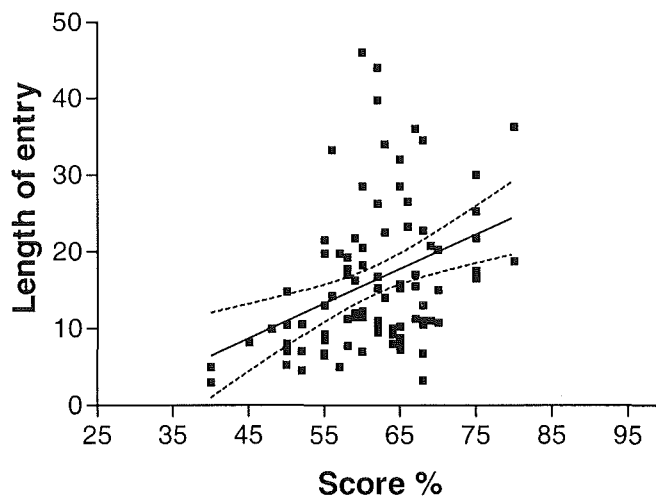
Figure 5: Relationship between length of log entry and score

The length of a log entry does have a significant relationship with the score. The graph above shows a linear regression line statistically different from zero (p=0.0004). It is worth noting though, that the length alone cannot explain the score, as can be seen in some very long entries scoring less than 60%. The vast majority of the logs were well written, had a balanced structure, and dealt with issues relevant to the PBL. The length is therefore best understood as a sign of dedication to the log, and thus it is to be expected that the length and score have a positive correlation.

Most entries had elements in 1-3 different categories, mostly at least one descriptive and one reflective. Table 36 shows proportionate length (in percentage) of each entry type for the Physiology and Sport Science, and Microbiology learning logs for term 1.

	Average proportion of entry (%)					
	Descriptive	Reflective-Decision-making	Reflective-Learning skills	Projective	Reflective - Ethics level 1	Reflective - Ethics level 2
4th entry (n= 90)	47.3	34.3	3.4	7.7	0.6	3.1
2nd last entry (n=90)	52.1	29.9	12.7	1.0	1.2	5.9

Table 35: Term1 learning logs: proportional lengths of different elements

The first notable element in these type lengths is the dominance of descriptive elements. It was emphasised in the introduction to learning logs and during the sessions that the logs should be an account of their thoughts, not what they had done. Still, approximately half of each entry comprised descriptive elements. There were also 12 students whose 4th entry was completely descriptive and 19 whose 2nd last entry had no other element types.

The second noticeable quality of these entries is the very small proportion of ethics reflection, including a complete lack of level 3 ethics reflection in all entries.

Table 36, for the Microbiology Term 2 logs (n=21), shows small but important differences:

	Average proportion of entry (%)					
	Descriptive	Reflective-Decision-making	Reflective-Learning skills	Projective	Reflective - Ethics level 1	Reflective - Ethics level 2
4th entry	36.0	31.5	14.0	6.3	5.9	6.3
2nd last entry	29.4	37.3	10.6	2.6	8.1	9.0

Table 36: Term 2 (Microbiology) learning logs: proportional lengths of different elements

The dominance of descriptive elements in the entries has been reduced, though it is still high in the light of the instructions given to students. The proportionate weight of ethics reflection has also increased in their second logs. The proportion

is still low, but the practice with this type of work seems to have increased both the students' ability to consider and write about ethical issues.

10.2.1 Summary of PBL experience

As experienced in the PBLs in this study, undergraduate science students have relatively low competence levels in using ethical tools in problem-solving. This has become apparent also in the discussion groups carried out with Level 3 students in other Honours courses. Students do not 'see' moral problems and they do not prioritise them in their decision-making. Thus, when allowed to approach problems on their own, students have a tendency to ignore sometimes even potent ethical problems facing them. For example, Microbiology students can without a moment's hesitation suggest that an ideal research protocol for a vaccine includes the use of over 500 primates in the initial animal trial phase. Not only is this financially crippling (resource allocation issues), but also ethically questionable in the magnitude of pain caused to animals. The students did not question the justification for using animals, unless directly questioned and even then the first knee-jerk reaction was that people who support animal rights are terrorists and use of animals is essential and right. Therefore, the facilitator participation in these PBLs was probably more active and directive than an ideal PBL design would suggest.

The learning logs reflect a relative lack of interest or ability to engage in ethical considerations. The high proportion of descriptive elements in the entries could have been a result of at least two things: 1) the instructions were not clear enough, or 2) the students had genuine difficulties in reflecting on their learning. I believe the problem was rather in the latter. The main benefit of the learning log and the feed-back the students received would have been in the realisation of the nature and difficulty of reflection on learning. A learning process can be observed in the proportional differences between Term 1 and Term 2 logs. In their second log, the students were more focused on reflection and increased the proportion of their log discussing ethical problems.

PBL has considerable potential as a good form of ethics teaching, but it is not without problems. First, it is probably not ideal to introduce students to both PBL and ethics at the same time. The learning of both a new method of study and team work, and a completely new subject may require students to choose where to focus their energies and attention. Judging from this experience, the energy is directed at mastering the PBL approach to learning, rather than at ethics. Dr Mary Tatner, who taught the science-based PBL to parasitology students (the control group), reported a similar experience. During the first PBL in term 1, students spent a considerable amount of their time in learning the method and gave the problem itself less of their time and attention.

If students are familiar with PBL and have had at least some very preliminary introduction to the nature of philosophical inquiry, I believe that PBL offers an excellent method for ethics teaching. It allows students to explore the issues independently and thus not pose problems with indoctrination and could enhance the understanding of ethics in difficult decision-making far better than discussion groups.

11. Analysis of the Moral Development Questionnaire

The Moral Development Questionnaire was designed to detect three distinct elements in moral development: moral sensitivity, moral cognitive skills, and meta-ethical conceptions. The results of the questionnaire, applied before and after the ethics education intervention, were expected to identify progression and regression in all of these three developmental elements. The effects of ethics teaching were the main focus of the analysis of these results.

The Moral Development Questionnaire was administered to each L3 Honours course during the first few weeks of Term 1 and then again either at the end of term 1 or at the start of Term 2, which ever was most applicable considering the teaching schedule. Approximately 450 students in the L3 Honours courses participating in ethics teaching were expected to fill in the questionnaire. Due to scheduling difficulties in organising a time for the questionnaire to be filled in and to variable student attendance at these sessions, the actual number of students filling in either the pre or post questionnaire was considerably lower (374 for the pre and 301 for the post questionnaire). The numbers were further reduced after the students filling in only one of the questionnaires were removed from the analysis (242 for DIT).

This reduction in number of questionnaires included in the analysis was unfortunate. The original plan to compare results between Honours courses had to be abandoned in most cases as the group sizes were too small. Also, some of the gender-based analysis carried out suffered from small group sizes.

All that could be done, was done in organising the questionnaire times as suitably as possible for the students. Variable attendance is a common phenomenon throughout the course structure, so the lack of students attending these sessions was not out of the ordinary. However, the fact that for most students, the ethics component was not part of the course assessment may have influenced the students' motivation to attend these sessions.

The main statistical test used to analyse the Moral Development Questionnaire data was the Wilcoxon paired sample t test. This is a nonparametric analogue to the paired-sample t test. The Wilcoxon t test is applicable whenever a paired t test is applicable, but also when we cannot assume normal distribution. The testing procedure involves the calculation of differences, which are then ranked according to absolute values of differences, from low to high (the smallest difference being assigned a rank of 1, the second smallest two etc.). Then the ranks with a plus sign are summed and those with a minus sign are summed. The higher raw value of rank sums indicates the direction of change. However, this is only indicative as different distribution patterns in the ranks can strongly influence the rank sums. The Wilcoxon test identifies differences in samples even when the mean differences are very small by comparing each pair individually (Zar,1999).

Other statistical tests used were an unpaired t test with Welch's correction and the Mann-Whitney test. Welch's correction takes into consideration that the two data sets may not have equal variances, when Mann-Whitney does not assume normal distribution. All data were tested for normality and variance before an unpaired t test was carried out. Normality was tested with Kolmogorov-Smirnov test. Pearson correlation coefficient was also employed.

As will become apparent in this chapter, these data have a high level of noise. Significant numbers of students both progress and regress in all measures. This leads to very small changes in the mean scores. Also, emerging trends are only moderate due to the movement of scores in both directions. This made the analysis complicated and created challenges in drawing firm conclusions.

This chapter looks at each of the three elements in the Moral Development Questionnaire separately first and then compares the results at the end.

The first element looked at is the Moral Sensitivity test, section 11.1. This is followed by DIT results in section 11.2. The DIT results are divided into six main sub-sections:

  11.2.1 P-score,

  11.2.2 N2-score,

  11.2.3 Gender variance in P- and N2-score,

  11.2.4 Type-score,

  11.2.5 U-score, and

  11.2.6 Gender variance in Type and U-score

The last element in the Moral Development Questionnaire is Perry development, results of which are detailed in section 11.3; the gender variance in Perry is summarised in sub-section 11.3.1.

The last section (11.4) in this chapter describes the relationships between all these three elements.

The Chapter ends in a summary of the results.

11.1 Moral Sensitivity results

The Moral Sensitivity Test measures development in the first element of moral decision-making - the ability to recognise moral problems. Without this recognition there can be no further solving of a problem.

The first part of the Moral Development Questionnaire was a Moral Sensitivity Test. Students were asked to write no more than five questions or issues they believed should be considered before making a decision on whether to start research on a genetically modified cow producing pharmaceutical milk to treat Cystic Fibrosis (see Appendix IV - Moral Development Questionnaire). The scoring guide for these responses is discussed in more detail in section 5.3.2.2 (Scoring method) and the guide itself can be found in Appendix X - Scoring Guide.

11.1.1 Moral Sensitivity response frequencies and averages

Student response forms (protocols) were included in the analysis if there was at least one issue or question raised. Protocols left blank were removed from the analysis, because it was not possible to know whether students had not recognised any issues (genuine blank) in the proposal or whether they had skipped the Moral Sensitivity Test altogether (artificial blank). Table 37 shows the numbers of protocols in each step of inclusion.

Moral Sensitivity Protocols	pre	post
All returned	374	301
At least one response	367	274
Paired	267	267

Table 37: Number of Moral Sensitivity Protocols

Table 38 details group sizes of included protocols both in relation to the male/female divide, and different teaching approaches.

		Teaching Approach		Gender	
	All	PBL	Discussion	Female	Male
Test	**133**	48	85	91	42
Control	**134**	49	85	98	36

Table 38: Gender and teaching approach in Moral Sensitivity

The mean number of questions or issues raised was approximately 3.2 in both pre and post questionnaire and for both Test and Control groups. Table 39 summarises the information on the mean number of issues raised in the pre and post Moral Sensitivity Tests.

Group	Mean number of responses	
	pre	post
Overall (n= 267)	3.21	3.22
Test (n=133)	3.15	3.23
Control (n=134)	3.28	3.21

Table 39: Moral Sensitivity test mean number of responses

If issues scoring zero (non-ethical responses, for further details see Appendix X - Scoring Guide) are removed from the frequency analysis, we find some differences between the mean number of issues/questions raised between Test and Control group pre and post tests. Table 40 summarises the mean numbers of responses after the exclusion of responses scoring zero.

Group	Mean number of responses	
	pre	post
Overall (n=267)	2.87	2.99
Test (n=133)	2.83	3.08
Control (n=134)	2.91	2.91

Table 40: Moral Sensitivity Test: mean number of responses scored above zero

The issues and questions scoring a zero are not ethical considerations. Thus a student could raise five issues/questions without showing any ethical sensitivity.

By removing the zero scoring responses, we get a more accurate understanding of the increase/decrease of *ethical* questions/issues raised by students.

The mean number of responses after zero-scoring responses were removed, given by Test and Control groups, was not statistically different (p=0.639 in the pre test and p=0.303 in the post test, unpaired t test with Welch's correction). The increase in the mean number of responses in the Test group was cancelled out due to high variation in the number of responses in both Test and Control group.

In both Test and Control group, there were significant differences in individual students' number of responses in the pre and post test (p<0.0001, Wilcoxon t test). In the Control group, the ranked changes in the number of responses cancelled each other out, while in the Test group the increase of responses was on average larger than the decrease.

The scoring guide for the Moral Sensitivity Test identified four logically independent types of responses:

I Risks
  - to human health
  - to animals
  - issues of supervision
  - testing and labelling products

II Cost and Benefit + Research issues
  - medical benefits to humans
  - opportunity costs in research
  - commercial involvement and access to treatment

III Issues of Basic Values
  - genetic research
  - animal rights

IV Public Opinion

Table 41 details the mean number of responses in each type. Responses scoring 0 (see Appendix X - Scoring Guide for details) are included in this table, while duplicate responses (more than one issue/question raised in the same sub-type) have been removed before means were calculated. The responses scoring zero were included in order to give an accurate picture of the types of issues students

recognise. It is worth noting, however, that types III and IV have no zero scoring elements.

Group	Types			
Pre	I	II	III	IV
Overall (n=267)	1.38	1.35	0.30	0.19
Test (n=133)	1.36	1.32	0.32	0.15
Control (n=134)	1.39	1.38	0.28	0.23
Post	I	II	III	IV
Overall (n=267)	1.34	1.48	0.25	0.16
Test (n=133)	1.33	1.44	0.31	0.15
Control (n=134)	1.35	1.53	0.17	0.17

Table 41: Mean response frequencies in each type

The number of responses in each category was very similar in Test and Control groups in both pre and post test. When the pre/post answering patterns are compared for each student there are no statistically differing patterns ($p=0.879$, paired t test). The biggest percentage difference is in the decline of mean number of responses by the Control group in type III.

11.1.2 Moral Sensitivity scores

The Moral Sensitivity Test was scored using the scoring guide (Appendix X - Scoring Guide). The level 0 answers did not accrue any points, level 1 answers were given 1 point, level 2 answers 2 points and level 3 answers 3 points. If a student had more than one answer in one sub-type, the highest scoring element was the only one included in the score.

The pre score mean for the entire group was 4.780. For the Control group the pre score mean was 4.890 and for the Test group 4.681. The difference between Control and Test group pre-scores was not statistically significant when tested using an unpaired t test with Welch's correction ($p=0.443$).

Table 42 summarises information on the direction of change in the Moral sensitivity Test scores in Test and Control groups, pre and post teaching.

Group	Direction of change (%)		
	Progressing	Regressing	No change
Test All (n=133)	**51.8**	**33.3**	**14.9**
Discussion (n=85)	49.4	32.9	17.7
PBL (n=48)	54.2	37.5	8.3
Control All (n=134)	**31.5**	**44.9**	**23.6**
Discussion (n=85)	31.8	48.2	20.0
PBL (n=49)	27.7	46.8	25.5

Table 42: Moral sensitivity progression/regression

It is noticeable that a large percentage of students progressed and regressed during the test period in both Test and Control groups. However, the Test group has a higher percentage of progressing students and a lower number of regressing students than the Control group. Also, a higher percentage of Control group students experienced no change between the pre and post Moral Sensitivity Questionnaire. The numbers of male and female students regressing/progressing in the Control and Test group were not significantly different from percentages for the whole group. When the change direction patterns between Control and Test groups are compared using chi-square, the patterns appear significantly different (p<0.0001, X^2=24.941, df=2).

Even though there were no great differences in the direction of change between Discussion and PBL groups, the trends are still interesting. Students in the PBL Test group progress more than students in the Test Discussion group, while fewer students in the Control PBL group progress than students in the Discussion Control group. The Test PBL group has also the smallest percentage of students not experiencing any change.

The fluctuations of students in both Test and Control group make further interpretation of moral sensitivity scores tentative. The level of noise may mask some significant differences between sub-groups and the other way round; significant changes may result from a coincidence in high noise data. However, the directions of change indicate that ethics education has encouraged positive

change in the Moral Sensitivity test in comparison to the encouragement provided by the academic experience in general.

Figure 6 shows the relationship between the Moral Sensitivity pre score and the size and direction of the change between pre and post questionnaire for both Test and Control groups.
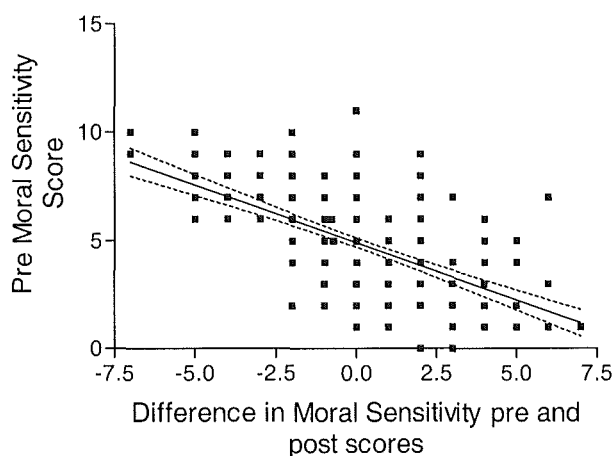
Figure 6: The linear regression between Moral Sensitivity pre scores and the pre/post difference

The linear regression is highly significant ($p<0.0001$, $r^2=0.362$). The higher the pre-score the more likely the student to regress and the lower the pre score, the more likely progression becomes. This is logical as those with a higher pre-score have more 'room' to regress than those with a low pre-score, and those with a low pre-score have more 'room' to progress. What is noticeable, however is that the Moral Sensitivity score is not a particularly static variable and students scoring high at one testing occasion may not do so next time. It would be interesting to measure students at several points and try to seek a truer estimate of their moral sensitivity and identify the issues that contribute to the variability.

Table 43 summarises the mean Moral Sensitivity Scores for both Test and Control groups in the pre and post questionnaires.

Group	Score (mean ± SD)	
	Pre	Post
Test (n=133)	4.68 ± 2.27	5.30 ± 2.25
Control (n=134)	4.89 ± 2.18	4.67 ± 1.95

Table 43: Moral Sensitivity scores

In both the Test and Control groups the pre and post sensitivity scores were found to be significantly different ($p < 0.001$, Wilcoxon t test), which again shows the nature of the measure. For the Test group the difference between sums of positive and negative ranks was 2112 (2574 and - 4686). This indicates that the larger changes in the group were regressive and the smaller changes progressive. The smaller number of regressive students can influence the Wilcoxon sum of ranks greatly in a large sample size. Therefore, considering a) the percentages of students progressing, regressing and experiencing no change (Table 42: Moral sensitivity progression/regression; b) the change in mean Moral Sensitivity score; and c) the Wilcoxon t test, I conclude that the overall change in the Test group was progressive, while at the same time the regressive differences were on average larger than the progressive differences.

For the Control group the difference between sums of positive and negative ranks was 576 (2665 and -2089). The positive sum of ranks being larger indicates bigger differences in the progressive sub-group. However, considering the percentage of students progressing, regressing and experiencing no change; the change in mean moral sensitivity score; and the Wilcoxon t test, I would conclude that the overall change in the Control group was regressive, while at the same time the progressive differences were on average larger than the regressive differences.

11.1.3 Gender differences in Moral Sensitivity

Gender is one possible variable in moral sensitivity. This section identifies the role gender plays in the Moral Sensitivity scores and moral sensitivity

development of the Glasgow University student sample participating in a minimal ethics education intervention.

Table 44: Moral Sensitivity scores by gender in the pre and post questionnaire Moral Sensitivity score means for female and male students in Test and Control groups.

Group	Score (mean ± SD)	
	Pre	Post
Female all (n=189)	4.86 ± 2.21	5.08 ± 2.18
Male all (n=78)	4.62 ± 2.27	4.77 ± 1.91
Test (n=133)	**4.68 ± 2.27**	**5.30 ± 2.25**
Female (n=91)	4.96 ± 2.26	5.34 ± 2.34
Male (n=42)	4.18 ± 2.23	5.22 ± 2.12
Control (n=134)	**4.89 ± 2.18**	**4.67 ± 1.95**
Female (n=98)	4.76 ± 2.17	4.78 ± 2.11
Male (n=36)	5.22 ± 2.21	4.39 ± 1.48

Table 44: Moral Sensitivity scores by gender

The pre and post score mean differences between male and female students (Test and Control groups combined) were not statistically significant (pre $p=0.414$ and post $p=0.780$, Mann Whitney unpaired t test). There no gender differences within Test and Control groups either (Test Pre scores $p=0.058$, Test Post scores $p=0.599$, Control Pre scores $p=0.323$, and Control Post scores $p=0.350$; Mann-Whitney unpaired t test).

Table 45 details the percentage of female and male students either progressing, regressing or experiencing no change in Moral Sensitivity score.

Group	Direction of change (%)		
	Progressing	Regressing	No change
Test All (n=133)	**51.8**	**33.3**	**14.9**
Female (n=91)	48.3	37.4	14.3
Male (n=42)	57.1	28.6	14.3
Control All (n=134)	**31.5**	**44.9**	**23.6**
Female (n=98)	31.6	48.0	20.4
Male (n=36)	25.0	47.2	27.8

Table 45: Moral Sensitivity direction of change by gender

When we look at the data from both tables (Table 43 and Table 44), the pre and post scores for the female students in the Control group are significantly different (p<0.0001, Wilcoxon t test), but the difference between the sum of ranks is only 15 (1271 and -1286). Thus, the sizes of positive and negative change within the group were close to equal in both directions. The difference between means is also very moderate. Including the percentage of students experiencing positive change (31.6%), the general trend in the Control group females seems to be dominantly one of equal changes in both directions, with a higher percentage of students regressing than progressing.

The difference between pre and post Moral Sensitivity Test for the female students in the Test group was also highly significant (p<0.0001, Wilcoxon t test), and the difference of their median ranks was 519 (1281 and -1800). Also, almost half of the Test group female students progressed in the Moral Sensitivity score and a smaller number regressesed. Thus the general trend is clearly progressive.

For the male students in the Control group, the difference between pre and post Moral Sensitivity scores was again highly significant (p<0.0001, Wilcoxon t test), and the difference of sum of ranks was 161 (256 and -95). This indicates a progressive trend in the scores, or that the progressive differences were of a larger impact and thus ranked lower (higher rank sum) in the Wilcoxon t test. At the same time the means of pre and post Moral Sensitivity scores for males in the Control group regressed from 5.22 to 4.39. Also, only 25% of males in the control group progressed, compared with 47.% regressing and 28% experiencing

no change at all. Considering all these data, the conclusion is that the males in the Control group had a regressive trend.

The males in the Test group also had a highly significant difference between pre and post Moral Sensitivity Scores ($p<0.0001$, Wilcoxon t test). The difference in sum of ranks was -462 (221 and -682), which indicates larger negative than positive changes. This is contradicted by the increase of means from 4.18 to 5.22. Also, 57.1% of test group males progressed, with only 28.6% regressing and 14.3% experiencing no change. Considering all these data, I conclude that the male students in the test group had both statistically and genuinely significant gains in their moral sensitivity scores after their experience of moral education.

11.2 DIT results

DIT measures moral cognitive skills and the ability/desire to use justice-based reasoning in moral decision-making (for further details see section 4.3).

The Defining Issues Test has a set analysis approach explained in section 4.3.3. The data were extracted from the optically read questionnaires first into an Access file. In this format, the computed results that were either incomplete (blanks) or had multiple entries were compared with the paper originals and either corrected, included with a note or disregarded from further analysis.

Table 46 summarises the inclusion steps and numbers of questionnaires in each step:

DIT protocols	pre	post
Optical reading		
complete	215	147
blanks	129	141
multiples	30	12
All	374	300
Pairing	256	256
DIT validity	195	195

Table 46: DIT protocols included

After the questionnaires that could be completed (blanks) or corrected (multiples) had been dealt with, students were paired and only those that had completed both the pre and post questionnaire were included in the analysis. The next step was the DIT validation procedure outlined by Rest (1986). Out of the protocols rejected at this stage, 33 were rejected due to an unacceptably high M-score (M>7), 11 because of more than 9 same ratings in a story (indicating completion without due consideration of items), 2 because of P-score and N2-score inconsistencies, and 15 because of inconsistencies in the ranking and rating of data. Thus 23% of paired protocols did not pass at least one of the validity tests. This figure is higher than the quoted average rejection percentage of 10-15% (Rest et al., 1997).

Table 47 summarises the sizes of sub-groups in the included valid DIT protocols.

	All	Teaching Approach		Gender	
		PBL	Discussion	Female	Male
Test	114	47	67	73	41
Control	81	25	56	59	22

Table 47: Sex and teaching approach ratios of valid DIT protocols

11.2.1 P-score

P-score is the widest used and one of the oldest DIT scores computed (see section 4.3.3 for further detail). The P-scores I quote here were validated by Dr Stephen Thoma from the University of Alabama. Dr Thoma offered to double-check my calculations after my first two efforts produced unusual results. The results Dr Thoma calculated matched the original results.

The first unexpected P-score result was the low P-score mean in the pre-questionnaire. Table 48 shows the results.

Group	pre P-score (mean ± SD)
Control	30.44± 14.18
Test	32.89± 15.83

Table 48: Pre P-score means for Test and Control groups

These P-score averages for the Glasgow University Level 3 Bioscience students correspond with international averages for high school students, not with students already in tertiary education. The international average for tertiary students is 42.3 based on a sample of 2886 (Rest 1986). 92 (24.6%) students in my sample had a pre P-score higher than 42.3, thus exceeding international averages for students with similar educational back-ground. The pre P-scores of Test and Control group were not significantly different (p = 0.1755, unpaired t test with Welch's correction).

The second, and even more unusual, element of the P-scores was the high percentage of students regressing between pre and post questionnaire i.e. their P-score was lower in the post-questionnaire. Table 49 shows the percentage of students changing in each direction in P-scores.

Group	Direction of change (%)		
	Progressing	Regressing	No change
Test All (n=114)	**51.8**	**46.5**	**1.7**
Discussion (=67)	55.2	44.8	-
PBL (n=47)	46.8	50.0	4.2
Control All (n=81)	**59.3**	**39.5**	**1.2**
Discussion (n=56)	55.4	42.9	3.6
PBL (n=25)	68.0	32.0	-

Table 49: P-score - direction of change

There is no published precedent for such a high percentage of students regressing in the DIT. Regression is not recognised in the original Kohlberg theory, though regression has been detected in DIT protocols in longitudinal studies, but never to this extent (personal e-mail correspondence with Dr Thoma). In particular, it is to be noted that more students in the Test group regressed in comparison to the students in the Control group.

Figure 7 details the sizes of P-score changes for Test and Control groups. The changes are divided into eight bands and the bars represent the percentages of students in Test and Control groups for each change band.

Figure 7: P-score differences between Test and Control groups

The differences in P-score change between the Control and Test groups are not statistically significant (p=0.896, unpaired t test with Welch's correction), though their variances are significantly different (Pearson correlation coefficient p<0.05). The difference in variance indicates different types of movement between the groups which can also be detected in Figure 7. In general, the Test group students have a higher percentage of students at both the positive and negative end of the scale. Almost half of the students in both groups have a P-score change size less than 10.

Table 50 details the mean progression and regression sizes for Test and Control groups.

Group	P change size	
	Progressing	Regressing
Test (n=114)	15.10	-11.40
Control (n=81)	11.60	-10.10

Table 50: P-score change means

The magnitude of these changes, more than 10 points on average, is very large. In P-scores, changes of this magnitude often represent, on average, progress gained from 5 years of formal education. Though there are educational interventions that report changes of similar size (Self and Ellison, 1998), change

of this magnitude was not expected from a minimal intervention (Schlaefli *et al.*, 1984).

Considering the number of students progressing and regressing between pre and post DIT, it is not surprising that the Wilcoxon t test for the pre and post tests for both Test and Control groups shows a highly significant difference (p<0.0001). The comparisons between positive and negative rank sums between Control and Test groups are very similar, indicating similar magnitudes of change in both directions. Also, there are no significant differences between Test and Control group Post P-scores (p=0.3359, unpaired t test with Welch's correction). Therefore, taking into consideration both the high variance in both Test and Control group, and the lack of significant P-score differences in the post scores, I conclude that the ethics intervention produced no noticeable effect on students moral decision-making skills when measured by DIT P-score.

11.2.2 N2-score

A newer DIT score is N2 (for more details refer to section 4.3.3). It has been found to detect similar developmental tendencies to the P-score, while it is more sensitive towards changes after educational interventions (Rest *et al.* 1997).

The N2-score follows the same trends as the P-score in relation to students regressing and progressing between the pre and post questionnaires. In only 17 protocols, the direction of change was different between P- and N2-scores. Out of these, 13 had P-score changes less than 0.1, proposing no difficulty in accepting the difference. In the remaining four there were two major shifts (change difference >10), one from a P-score difference of -9.97 compared with a N2 difference of +1.06 and another from P-score difference +26.7 compared with a N2 difference of -6.21; the other two protocols had a change difference less than 3. The two highly irregular protocols were removed from further analysis on account of inconsistencies in the original data.

The average Pre N2-score for the Control group was 30.64 and for the Test group 34.13. This is again lower than international averages for university students - N2 scores have similar averages to P-scores and thus the N2 international averages for tertiary students are in the low 40s (personal correspondence with Dr S. Thoma). The Test and Control N2 pre-scores were not statistically different (p=0.90, unpaired t test with Welch's correction).

Figure 8 details the sizes of N2-score changes for Test and Control groups. The changes are divided into eight bands and the bars represent percentage of students in Test and Control groups for each change band.
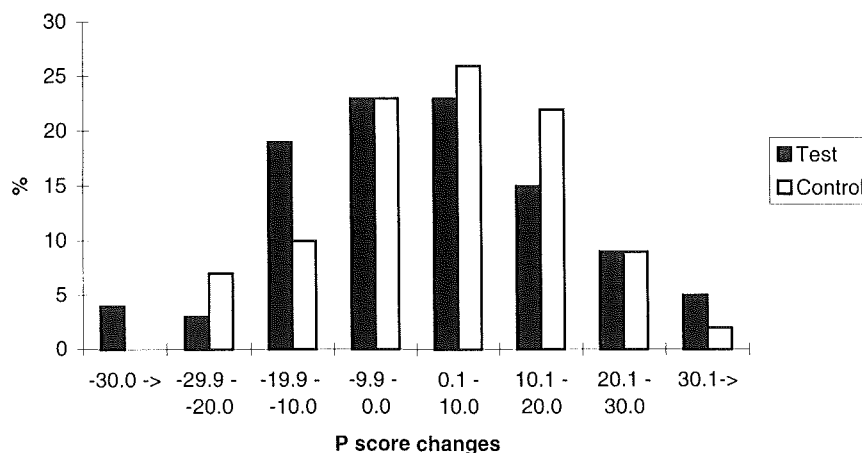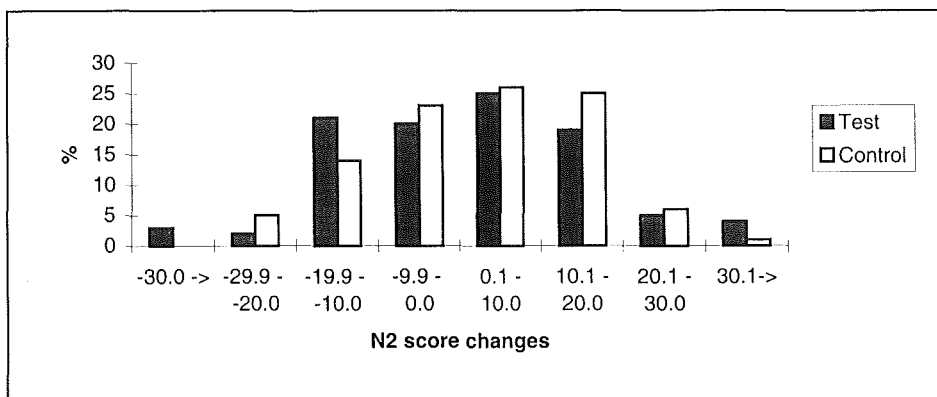

Figure 8: N2-score difference in Test and Control groups

The difference between Test and Control group N2 changes (before and after the ethics intervention) is not statistically significant (p=0.514, unpaired t test with Welch's correction). The N2 scores resemble P-scores in there being proportionally more Test group students at the far ends of the scale and in that most students experience more moderate change.

Table 51 details the N2 mean size of change for regressing and progressing students separately for Test and Control groups.

Group	N2 mean change size		
	Overall	Progressing	Regressing
Test (n=114)	1.63	13.00	-11.93
Control (n=81)	2.99	12.20	-9.75

Table 51: N2 mean change size

The overall mean change size for N2 is very small, when compared with pre-post change sizes reported in Rest *et al.*, (1997). There the mean change size (including both regressive and progressive students) is 13.94 (n=502). The Rest *et al.* study does not specify regressive and progressive means, but as regression is rare in DIT, it seems safe to assume that the low mean N2 score for the Glasgow University sample is due to the large number of regressing students.

Taking into the consideration the large movement between pre and post N2 scores, it was to be expected that the Wilcoxon t-test would confirm the difference between pre and post scores as highly significant for both groups (p<0.0001). Also, the sum of positive and negative ranks was very similar between Test and Control groups.

The pre score was identified as a possible predictor of DIT change size and direction between pre and post questionnaires. In order to find out whether the pre-score could be used as an indicator of this sort, linear regression was carried out (Figure 9).

Test N2
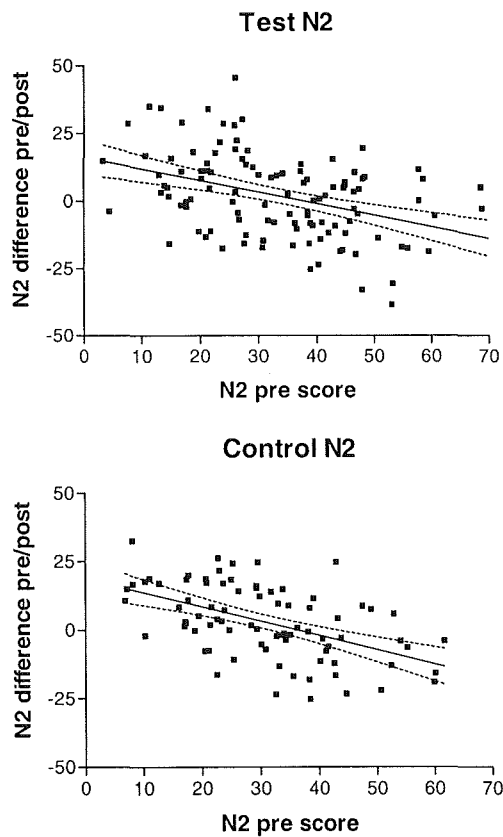
Control N2

Figure 9: N2 pre/difference regression

The linear regressions for both Test and Control groups are statistically significant (p<0.0001). The regression was also statistically linear with the residual p value >.05. Still, even though the regression is significant the power of this prediction is relatively small (Control r^2=0.268 and Test r^2=0.178).

These regressions indicate that we may not expect significant change in N2 - scores from students whose N2 pre-score falls close to the average score for the group (35-45). A more significant regression or progression can be, on the other hand, expected from the subjects scoring either lower than average in the pre-questionnaire (progression) or higher (regression). This seems only logical, as the students scoring low can only either progress or experience no change, while the students scoring high have an option either to maintain their score levels or to regress. This might indicate, in the light of regression being relatively rare in DIT, that the high pre-scores in this population might have been coincidental rather than true indications of a higher level of moral development.

.

The post N2 scores were not significantly different between Test and Control groups (p=0.3218, unpaired t test with Welch's correction). Considering this and the high number of students regressing and progressing in both groups, I conclude that when measured by DIT N2, no difference could be detected in students' moral decision-making skills in relation to whether they attended an ethics course or not.

11.2.3 Gender variance in P- and N2-score

There has been a lot of controversy about possible gender bias of DIT. This is one reason to look into the DIT results from a gender specific point of view. Also, it is important to find out whether genders benefit differently from this type of ethics teaching.

Table 52 summarises P and N2 pre scores for both genders.

Group	Mean scores	
	P	N2
Pre		
Female (n=132)	34.20	36.39
Male (n=63)	27.56	28.42
Post		
Female (n=132)	36.32	36.35
Male (n=63)	31.19	31.78

Table 52: P and N2 mean scores by gender

The differences between male and female students' pre scores were statistically significant for both P and N2 (p<0.005, unpaired t test with Welch's correction), with female students scoring higher in the pre test. This does not support the view that DIT discriminates against female-specific decision-making methods, or that females would be less capable in using justice-based moral decision-making.

In the post scores the difference between genders, when measured as a unified group, was reduced in both measures. For the P-score the difference is

statistically significant at p<0.05 (unpaired t test with Welch's correction), but the N2 score difference is just below statistical significance (p=0.058, unpaired t test with Welch's correction).

Table 53 summarises the gender differences in P and N2 scores separately for Test and Control groups. The asterisk in the table indicates significance: * p<0.05 and ** p<0.01.

Groups	P/Control	P/Test	N2/Control	N2/Test
Pre				
Female	**32.75	35.38	*32.64	36.39
Male	**23.32	29.84	*25.27	33.05
Post				
Female	33.92	37.89	34.44	37.89
Male	30.9	31.97	31.43	31.97

Table 53: P and N2 scores by gender and Test/Control

The differences between male and female students are evened out when the comparisons are made within Test and Control groups. The only statistically significant gender differences can be detected in the Control group pre-scores (P-score p=0.005 and N2-score p=0.018, unpaired t-test with Welch's correction). What is particularly worth noting in Table 53, is that male students in the Control group catch up with the female students by the time the post test was applied. Further, it is worth noting that in ALL female groups the averages are slightly higher in the post questionnaire, while the average for Test group male students actually goes down in the N2 measure. Female students also continue to score higher in all measures in the post questionnaire.

Table 54 summarises the proportions of female and male students progressing and regressing in the N2 score. Only N2 is presented here for the sake of simplicity and no important information is likely to be lost considering that the direction of change was found to be very similar in the N2 and P scores above.

Groups	N2 change (%)	
	Progressing	Regressing
Test All (n=114)	54.4	45.6
Female (n=73)	54.8	45.2
Male (n=41)	53.7	46.3
Control	58.0	42.0
Female (n=59)	52.5	47.5
Male (n=22)	72.7	27.3

Table 54: N2 change direction by gender and Test/Control

In the Test group the direction of change is very similar between male and female students. In the Control group, on the other hand, we notice highly different patterns of change. The female students follow the pattern of students in the Test group, while a considerably higher proportion of the Control group males progress. This could be explained by the low pre score average of Control group male students (25.27 compared with the overall pre N2 average of 34.50). The low Control male pre score is equivalent to high-school scores (Rest 1994). This strong change and very low pre-score indicates a confusion within the student population about the moral decision-making tools they use: a genuine ten-point change, at least without intervention, is highly unlikely.

11.2.4 Type-score

One of the newest DIT measures is the Type -score (for more details refer to section 4.3.3). It is calculated from two components, a C-score measuring consistency in the choice of DIT items, and the dominant scheme a student uses.

Table 55 details the Type results from both pre and post questionnaires for Test and Control groups.

Group		DIT Types (% of students)				
	Type (% from Rest et al. 1999a)	Type 2 (4.4)	Type 3 (40.5)	Type 4 (13.8)	Type 5 (29)	Type 6 (12.2)
Test (n=123)						
	pre	12.4	45	1.6	39.5	1.6
	post	10.6	48.8	4.1	33.3	3.3
Control (n=106)						
	pre	13.6	38.2	1.8	44.5	1.8
	post	11.3	54.7	1.9	31.1	0.9

Table 55: DIT Types pre/post

First, the Test and Control group pre Types are not statistically different (p=0.494, Mann-Whitney). The second thing to note from the Type summary, is the low percentage of 'consolidated' subjects in either pre or post questionnaires (Types 1, 4, and 6). Consolidation refers to a stage where a subject is using dominantly one of the decision-making schemas, unlike transitional subjects, who use several schemas in a more interlaced manner. The C-scores (see Appendix V), which identify whether a subject is consolidated or transitional, ranged between 0.014 and 40.322, where the cut-off point for consolidated subjects is C>15.705. The standard deviation of all C-scores (Test/Control, Pre/Post) was 1.172 which is considerably lower, for example, than the SD of 9.09 in a 505 subject sample presented by Rest *et al.* (1999a). This again supports the suggestion that the students are experiencing a stage of transition and possibly even confusion in relation to their preferred moral decision-making method.

The high percentage of Type 2 students in the Glasgow University Bioscience population is one of the main differences between these two samples (Glasgow and Rest's 505 sample). This again indicates less mature stages of moral development in the Glasgow University sample in comparison to published international data. Also, the very low percentage of Type 4 students in the Glasgow University sample is worth noting.

In Type, as in P- and N2-scores, students both regressed and progressed between the pre and post questionnaire. Table 56 details the percentages of students in each group either progressing, regressing, or experiencing no change.

Group	Direction of Type-score change (% of students)		
	Progressing	Regressing	No change
Test	32.79	28.69	38.52
Control	26.21	33.98	39.81

Table 56: Type-score change %

Again, we find a large percentage of students regressing. But as the Type-score identifies more slowly occurring changes in the moral decision-making patterns than P-score and N2, there are more students who experience no change in Type-score between pre and post questionnaire in both Test and Control groups.

The changes between pre and post Types are statistically significant for both Test and Control group when measured using the Wilcoxon t-test ($p < 0.0001$). There is no significant correlation (Spearman's rank-order correlation co-efficient) between the pre and post scores for either group, thus a pre Type cannot be used to predict post Type. However, it is worth noting that the non-significant correlation was negative ($rs = -0.0430$) in the Control group and positive ($rs = 0.0979$) in the Test group - the test group students' pre-scores indicating a slight progression on average in the Test group in comparison with the slight regression on average in the Control group.

11.2.5 U-score

U-scores measure the level of utilisation of justice based items in making a decision in the DIT (for more details refer to section 4.3.3). U-scores for the sample were calculated by Dr Steven Thoma. A large number of students had not chosen an action for all three stories and the U-score could be calculated for only a sub-group of the DIT sample. Table 57 details the numbers of students in each DIT sub-group and the averages of their U-scores.

Groups	U-score	
	Pre	Post
Test	0.217 (n=79)	0.197(n=71)
Control	0.207 (n=59)	0.189 (n=56)

Table 57: U-score mean

The U-scores are relatively low. U-scores have a potential range of -1 (low utilisation) to +1 (high utilisation). However, large sample estimates of utilisation suggest an actual range of -.40 to .77 (Thoma and Rest 1999). These results indicate that the Glasgow University Bioscience students are using justice based reasoning only moderately as the basis of their decision-making. This coincides with the transitional Types dominant in the sample.

The differences between pre and post U-scores for the Control group were statistically highly significant ($p<0.001$, Wilcoxon t test) and the positive changes were larger than the negative changes. For the Test group the difference was also significant ($p=0.029$, Wilcoxon t test) and the difference of changes in both direction was relatively equal.

11.2.6 Gender variance in Type and U-score

The gender differences in Type-score in either Test or Control group, pre and post are not statistically significant (Control $p=0.152$ and Test $p=0.026$ both with Mann-Whitney).

There are no significant differences between male and female students in U-score either pre or post (Pre $p=0.799$ and Post $p=0.969$ both with unpaired t test with Welch's correction).

11.3 Perry Results

The Perry questionnaire was designed to measure students' meta-ethical developmental level. In this continuum a more advanced meta-ethical approach appreciates uncertainty in the application of moral values and recognises the importance of personal commitment and responsibility as part of being an independent moral agent (for more details see Appendix III - Perry's Scheme).

Table 58 details the inclusion criteria and the number of protocols at each step.

	Pre	Post
All	374	300
Complete Perry	358	282
Valid	326	278
Paired	225	225

Table 58: Inclusion steps for Perry

The validation of the Perry questionnaire was based on a comparison of questions 6 and 9, which represent the same items with different wording. If the difference between the choices of these two items was more than one place on the Osgood scale, the protocol was rejected on the assumption that the student had not read the questions properly.

Table 59 details the sizes of Test and Control groups, and their sub-groups in relation to teaching approach.

Groups	Perry Protocols		
	All	PBL	Discussion
Test	118	48	70
Control	107	36	71

Table 59: Perry protocols

In the pre questionnaire, similar trends to those detected in the trial Perry questionnaire (refer to section 5.3.1.2 and Appendix VII - Preliminary Perry Questionnaire for further details), could be detected in the dominance of Perry

type statements chosen for element IV in comparison to overall choices (see Figure 10: Perry Type usage in the pre questionnaire). Students who otherwise chose dominantly C or Cb items in the questionnaire (indicating the highest developmental levels in the Perry scheme) , chose B items (indicating lower levels of development in the Perry scheme) far more readily than other items for element IV - Commitment. This confirms the earlier hypothesis that commitment is the last step in the Perry developmental scale and that the students in Level 3 have made or are still in the process of making that transition.
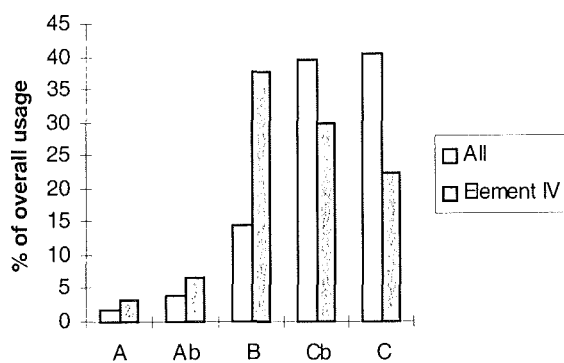
Figure 10: Perry Type usage in the pre questionnaire

A Personal Perry Score (PPS2) was calculated for all paired subjects, first for the entire questionnaire and then individually for the five distinct Perry elements :

     I Source and type of moral answers (Q2, Q4, and Q7)

     II Role of Authority (Q3)

     III Nature of multiplicity (Q8)

     IV Personal responsibility and relationship with multiplicity (Q1, Q5, and Q10)

     V Purpose of moral discussions (Q6, and Q9)

Table 60 shows the results for Test and Control group pre test. All the scores in Table 60 are means per statement, so that different elements can be more easily compared.

Groups	PPS2 scores					
	PPS2	PPS2 I	PPS2 II	PPS2 III	PPS2 IV	PPS2 V
Test (n=118)	17.49	18.81	15.79	18.97	14.41	20.28
Control (n=107)	16.87	18.17	15.14	18.03	14.04	19.50

Table 60: Pre PPS2

The overall pre PPS2 scores for Test and Control groups were not statistically different (p=0.155, unpaired t test). It is worth noting from this table that the average is below Cb threshold (16) only in elements II and IV. This supports data shown in Figure 10, that students choose on average less advanced statements in element IV. Element II was found previously in the preliminary Perry questionnaires to have a higher percentage of straight A responses, which brings the average down. In this sample, 49 students (21.7%) chose either an A or Ab item for element II. This is considerably more A items than for any other element (3 only for element III and 1 only for element V).

The direction of change in the Perry questionnaires (PPS2 for all elements) between pre and post ethics teaching intervention can be found in Table 61.

Groups	% of students (average change)		
	Progressing	Regressing	No change
Test (n=118)	**40.7 (+2.4)**	**54.2 (-2.5)**	**5.1 (0)**
Discussion (n=70)	41.4 (+2.5)	51.4 (-2.8)	7.1(0)
PBL (n=48)	39.6 (+2.2)	58.3 (-2.2)	2.1 (0)
Control (n=107)	**49.5 (+1.7)**	**48.6 (-2.9)**	**1.9 (0)**
Discussion (n=71)	52.2 (+1.4)	46.5 (-2.3)	1.4 (0)
PBL (n=36)	44.4 (+2.4)	52.8 (-3.4)	2.8 (0)

Table 61: PPS2 direction of change

Again, approximately half of the students regressed between the two questionnaires. Though pre and post PPS2 scores are significantly different (Wilcoxon t test, p<0.0001) for both Test and Control groups, the average changes are only small. Also, when the size of change was compared between Test and Control groups, they were not found to be statistically different

(p=0.0630, unpaired t test with Welch's correction). Still, Table 61 shows an interesting trend of Test group students being more likely to regress than Control group students.

In order to identify the area of change, changes in different Perry developmental elements were compared pre and post for both Test and Control groups (Table 62). In the table the higher PPS2 mean is highlighted.

Test (n=118)	PPS2	PPS2 I	PPS2 II	PPS2 III	PPS2 IV	PPS2 V
Pre	**17.49**	**18.81**	15.79	**18.97**	14.41	**20.28**
Post	17.09	17.58	**17.98**	18.92	**14.58**	19.43
Control (n=107)						
Pre	**16.87**	**18.17**	15.14	**18.03**	14.04	**19.50**
Post	16.30	17.07	**16.41**	17.63	13.85	18.59

Table 62: PPS2 mean scores: Test and Control groups, pre and post

The only difference between the Test and Control groups in the mean direction of change is in element IV - commitment and responsibility. The average magnitude of change is very small in PPS2 element IV - in the Control group - 0.184 and in the Test group +0.169. The range of change on the other had is quite large - from +12.33 to -10.33 in the Test group and from -10.0 to +8.33 in the Control group. However, the difference between the PPS2 element IV change between Test and Control groups is not statistically significant (p=0.461, unpaired t test with Welch's correction).

It is also noticeable that both groups progress in element II, the role of authority in moral questions. In the Post test the number of students choosing type A or Ab items in element II had reduced from 49 to 34 (15%), which accounts for much of the increase in average scores.

The post PPS2 scores for Test and Control groups were not statistically different (p=0.9142, unpaired t test with Welch's correction). Considering this, the high variation in the data, and the lack of clear difference in progression/regression, I conclude that the ethics intervention did not make a significant difference to Perry scores in general.

11.3.1 Gender variance in Perry

Gender differences have been found in both Moral Sensitivity test and DIT, so it seems only plausible to suspect that similar differences could be found in the Perry Test.

Table 63 details the numbers of male and females students in Test and Control groups.

	Test	Control	All
Female	81	74	155
Male	39	31	70

Table 63: Gender groups in Perry

Table 64: Pre PPS by gender details the mean pre Perry scores by gender.

Groups	Pre PPS2					
	PPS2 All	PPS2 I	PPS2 II	PPS2 III	PPS2 IV	PPS2 V
Female	**17.43**	**18.85**	**17.07**	**18.70**	**14.35**	**20.09**
Male	16.66	17.73	15.98	18.11	13.98	19.51

Table 64: Pre PPS by gender

In the pre scores gender differences were not found to be significant ($p= 0.096$ unpaired t test with Welsch's correction). It is worth noting that in the pre test, female students score higher in all five Perry elements. Element IV is again the lowest element for both genders, reflecting the overall result.

Table 65 summarises the average pre and post PPS2 scores for Test and Control group females.

Female	PPS2 All	PPS2 I	PPS2 II	PPS2 III	PPS2 IV	PPS2 V
Test						
Pre	**17.77**	**19.13**	**18.17**	19.33	**14.51**	**20.57**
Post	17.39	18.25	18.13	**19.72**	14.21	19.78
Control						
Pre	**17.08**	**18.57**	16.00	18.05	**14.18**	**19.59**
Post	16.62	17.19	**17.33**	**18.08**	13.91	19.22

Table 65: Female PPS2 scores

Although Test group females score higher than Controls in all five Perry elements, the differences between the groups are not significant in the pre-scores (p=0.159 unpaired t test with Welch's correction). The pre/post difference in PPS2 in the Test group is, however, highly significant (p<0.0001, Wilcoxon t-test). The direction of change is regressive - 38% of students progressed, 54% regressed and 8% had no change. In the Control group the difference was also highly significant (p<0.0001, Wilcoxon t-test) and the direction was more ambiguous - 52% progressed, 46% regressed and 2% had no change. The regressions were larger in size in the Control group than the progressions, and the general trend in the t test was thus regressive.

Table 66 summarises the Male students PPS2 scores in pre and post test. It is worth remembering that the group sizes for male students were less than half of those of the female students, and that this has an effect on the power of the statistical tests.

Male	PPS2 All	PPS2 I	PPS2 II	PPS2 III	PPS2 IV	PPS2 V
Test						
pre	**16.91**	**18.16**	16.47	**18.23**	14.21	**19.71**
post	16.48	16.21	**17.65**	17.31	**15.32**	18.73
Control						
pre	**16.35**	**17.18**	**15.41**	**17.97**	13.70	**19.26**
post	15.52	16.76	14.10	16.52	**13.72**	17.03

Table 66: Male PPS2 scores

Again, the Test group males scored higher in all five Perry elements in the pre-score. These differences were not significant however (p=0.602 unpaired t test

with Welch's correction). The differences between pre and post PPS2 are highly significant for both Test and Control group males (p<0.0001) and the trend in both groups is regressive. When comparing the sum of positive and negative ranks, it appears that the regressive trend was stronger in the Control group. In the Test group 45% of the male students progressed and 55% regressed, while in the Control group 41% of the male students progressed and 55% regressed and 3% stayed the same.

From the data above, I conclude that there was no significant gender variation in relation to the Perry scores and the teaching style did not have a gender-specific impact on students.

11.4 Comparisons across the Moral Development Questionnaire

161 students completed all three elements in the questionnaire and passed all validity tests for both pre and post test. This sub-group of students allows comparisons to be made between the different elements in the questionnaire.

Table 67 summarises the sizes of groups for males/females and Test/Control.

	All	Male	Female
Test	96	33	63
Control	64	18	46

Table 67: Groups in cross questionnaire comparison

Because different elements in the moral development questionnaire are not directly comparable (different scales), the results in different elements cannot be compared by employing t tests. One method is to compare the directions of change. Table 68 summarises the results of this type of analysis. The first two lines in the table identify the direction of change of each element in the Moral Development Questionnaire, and the last two lines list the percentage of students in Test and Control groups for each direction of change pattern .

Regress		All	PPS2 and Sensitivity	N2	N2 and Sensitivity	Sensitivity	N2 and PPS2	PPS2
Progress	*All*		*N2*	*PPS2 and Sensitivity*	*PPS2*	*N2 and PPS2*	*Sensitivity*	*N2 and Sensitivity*
Test (n=96)	21.9	24.0	27.1	16.7	5.2	3.1	2.1	N/A
Control (n=64)	26.6	20.3	29.7	18.8	1.5	1.6	N/A	1.6

Table 68: Direction of change in cross-reference

Close to half of the students in both Test and Control group either progress or regress in all three measures. When there are disparities between the different measures, N2 stands out as the one moving in a different direction to Moral sensitivity and PPS2. The last four directional combinations have six or less students in them, and are thus less significant.

The differences between Test and Control groups in changes of direction are small and from these data it is impossible to suggest that ethics teaching would have had a significant impact on the directional changes between parts of the moral development questionnaire.

An alternative analysis for comparing different elements of the Moral Development questionnaire is to carry out linear regression analysis between different elements either pre or post. This was done for the following pairs of Moral Development Questionnaire elements:

- Pre N2/PPS2 (p=0.010, r^2=0.041)**
- Pre N2/Moral Sensitivity (p=0.038, r^2=0.027) *
- Pre Moral Sensitivity/PPS2 (p=0.825, r^2=0.000) NS
- Post N2/PPS2 (p=0.033, r^2=0.028)*
- Post N2/Moral Sensitivity (p=0.078, r^2=0.019) NS
- Post Moral Sensitivity/PPS2 (p=0.395, r^2=0.005) NS

Statistically significant linear relationship could be found in three pairs: Pre N2 and PPS2, Pre N2 and Moral Sensitivity, and Post N2 and PPS2. For all of these three pairs the goodness of fit of the linear regression line is very low (r^2<0.041). The explanatory power of these regressions is thus very weak.

These data suggest that when there is a moderate link between the direction of developmental change in moral cognitive skills (N2), moral sensitivity, and meta-ethical understanding. When the different elements are not connected, moral cognitive skills proceed as an independent element. When the scores are compared pre and post, only very moderate relationships between different elements could be demonstrated. Thus students tend to develop in these skills more often than not on all fronts, but there is no link between the levels of skills in each element.

11.5 Summary of Moral Development Questionnaire results

The results from the Moral Sensitivity Test indicate that ethics education in the form of discussion groups and PBL encouraged development of moral sensitivity. In the Test group more than half of the students progressed in the Moral Sensitivity score during the research period and these changes were statistically significant. This is highlighted in the mean score change from 4.68 to 5.30. In the Control group only 31.5 percent of the students progressed during the research period and the mean Moral Sensitivity score regressed from 4.89 to 4.67. The Control group pre and post Moral Sensitivity scores were significantly different.

There were no significant differences between the mean Moral Sensitivity pre scores of male and female students. However, the male students in the Test group improved more as a result of the intervention than the female students. At the same time, the males in the Control group regressed more during the research period than the female students.

The DIT scores returned highly unexpected results, with close to 50% of the students regressing between pre and post questionnaire in both Test and Control group. Regression is recognised in DIT, but for this level of regression there is no published precedent. Possible explanations for this pattern are explored in the Discussion to follow. Also, it is worth noting that the pre P- and N2 scores for the Glasgow University sample are significantly lower than international averages for university students.

Differences between groups were in many respects insignificant: there was no statistical difference between the following data sets:

- Pre-questionnaire P-score between Test and Control groups
- P-score change in size between Test and Control groups
- Post P-scores between Test and Control groups
- N2 pre and post difference between Test and Control groups

- N2 post scores between Test and Control groups
- Female and male students (when divided into Test and Control groups) for pre and post P- and N2-scores (apart from Control pre scores)
- Gender difference for Type-score
- Gender difference for U-score

On the other, the following statistical analyses revealed significant differences:

- Paired pre and post P-scores for both Test and Control groups
- Paired pre and post N2-scores for both Test and Control groups
- Pre N2 and N2 change negative regression for both Test and Control group
- Pre N2 and P between male and female students including both Test and Control group students (females scoring higher)
- Only Control group female/male students in pre N2 and P scores
- Pre and post Type-scores for both Test and Control group
- Pre and post U-scores for both Test and Control group

From this we can draw three main conclusions:

1. Ethics teaching did not have a statistically significant impact detectable by P- or N2 score on the moral development of Test group students.
2. The high fluctuation of scores, dominance of transitional decision-making schema and the relatively low U score indicate confusion among students in choosing ethical decision-making methods in a consistent manner.
3. There are no significant gender differences in DIT scores.

The Perry scores support the original hypothesis that commitment and responsibility are the final challenge in the Perry development scheme. In elements IV and II (authority) the average pre scores were less than the Cb threshold.

In the Perry questionnaire, the students moved in both directions - progressed and regressed - in very similar proportions. Regression was, however, slightly more common within the Test group. Due to this high level of movement there were highly significant changes between pre and post Perry scores within both Test and Control groups, but no significant differences between them. There was also no significant gender variation in the Perry score.

These results support the data from the other questionnaire parts in describing the students as in transition and search for moral decision-making tools. In Perry, that is particularly highlighted by the low scores for commitment and moral responsibility.

When comparing different elements of the Moral Development Questionnaire two main results emerged:

- Approximately half of the students either regress or progress in all three measures and for most of those students who have less consistent developmental patterns, the direction of N2 changes is different from Moral Sensitivity and PPS2.
- There are only very weak relationships between the different elements of the Moral Development Questionnaire indicating that the parts are measuring genuinely different elements of moral development.

12. Moral Development Questionnaire - Level 1 results

The original research plan included ethics teaching in L1. It turned out to be technically impossible to introduce any significant amount of ethics teaching into the L1 curriculum and the full inclusion of L1 in the research had to be abandoned. Moral Development Questionnaire data was, however, collected during the first week of term 1 in order to study natural developmental patterns in university students, by comparing the L1 data with the L3 pre-teaching data. There is no particular reason to expect session 1999-2000 L1 students to be different in terms of moral sensitivity or moral cognitive skills from the equivalent group two years before, and the relatively large sample size I was able to compare, gives confidence in the comparison. Ideally, it would have been desirable to follow moral development in a group of students through from L1 to L3. However, time constraints did not allow this.

The L1 sample consists of 253 questionnaires of which 172 (68%) are female and 81 (32%) male. The sample is approximately a quarter of the entire L1 population. The students were chosen at random: five laboratory groups filled in the questionnaire.

12.1 Moral Sensitivity results

The Moral Sensitivity scores were calculated for a randomly chosen representative sample of the L1 Moral Development Questionnaires. The sample consisted of 50 questionnaires, 36 (72%) female students and 14 (28%) male students.

The mean number of responses (including zero scores) was 3.90. After the zero scoring results were removed from the analysis, the mean number of responses was 3.18. The mean number of non-zero scoring entries was not significantly different from the L3 pre mean of 2.87 (p=0.77, unpaired t test).

When the response frequencies are compared per types in the scoring guide
(Table 69), the differences between L1 and L3 are statistically significant
(p<0.0001, unpaired t test). The differences appear in types I and II, with L1
scores significantly higher.

Group	Moral sensitivity types			
	I	II	III	IV
L1 (n=50)	1.72	1.75	0.29	0.14
L3 (n=267)	1.38	1.35	0.30	0.19

Table 69: L1 and L3 Moral Sensitivity response frequencies by type

The mean Moral Sensitivity score for the L1 sub-sample is 4.275. Even though
the L3 pre score mean is 4.780, the means of the two scores are not statistically
different (p=0.097, unpaired t test). The higher mean of non-zero scoring
responses in L1 and the lower Moral Sensitivity scores indicate that the L1
students had more lower level scoring responses when compared with L3.

12.2 DIT

Out of the 253 completed L1 DITs, 37 (15%) did not pass one of the validity
tests. Out of the rejected questionnaires, 8 failed the M-test and the rest (29) the
internal consistency test. This is a considerably lower rejection rate than for L3
(23%), and within the average rejection rate quoted by Rest *et al.* (1997). Later, a
further 23 DIT questionnaires were abandoned due to incompleteness of rating
data.

The mean P-score for L1 is 28.94 and it is not statistically different from the L3
pre score of 30.94 (p=0.513, unpaired t test). The mean N2 score for L1 is 28.49.
Again, there is no significant difference between the L3 pre N2 score (31.09) and
L1 N2 score (p=0.504, unpaired t test).

Figure 11 shows the differences in Type-score between L1 and L3 students.
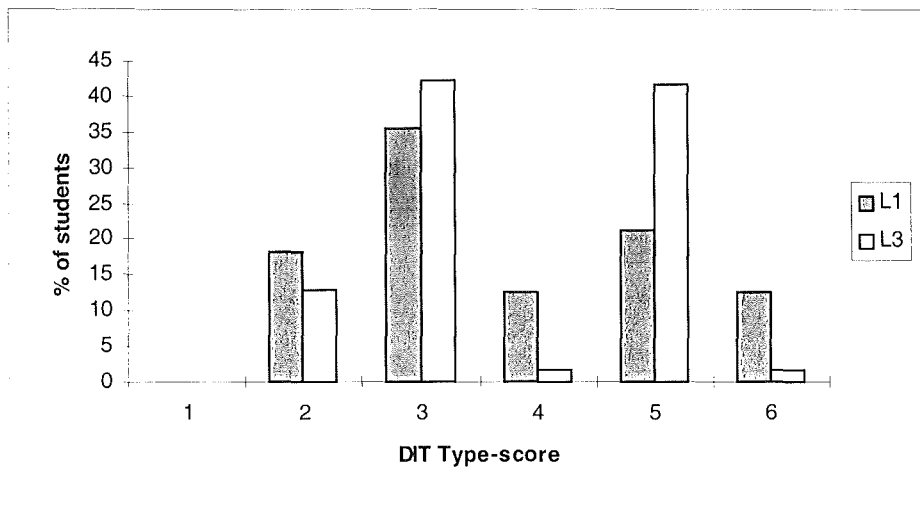
Figure 11: Type score for L1 and L3

Figure 11 shows clear differences between the L1 and L3 Type-scores. The scores are significantly different (p=0.010, Mann-Whitney test). What is noticeable in this graph is the higher percentage of L1 students in both extreme Types 2 and 6. L1 students are also across the board more consolidated (types 4 and 6) than L3 students who are dominantly transitional (types 2, 3, and 5).

12.3 Perry

Out of the 253 L1 Perry questionnaires, 19 (7.5%) did not pass the internal validity test.

Figure 12 shows the Perry scheme patterns both for all Perry elements and then separately for Perry IV elements in L1 and L3.
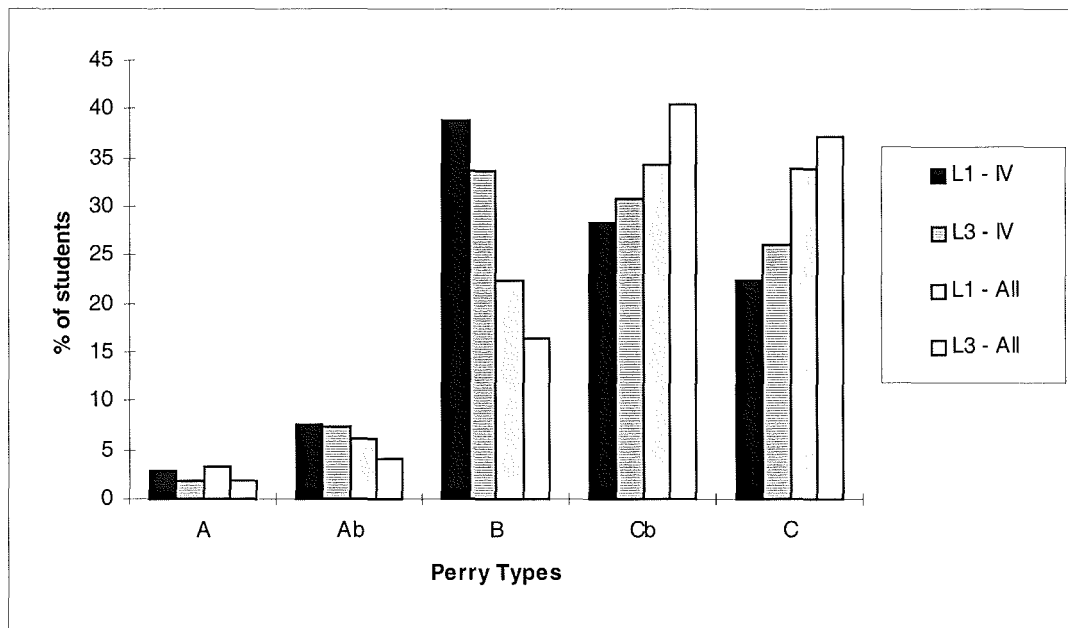
Figure 12: L1 and L3 Perry usage

The L1 students follow the pattern of choosing lower level items for Perry IV elements. It is noticeable, however, that the L1 students have chosen Type B items more often than L3 students and correspondingly less C and Cb items in all Perry elements.

Table 70 details the PPS2 data for L1 and L3 (Test and Control group pre data).

	PPS2 - all	PPS2 I	PPS2 II	PPS2 III	PPS2 IV	PPS2 V
L1 (n=234)	15.26 ± 3.05	15.29 ± 4.52	14.25 ± 8.84	18.42 ± 6.93	13.63 ± 3.32	16.58 ± 5.56
L3 (n=225)	16.72 ± 3.23	17.34 ± 4.63	16.14 ± 7.89	18.31 ± 6.13	14.23 ± 3.73	19.03 ± 5.14

Table 70: L1 and L3 Perry scores

The L1 and L3 Perry scores per type are statistically different (p=0.0095, unpaired t test). The L1 students scored lower in all elements except III, where the mean score is only a fraction higher than that of L3 students.

12.4 Gender variance in L1 scores

The moral sensitivity scores were calculated for 37 female students and 14 male students. The scores were not significantly different (p=0.6001, unpaired t test)

In L1, 137 female students and 56 male students completed a valid DIT protocol. There were no significant differences between the male and female students' P scores (p=0.971, unpaired t test), or N2-scores (p=0.985, unpaired t test).

The Type-scores were not significantly different (p=0.355, Mann Whitney test).

The PPS2 was calculated for 160 female students and 74 male students. There was a significant difference (p=0.008, unpaired t test) between the male and female student PPS2 scores when they are compared in all five elements.

Table 71 shows the PPS2 for L1 female and male students separately.

	PPS2	PPS2 I	PPS2 II	PPS2 III	PPS2 IV	PPS2 V
Female	16.64 ± 3.36	17.42 ± 4.94	15.31 ± 8.03	18.88 ± 6.76	14.30 ± 3.70	18.55 ± 5.69
Male	15.35 ± 3.16	15.52 ± 4.86	11.59 ± 8.59	19.50 ± 6.34	13.23 ± 3.75	18.11 ± 5.79

Table 71: L1 PPS2 by gender

The female students scored consistently higher (apart from element III - Nature of multiplicity) and the difference was particularly marked in element II - Role of authority.

12.5 Summary and conclusions of L1 results

In the Moral Sensitivity test, L1 students raised on average more questions and issues than L3 students, but scored lower. The number of responses raised by L1 students was significantly higher than the L3 figure, while the difference between the Moral Sensitivity scores was not. This suggests two conclusions:

1. L1 students did take the questionnaire seriously and thus we can assume that the scores are a true reflection of their moral sensitivity.

2. Even though the Moral Sensitivity scores were not significantly different, the fact that the L3 students have a higher average score per raised issue/question indicates that there is a natural progression towards more sophisticated moral sensitivity during the first years of tertiary education.

In DIT the differences between L1 and L3 students were not statistically different when P- and N2 -scores were calculated. The Type-scores, on the other hand, were significantly different, with L1 students more consolidated than L3 students. This suggests that the university experience gained by the L3 students does not, *per se,* encourage moral development as far as it is measured by the DIT, while it does seem to increase the mix of schemas the students use. The increase in transitional types could be interpreted as increased confusion about moral decision-making tools, which is apparent in the L3 pre and post questionnaire data.

In the Perry questionnaire, L1 students follow the same pattern as L3 of choosing lower level items more often in element IV than in any other element. The PPS2 scores of L1 and L3 students are statistically different with L3 students appearing more advanced than the L1 students. This suggests two conclusions:

1. When introducing ethics, we need to take into account students' ability to deal with complex moral issues. If students are dominantly Perry types A and B, teaching needs to be designed to support the learning process of students at those stages.

2. Perry development seems to occur spontaneously, possibly both as a result of the academic experience and of general life-experiences.

The gender differences found in the Perry questionnaire, with the male students choosing significantly more often lower level Perry items than the female students, suggest that the male students would have more difficulty following

ethics teaching that is not designed to promote growth in meta-ethical thinking, but rather further development of students who have reached Perry type C.

Part V - Discussion and Recommendations

The ethics intervention at Glasgow University during the academic years of 1999-2000 was a challenging project. Not only was there no known precedent of science ethics education in UK, there were also academic and practical challenges in setting up a manageable ethics course and a suitable assessment method. In this discussion I hope to draw together the different elements in what I would regard as a successful ethics course based on the experience at Glasgow - from creating a suitable teaching method to using assessment methods as supportive tools in the process. In the end, I will summarise the essential criteria for an effective ethics course for biological science students.

13. Teaching Ethics

13.1 Creating an Interest

Maximising the opportunities for student participation was a successful pedagogical choice in teaching ethics. The Glasgow University L3 Bioscience students participating in the ethics intervention clearly expressed their interest in ethics and enjoyment in learning by doing.

I have once heard a bold claim that 70% of learning is based on motivation and 30% on intellectual/cognitive abilities (lecture by Professor Arto Mustajoki, University of Helsinki, 1996). If this is anything close to the truth, we cannot ignore the student perception of teaching. This is particularly the case when the teaching involves a subject area not directly chosen by the students, like ethics in a science curriculum. This is in contrast to more obviously subject-based material students may see the need to learn such as, say, molecular techniques. There is a greater need for personal motivation for a subject students regard as less relevant to their future careers. If the students are not motivated, ethics courses will fail to generate learning. There is increasing support and demand for ethics education for science students, but there is no necessary link between that and student motivation. It is a task for the educators to encourage and nurture motivation towards what is considered important, both in life and within the degree-course.

I also believe that it would be difficult, if not impossible, to create a motivational ethics course which aimed at inculcating the 'right' values. This approach, as was discussed in Part I, is not only morally questionable, but also demeaning to the students. It is certainly possible that some students would have an interest in receiving the 'right' answers from an ethics course. Science students often find the lack of 'right' answers very uncomfortable - saying things like 'I've found my degree course till now very effective: teachers have told me what to learn and I've done it. Now, you are asking me to read all this complicated and contradictory stuff which simply confuses me ... why don't you just tell me the

answer!' Unfortunately such a course would not spark motivation in the students to develop their skills further and to explore unexpected new problems with confidence and skill.

Students in this project arrived at the ethics sessions with their own pre-conceptions of what it was for and how highly they valued these expected benefits. My perception is that most of them were at the best curious about it, but rarely enthusiastic. It was then my task as a facilitator to describe the reasons for the ethics course, the benefits I believed were to be gained from it, and to create an environment where the students could build their own motivation. For me, the key in motivating students towards ethics was a respect for students' own views, and provision of an intellectual challenge.

Reluctance to discuss, poor motivation, shyness, and lack of skill needed to be overcome before the sessions could achieve their core objective - to ignite a spark to think and consider ethical issues. I certainly did not succeed in overcoming these obstacles with every group, but with the majority I believe I did. It was enjoyable to see the interest grow, the views develop, and, quite literally, learning to take place during the discussions. This subjective impression of the discussions is supported by 74% of the students highlighting challenges to personal ethical views as one of the main benefits of ethics sessions.

13.2 Developing skills

Once the motivation is in place, I believe that the next most important thing in teaching ethics is to actively engage students in using and improving their skills. Moral decision-making is not about having the 'right values', but a skill to recognise moral issues, to include all shareholders in the consideration, to use adequate methods for comparing different routes of action, and having the personal strength to follow the decision through. We cannot teach skills without creating an environment where they can be practised.

Creating an opportunity for students to practice their moral decision-making skills is not about stepping back and letting them get on with it: very emphatically not. In order to create an environment for growth, you need to nurture the skills, provide students with the basic tools they need, and give them explicit support and encouragement. A logical and chronological order might call for a series of introductory sessions for 'giving the tools' - i.e. explaining the methods of moral decision-making. I believe that this would be, in most instances, detrimental to motivation. Starting an ethics course with students in a passive role, creates an unnecessary uphill struggle when later trying to re-activate the students to practice and use their skills. It is also a subtle message of distrust - we believe you cannot make good/right moral decisions, so we are here to tell you how. It escapes me how that is supposed to motivate anyone. People have a tendency to want to live up to the expectations they perceive, and it can well be that when you expect a little, you only get a little.

In this aspect ethics education seems to begin at the stage that traditional science teaching tends to reach at advanced stages i.e. an elementary approach in science is to teach the 'facts' first and only later to emphasise areas of doubt where hypotheses compete and cannot be easily resolved. The contrast does not need to be this sharp. If students lack the basic skills of ethical decision-making, an ethics course requiring advanced skills would not be the most successful one. Ethics, like science, is not a discipline where advanced levels can be successfully understood without basic knowledge. Further, ethics and science are similar in their emphasis at the more advanced levels on the value of finding out for yourself rather than relying on an 'authority' for the truth. But what is important, however, is that the basic tools of ethics do not need to be taught using didactic teaching methods. Student-centred teaching can be designed to encourage moral development even when students still have a long way to go before reaching maturity as moral agents. David Shapiro in his wonderful book 'Choosing the Right Thing to Do' (1999) describes several such methods.

Inclusion of support and supply of tools in student-led learning is not a straight-forward process. It involves striking a fine balance between offering too much

and too little support, and timing it right. Each group is different; every meeting with the students is unique; so it is impossible to have a set plan to deliver the pieces of golden wisdom you might perceive that the students need. Instead, what a facilitator can do is to have a collection of methods and tools ready and a perceptive eye to see when and how to deliver them. Facilitation is a skill that requires training and practice. Even after almost 200 hours of facilitating ethics discussions, my learning curve as a facilitator did not reach a plateau during this project - there remain many areas where I can see improvements both as possible and necessary.

13.3 Teaching in Glasgow University

The two different teaching methods chosen for ethics teaching both worked well, but they were distinctly different. From the comfort of hindsight, I would now say that the discussion approach had distinct advantages over PBL. One of the main disadvantages of the PBL programme in this project was the concurrent introduction of both PBL as a method and ethics as a subject. Even though the teaching practice at Glasgow University is changing towards student-centred approaches, didactic teaching is still the dominant form. Further, students have over 10 years of school experience, which has more often than not encouraged passive learning techniques. Against this back-drop, PBL is a radical approach to learning. Students who are used to being told what to learn, how to learn it, and when to learn it may find the responsibility for learning daunting and unnerving. There is safety and comfort in being told. The transition from a passive absorber to an active searcher of knowledge takes energy and attention, which at first may detract from effective learning of the subject matter.

In this project, to make things even more demanding, the subject matter was also difficult in its essentially novel approach to problems. The methods of doing good ethics and the methods of doing good science are not identical: in some places they are even distinctly foreign to each other. In science the quest is for the best approximation of reality based on hard facts in numbers from experiments. Ethics, on the other hand, is not based on hard facts, but an

appreciation of reality based on a coherent application of value systems. Where science describes reality, ethics evaluates it. Glasgow University students have some exposure to ethical decision-making before L3 (see chapter 8 Teaching strategy), but the ethics course designed for this project requires students to adopt a more focused approach to ethics than before.

The ethics PBLs did not achieve their full potential because the students needed to direct their attention and energy towards mastering the method at a cost to concentration and focus on the subject. The students were also genuinely at a loss to understand and recognise the ethical dimensions of the problems they were given. It is my belief that without active facilitation most groups would have proceeded to solve their problems with only minimal consideration of ethics. The main learning gain would have thus been the process of PBL, not ethics. Ideally then, PBL should be used in ethics teaching only when either 1) the students are already familiar with the PBL method, or 2) they have had an introduction to ethics as a form of inquiry, possibly structured ethics discussions. In my experience, the chosen problems functioned as well as they could in integrating ethics and science in a meaningful way and allowing students to approach the problem with the level of detail and dedication they chose. Thus, the PBLs were not unsuccessful, while they did not meet their full potential as a method for teaching ethics.

The structured discussion programme provided a vehicle for students to explore their own moral values and moral decision-making models and to learn from those of others. In most instances the discussion seemed to capture the attention of the majority of students and the learning gains were both in increased understanding of the issues, and in improved personal decision-making skills.

I believe that the preliminary reading was essential to the success of the discussion programme. It gave the sessions a focus, a reference point, and highlighted alternative ways of looking at a problem. The choice of the reading was difficult, however. The piece needed to be short and challenging, while at the same time an understandable overview of the key ethical issue. In my view, it

is important to read difficult texts as well as easy ones, both truly philosophical and scientific ones from which to distil the ethical issues. The reading should provide a challenge appropriate to the skills of the students. I chose to write the preliminary reading for a few topics when it proved impossible to find a suitable paper, but I believe it would not have been ideal to prepare the material specifically for all courses. Important learning opportunities could have been lost, if the reading material had been too homogeneous.

I also believe it was essential to have a facilitator with a strong ethics background. The students needed a lot of facilitation in recognising and analysing ethical issues. I believe this would have been difficult provide if the facilitator had only a limited understanding of ethical theories and had not had a formal opportunity to practise ethical decision-making skills. Further, it was extremely important that the students were actively encouraged to explore ethical issues and that their efforts were supported in a friendly and appreciative environment.

If I were to design an ethics course afresh with what I know now, and importantly, if I had more time with the students, I would not design a straight structured discussion course, nor a PBL one. A mixture of these two could prove beneficial - structured discussions preceding PBL. An alternative would be a discussion programme including more varied forms of interaction - role plays, debates, games, and presentations. These could add a new dimension to the understanding of ethical issues and make the learning more fun - possibly a key element in motivation.

13.4 Assessment

I believe it is crucial for the success of an ethics course that it should be included in the assessment procedures. Assessment is an agreed sign of importance and students are apt and accustomed to make their time and effort-related choices based on these signs. If ethics is a valued and important part of the curriculum, we need to attach the appropriate credentials to it. Assessment also provides

external motivation, possibly not sufficient on its own to make the most of ethics teaching, but it might be a good starting point.

Before implementing an assessment procedure, we need to ask ourselves what aspects of a bioethics courses are appropriate for assessment? We can assess, among other things, the knowledge of ethical theories, the ability to apply different moral decision-making methods, and participation and/or attendance. I do not believe that measuring knowledge of ethical theories is appropriate for bioscience ethics courses, for the simple reason that the understanding of theories should not be an important objective of a bioethics course. An ability to apply different moral decision-making methods coincides better with the objectives, in particular if recognition of the issues is also included in the task. Attendance would be the bare minimum, guaranteeing that the students would at least allow themselves to be exposed to new thoughts.

Including ethical considerations in what is considered a successful exam-answer is one of the best available methods to assess ethics. It would highlight the interaction between the scientific curriculum and ethics; it would require students to use their moral decision-making skills in context, and it would give ethical concerns an unquestionable seal of inclusion in the important elements of science education. At the same time, it should not present a barrier in an ethics discussion, which would decrease the freedom students perceive they have in expressing their views.

Assessing students before an ethics course may also be essential to a successful ethics course. Students' ability to recognise moral issues, to deal with the information, and how they perceive moral reality influences the way they perceive ethics teaching. It may well be that certain teaching approaches are more likely to encourage moral development of students in a particular stage of their moral development and even halt progress of students at another stage. For instance, it may be that the traditional ethics education process of challenging existing moral understandings has different effects on students based on whether they use a high mix of schemas (transitional) or a low mix (consolidated), or on

their U-score (utilising justice-based considerations in moral decision-making). A transitional individual with a low U-score may not benefit from a traditional ethics education approach, which challenges different moral views, since the intervention may serve mainly to highlight confusion and the limited utility of moral concepts. On the other hand, more consolidated subjects with a high U-score may respond in the intended manner to the intervention and find it conducive to re-considering their chosen moral views. It would be fascinating to study further the possibility of tailor-making ethics education to suit students in different stages of their moral development.

It is not particularly common to test students at university before designing teaching. Some methods have been employed, particularly in IT, to make teaching appropriate to the existing skill levels, but in more academic subjects it is not a widespread approach. The benefits of testing students before an ethics course are significant however. The testing methods used here tell us more than just students' moral decision-making levels - they describe some of the fundamental approaches students use in moral decision-making. These approaches may influence directly, as suggested by Thoma and Rest (1999), the benefits students can receive from an ethics course using a particular method.

This is similar to testing students on their approaches to learning in general - do they use deep/surface learning approaches; how versatile are they in mixing different learning approaches (Entwistle, 1988); do they excel in private study or in group work etc? Designing a course based on the knowledge you have gained about how your students learn, can enhance the impact of your course greatly. I believe this to be particularly true for ethics teaching.

14. Results

Measuring moral development is difficult. Even after an agreement on what moral development consists of, the development of a measuring approach and tools is complex and demanding. My choice in this research was to include one existing and validated measure (DIT) and to develop two new methods to study other areas of moral development. This was important for two reasons: 1) because the scope of DIT is very limited, and 2) because a reliable measure would provide an important comparison point to the new methods.

Even though only half of the students participating in the study filled in both questionnaires, the numbers remained comfortably large enough for most planned analysis. The results from the Moral Development Questionnaire proved complex and unexpected. The main result expectation - a clear difference between the Test and Control groups - did not materialise in most instances, but the results revealed interesting trends in student development.

The Moral Sensitivity Test was designed to measure students' ability to recognise moral problems. The fact that the mean number of questions/issues raised by students was 3.2 indicates that the students responded with an intention to give a genuine account of their thoughts. This is supported by the relatively small number of students leaving the Moral Sensitivity Test completely blank. This measure is, however, more subject to time-pressures and motivation than other parts of the Moral Development Questionnaire. Still, I believe that there is enough evidence to confirm that the results reflect students' genuine ability to recognise moral issues.

The overall mean Moral Sensitivity scores for all groups both pre and post teaching are relatively low, only a third of the theoretical maximum. Students, when they recognised an ethical issue, recognised it only in its most basic terms, accruing a low score. The scoring guide was designed to minimise the importance of length in answers, and thus reduce the impact of style in the score. The variation of scores within the group was also relatively high. In my

understanding these results highlight two main issues: 1) that moral sensitivity is an area where students need to progress further in order to call their moral decision-making skills excellent, and 2) moral sensitivity is a very individual measure, where large differences between students exist, while gender is not a key identifier.

There was a statistically significant difference between the Test and Control groups in Moral Sensitivity. The Test group had a progressive trend, while the Control group had a regressive one. The difference was not clear cut between the groups, as a third of the Test group students regressed during the study period and a third of the Control group progressed. Still, the benefits of ethics teaching were clear and distinct.

Moral sensitivity is the first element in moral decision-making. Before recognition of problems, there cannot be a decision-making process, which makes moral sensitivity not only the first, but an essential part of the process. Progress in moral sensitivity is increased awareness. It is adopting new ways of looking at a problem and including more shareholders into the considerations. The teaching methods in this project concentrated on the recognition and discussion of the nature of moral problems. The results showed that the approach was successful in meeting its objectives in increasing student awareness of ethical issues.

The DIT measures the use of justice-based moral decision-making tools. The DIT results were unexpected in two ways; 1) the students scored lower than expected for their age group, and 2) regression was much more common than previously reported for DIT (Self et al. 1998; Schlaefli 1984; Thoma 1984; Rest 1986) .

There were no distinct differences between Test and Control groups in any of the DIT scores calculated. When we add to these results the significant changes between pre and post scores for all student groups, we get a picture of a student

body that is very much unsettled and immature in their use and choice of moral decision-making tools.

The high rate of progressing and regressing in both groups, and the magnitude of change they experience, indicates that the students did not have a consistent approach to making moral decisions. They also utilised justice-based considerations only very moderately. The best explanation for this high level of fluctuation in both directions on the DIT scale, is that the students were profoundly confused about moral decision-making methods and as a result employed methods at their disposal in an inconsistent manner. The low P and N2-scores add to this confusion. The students were not only confused, but they predominantly used less sophisticated decision-making methods than expected for their age.

It is worth considering why these young adults appear to experience this intense confusion. As was found by Spickelmier (Rest and Deemer 1986), moral development in higher education is not dependent on specific skills, but rather the personal orientation to learning and development in general. When we consider the increasing pressures on students today to juggle part-time jobs in order to finance their studies and that teaching is becoming increasingly goal orientated with specific learning objectives for each course, it seems possible that universities are no longer offering the best possible environment for moral development to take place. There seems to be a decreasing amount of time and encouragement for following individual interests in studying which may create an atmosphere more suitable for training than education. This corresponds poorly with the ideal of a higher education which supports deep learning, students' competence in critique and self-reflection, and provides an opportunity for students to engage in an enquiry process in open dialogue and co-operation, freed from unnecessary direction (Barnett, 1990). Ethics education seems to have a role to play in reaching this ideal, but it cannot alone counter-balance the strong training-like characteristics students face in so many of their courses.

I have not been able to locate any DIT studies carried out exclusively on bioscience students, but I do not believe there to be any inherent difference between bioscience students and higher education students in general that would explain the low DIT scores in the Glasgow University bioscience student sample. Neither do I believe that the Glasgow University sample has some inherent disadvantages that would be reflected in the low DIT-scores. Based on these assumptions, I would suggest that the low DIT-scores in the Glasgow University sample represent a more general trend in young adults in the UK today. This is something that might be worth further investigation.

The ethics education intervention in this study was from the out-set a minimal one. It did not compare well with the most successful interventions lasting 4-12 weeks with weekly sessions (Schlaefli 1984). To reach this level, the contact hours would have needed to be at least doubled if not even tripled. For this reason also, it is not surprising that the DIT results did not detect significant benefits of ethics teaching in the Test and Control group scores.

The Perry questionnaire measures students' meta-ethical development. The results resemble the Moral Sensitivity and DIT results in having a high percentage of students both regressing and progressing between pre and post questionnaires. Out of the five Perry elements, in only one (personal responsibility and relationship with multiplicity) were the L3 students still experiencing major developmental shifts. In all other areas the students had predominantly reached the highest levels of development.

This result is comforting. The students have the basic understanding of the nature of moral questions, which is important for the type of ethics teaching employed in this research. If the students did not accept the multiplicity of moral answers and the relativity of right and wrong in applying moral principles, the teaching would have to address these issues before it could concentrate on moral sensitivity and cognitive skills.

The ethics intervention did not make a significant difference between Test and Control group Perry scores. This is understandable for two reasons: 1) Perry development is a fundamental process involving a change in a larger perspective of life and a short ethics course is unlikely to make a great impact on that, though it can be one of the elements encouraging change, and 2) most students had reached the highest levels of Perry development already, before the ethics intervention, and thus there was only limited room for development.

The benefit of the Perry questionnaire was the gained knowledge of students' Perry developmental stage. It is reassuring to have evidence that the students did not perceive moral reality as black-and-white and thus were within the reach of ethics education as it was designed in this intervention.

Analysis of the Moral Development Questionnaire results for Level 1 students suggests that there is only a limited amount of spontaneous moral development between L1 and L3. The Moral Sensitivity scores were not significantly different between L1 and L3 students, and neither were the DIT P- and N2-scores. The DIT Type-scores, on the other hand, indicated a shift from more consolidated schemas to more transitional ones as students progressed from L1 to L3. Also, there was a significant change in the L1 and L3 Perry scores, which suggest that students' views on moral reality have matured during their first years at university. This indicates that where students did not gain any major advantage in moral sensitivity or moral cognitive skills, their pattern of approach did change. The two extra years of academic life seem to reduce students' confidence in one moral decision-making method being sufficient. Thus movement has taken place and needs to be captured by moral education to use it productively towards moral development.

Different elements in the Moral Development Questionnaire, when studied from the L3 data, were not connected. Scoring high in one element did not indicate a high score in another element. This shows that moral sensitivity, moral cognitive skills, and meta-ethical conceptions are distinct elements of moral development. However, the elements are linked dynamically. Progressing/regressing in one

element is more often than not linked with a progression/regression in the other two elements as well. In other words the developmental patterns are connected. When the connection is broken, moral cognitive skills move in the opposite direction to moral sensitivity and meta-ethical conceptions. This highlights moral cognitive skills as separate in nature from the basic understanding and recognition of moral issues.

The teaching should thus include exercises in recognising the issues - explicitly working through a problem in order to highlight all the possible ethical questions that arise. Also, students need to engage in serious decision-making processes - trying different methods within one problem to see the differences between outcomes and thus assisting them in choosing the most appropriate methods. These two tasks can naturally be linked chronologically and around the same problem. Structured discussion programmes and PBLs allow the inclusion of these elements naturally and provide a suitable setting for both types of learning.

15. Conclusion and recommendations for the future

Ethics teaching is important and necessary. The external pressures to guarantee that graduates have at least elementary understanding of ethical decision-making methods and ability to recognise ethical issues are increasing and students see this need themselves. At the same time, the results from this study clearly and unmistakably demonstrate that the students do not have these skills as yet. They use relatively unsophisticated moral decision-making methods in an inconsistent manner. Ethics teaching is needed to meet the needs of both students and the society they are to live in.

The minimal intervention designed for Glasgow University Bioscience students has proven to be a suitable approach, but an inadequate one. It has given distinct advantages to students in supporting the development of their moral sensitivity, but it was not extensive enough to make a difference in their choice of decision-making methods. The students' self-reported benefits tell of motivation and interest in the way ethics has been taught, which is half-way to a successful outcome.

For ethics teaching to achieve more substantial benefits in both moral sensitivity and moral cognitive skills, the following elements would be essential in course design:

1. The minimum extent of the course should be five meetings + independent study.

2. Teaching methods should be student-centred and the facilitator have a strong knowledge-base in ethics.

3. Ethics should be included in course assessment procedures.

Further, the impact of ethics teaching is likely to be highlighted if the following elements support the course structure:

1. Ethical considerations are raised through-out the course structure by staff members who have participated in an introductory ethics course themselves, as well as in dedicated ethics sessions.

2. Ethics is included as part of the course through-out the four year university degree. The quantity of the teaching should increase towards the end of the degree with proportionally more ethics each year as students mature in their ethics decision-making abilities.

Ethics education has an opportunity to make a difference in students' lives, so let's do it!

Bibliography

ALBANESE M.A., MITCHELL S. (1993). Problem-based Learning: A Review of Literature on Its Outcomes and Implementation Issues. Academic Medicine 68:52-81.

BAKKEN L., ELLSWORTH R. (1990). Moral Development in Adulthood: Its Relationship to Age, Sex, and Education. Educational Research Quarterly 14:2-9.

BALDWIN D.C., DAUGHERTY S.R., SELF D.J. (1991). Changes in Moral Reasoning during Medical School. Academic Medicine 66:9, Sept. supplement:S1-3.

BARNETT R. (1990). The Idea of Higher Education. Buckingham: The Society for Research into Higher Education & Open University Press.

BARNETT R. (1997). Beyond Competence. In F. COFFIELD, B. WILLIAMSON (editors): Repositioning Higher Education. Buckingham: Society for Research into Higher Education & Open University Press, pp. 27-44.

BARNETT R., EVANS J., REST J. (1995). Faking Moral Judgement on the Defining Issues Test. British Journal of Social Psychology 34:267-78.

BARROWS H., TAMBLYN R. (1980). Problem-Based Learning: An Approach to Medical Education. New York: Springer.

BARROWS H.S. (1992). The Tutorial Process. Springfield IL: Southern Illinois University Medical School.

BARROWS H.S. (1994). Practice-based Learning: Problem-based Learning Applied to Medical Education. Springfield IL: Southern Illinois University School of Medicine.

BEBEAU M.J., BRABECK M.M. (1987). Integrating Care and Justice Issues in Professional Moral Education: A Gender Perspective. Journal of Moral Education 16:189-203.

BEBEAU M.J., REST J.R., YAMOOR C.M. (1985). Measuring Dental Students' Ethical Sensitivity. Journal of Dental Education 49:225-35.

BEBEAU M.J., THOMA S.J. (1996). The Impact of a Dental Ethics Curriculum on Moral Reasoning. Journal of Dental Education 58:684-691.

BELBIN R.M. (1993). Team Roles at Work. Oxford: Butterworth-Heinemann.

BLASI A. (1994). Bridging Moral Cognition and Moral Action: A Critical Review of the Literature. In B. PUKA (editor): Fundamental Research in Moral Development. New York & London: Garland Publishing Inc. pp. 123-67.

BLUM L.A. (1993). Gilligan and Kohlberg: Implications for Moral Theory. In M.J. LARRABEE (editor): An Ethic of Care : Feminist and Interdisciplinary Perspectives. New York: Routledge, pp. 49-68.

BOUD D. (1985). Problem-based Learning in Perspective. In D. BOUD (editor): Problem-Based Learning in Education for the Professions. Sydney: Higher Education Research and Development Society of Australia.

BOYD K. (1987). The Pond Report on Teaching Medical Ethics: Report of a Working Party on the Teaching of Medical Ethics. London: Institute of Medical Ethics.

BRABECK M. (1993). Moral Judgement: Theory and Research on Differences between Males and Females. In M.J. LARRABEE (editor): An Ethic of Care : Feminist and Interdisciplinary Perspectives. New York: Routledge, pp. 33-48.

BRANSFORD J.D., STEIN B.S. (1993). The Ideal Problem Solver: A Guide fo Improving Thinking, Learning, and Creativity. 2nd ed. New York: W.H.Freeman and Company.

BREDEMEIER B.J., SHIELDS D.L. (1994). Applied Ethics and Moral Reasoning in Sport. In J.R. REST, D. NARVAEZ (editors): Moral Development in the Professions: Psychology and Applied Ethics. Hillsdale, New Jersey: Lawrence Erlbaum Associates, pp. 173-88.

BREZINKA W. (1997). Educational Aims, Educational Means, Educational Success : Contributions to a System of Science of Education. Hants: Aldershot.

BRUCE D., BRUCE A. (1998). Engineering Genesis: The Ethics of Genetic Engineering in Non-Human Species. London: Earthscan.

CALLAHAN D. (1980). Goals in the Teaching of Ethics. In D. CALLAHAN (editor): Ethics Teaching in Higher Education. New York: Plenum, pp. 61-80.

CALMAN K.C., DOWNIE R.S. (1988). Education and Training in Medicine. Medical Education 22:48-91.

CAPLAN A.L. (1980). Evaluation and the Teaching of Ethics. In D. Callahan, S. BOK: Ethics Teaching in Higher Education. New York, London: Plenum Press.

CLARKEBURN H., BEAUMONT E., DOWNIE R., REID N. (2000). Teaching Biology Students Transferable Skills. Journal of Biological Education 34:133-138.

CURTIS C.F. (1994). Should DDT Continue to be Recommended for Malaria Vector Control? [review]. Medical and Veterinary Entomology 8:107-12.

DAVIS W.K., NAIRN R., PAINE M.E., ANDERSON R.M., OH M.S. (1992). Effects of Expert and Non-expert Facilitators on the Small-group Process and Student Performance. Academic Medicine 67:470-474.

DEARING R. ET AL. (1997). The Dearing Report: Higher Education in the learning society. London: HMSO.

DE VOLDER M.L., DE GRAVE W.S., GIJSELAERS W. (1985). Peer Teaching: Academic Achievement of Teacher-Led versus Student-Led Discussion Groups. Higher Education 14:643-650.

DOWNIE J.R. (1993). The Teaching of Bioethics in the Higher Education Biologists. Journal of Biological Education 27:34-38.

DOWNIE J.R., ALEXANDER L. (1989). The Use of Animals in Biology Teaching in Higher Education. Journal of Biological Education 23:103-111.

DOWNIE R.S., LOUDFOOT E.M., TELFER E. (1974). Education and Personal Relationships. London: Methuen.

DREYFUS H.L., DREYFUS S.E. (1986). Mind over Machine : The Power of Human Intuition and Expertise in the Era of the Computer. Oxford: Basil Blackwell.

EISENBERG V.H., SCHENKER J.G. (1997). Preembryo research: Medical aspects and ethical considerations. Obstetrical & Gynecological Survey 52:565-574.

EMLER N., RENWICK S., MALONE B. (1983). The Relationship between Moral Reasoning and Political Orientation. Journal of Personality and Social Psychology 45:1073-1080.

ENTWISTLE, N. (1988). Styles of Learning and Teaching. London: David Fulton Publishers.

ENTWISTLE N. (1998). Motivation and Approaches to Learning: Motivating and Conceptions of Teaching. In S. BROWN, S. ARMSTRONG, G. THOMPSON (editors): Motivating Students. London: Kogan Page, pp. 15-23.

FINSTER D.C. (1989). Developmental Instruction: Part I Perry's Model of Intellectual Development. Journal of Chemical Education 66:659-661.

FINSTER D.C. (1991). Developmental Instruction: Part II. Application of the Perry Model to General Chemistry. Journal of Chemical Education 68:752-756.

FLANAGAN O., JACKSON K. (1993). Justice, Care, and Gender: The Kohlberg-Gilligan Debate Revisited. In M.J. LARRABEE (editor): An Ethic of Care: Feminist and Interdisciplinary Perspectives. New York: Routledge, pp. 69-84.

FRIEDMAN M. (1993). Beyond Caring: The De-Moralization of Gender. In M.J. LARRABEE (editor): An Ethic of care : feminist and interdisciplinary perspectives. New York: Routledge, pp. 258-76.

FRISCH N.C. (1987). Value Analysis: A Method for Teaching Nursing Ethics and Promoting the Moral Development of Students. Journal of Nursing Education 26:328-332.

GIBBS J.C., WIDAMAN K.F. (1982). Social Intelligence: Measuring the Development of Sociomoral Reflection. Enlgewood Cliffs, New Jersey: Prentice Hall.

GILLIGAN C. (1982). In a Different Voice : Psychological Theory and Women's Development. Cambridge, Massachusetts: Harvard University Press.

GILLIGAN C. (1993). Reply to Critics. In M.J. LARRABEE (editor): An Ethic of Care : Feminist and Interdisciplinary Perspectives. New York: Routledge, pp. 207-214.

GRAY C. (1997). A Study of Factors Affecting a Curriculum Innovation in University Chemistry. Unpublished PhD manuscript, University of Glasgow.

HACKETT E.J. (1993). A new Perspective on Scientific Misconduct. Academic Medicine 68:September supplement:S72-6.

HALSTEAD J.M. (1996). Liberal Values and Liberal Education. In J.M. HALSTEAD, M.J. TAYLOR (editors): Values in Education and Education in Values. London: The Falmer Press, pp. 17-32.

HARTWELL S. (1995). Promoting Moral Development through Experiential Teaching. Clinical Law Review 1:505-539.

HARVEY, J.M. (1994) An Investigation into Ways of Encouraging the Development of Higher Level Cognitive Skills in Undergraduate Biology Students with Reference to the Perry Scheme of Intellectual Development. Unpublished PhD manuscript, Napier University.

HAU K.T. (1990). Moral Development and the Ability to Fake in Moral Judgement Tests among Chinese Adolescents. Psychologia 33:106-111.

HIRST P.H., PETERS R.S. (1970). The Logic of Education. London: Routledge & K. Paul Kegan.

INHELDER B., PIAGET J. (1958). The Growth of Logical Thinking from Childhood to Adolescence : An Essay on the Construction of Formal Operational Structures. London: Routledge and K. Paul Kegan.

KATUNG M., JOHNSTONE A.H., DOWNIE J.R. (1999). Monitoring Attitude Change in Students to Teaching and Learning in a University Setting: A Study Using Perry's Developmental Model. Teaching in Higher Education 4:43-59.

KATZ G. (1995). Facilitation. In C. ALAVI (editor): Problem-Based Learning in a Health Sciences Curriculum. London and New York: Routledge, pp. 52-70.

KATZ E. (1997). Defending the Use of Animals by Business: Animal Liberation and Environmental Ethics. In E. KATZ (1997): Nature as a Subject: Human Obligation and Natural Community. Lanham, MD London : Rowman & Littlefield, pp. 79-90

KING P.M., KITCHENER K.S., WOOD P.K., ET AL. (1989). Relationships Across Developmental Domains: A Longitudinal Study of Intellectual, Moral and Ego Development. In M.L. COMMONS, J.D. SINNOTT, F.A. RICHARDS, C. ARMON, (editors): Adult Development - Comparisons and Applications of Developmental Models. New York: Praeger, pp. 57-72.

KOHLBERG L. (1976). Moral Stages and Moralization: The Cognitive-Developmental Approach. In T. LICKONA (editor): Moral Development and Behavior: Theory, Research, and Social Issues. New York: Holt, Rinehart and Winston, pp. 31-53.

KOHLBERG L. (1994). The Development of Modes of Moral Thinking and Choices in the Years 10 to 16. New York & London: Garland Publishing.

KRAWCZYK R.M. (1997). Teaching Ethics: Effects on Moral Development. Nursing Ethics 4:57-65.

LOCKWOOD A.L. (1976). Moral Reasoning and Public Policy Debate. In T. LICKONA (editor): Moral Development and Behavior: Theory, Research, and Social Issues. New York: Holt, Rinehart and Winston, pp. 317-325.

LØVLIE L. (1997). The Uses of Example in Moral Education. Journal of Philosophy of Education 31:409-425.

LURIE P., WOLFE S.M. (1997). Unethical Trials of Interventions to Reduce Perinatal Transmission of the Human Immunodefiency Virus in Developing Countries. New England Journal of Medicine 337:853-856.

MCGEORGE C. (1975). The Susceptibility of Faking of the Defining Issues Test of Moral Development. Developmental Psychology 11:108.

MCNEEL S.P. (1994). College Teaching and Student Moral Development. In J.R. REST, D. NARVAEZ, (editors): Moral Development in the Professions: Psychology and Applied Ethics. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers, pp. 1-26.

MATTHEW, B. (1999). Problem-based Learning - A Description of the Process. Unpublished course notes for the Post-graduate Certificate for Higher Education, University of Glasgow.

NORMAN R.G., SCHMIDT H.G. (1992). The Psychological Basis of Problem-based Learning: A Review of the Evidence. Academic Medicine 67:557-565.

NUFFIELD COUNCIL OF BIOETHICS (1999). Genetically Modified Crops: The Ethical and Social Issues. Downloaded on September 5th from: http://www.nuffield.org/bioethics/publication/modifiedcrops/index.html

NUNNER-WINKLER G. (1993). Two Moralities? A Critical Discussion of an Ethic of Care and Responsibility versus Ethic of Rights and Justice. In M.J. LARRABEE (editor): An Ethic of Care : Feminist and Interdisciplinary Perspectives. New York: Routledge, pp. 143-156.

OLEN J., BARRY V. (1996). Applying Ethics: A Text with Readings. 5th ed. Belmont, CA: Wadsworth Publishing Company.

PAGE R., BODE J. (1979). Degree of Susceptibility of Faking of the Ethical Reasoning Inventory. Journal of Educational Research 72:355-356.

PENN W.P.J. (1990). Teaching Ethics - A Direct Approach. Journal of Moral Education 19:124-138.

PERRY W.G.J. (1999). Forms of Ethical and Intellectual Development in the College Years: A Scheme. San Fransisco: Jossey-Bass Publishers.

PETERS R.S. (1970). Ethics and Education. Oxford: George Allen&Unwin Ltd.

PONDER B. (1997). Genetic Testing for Cancer Risk. Science 278:1050-1054.

PRIDEAUX D., FARMER E. (1994). What Is a Good PBL Case? Some Principles For Case Selection. In S.E. CHEN, R. COWDROY, A. KINGSLAND, M. OSTWALD, (editors): Reflections on Problem-based Learning. Sydney: Australian Problem Based Learning Network, pp. 125-140.

PUOLIMATKA T. (1997). The Problem of Democratic Values Education. Journal of Philosophy of Education 31:461-476.

RABOW J., CHARNESS M.A., KIPPERMAN J., ET AL. (1994). William Fawcett Hill's Learning Through Discussion. 3rd ed. Thousand Oaks: Sage Publications.

RACE P. (1998). Teaching: Creating a Thirst for Learning. In S. BROWN, S. ARMSTRONG, G. THOMPSON, (editors): Motivating Students. London: Kogan Page, pp. 47-57.

RAWLS J. (1972). A Theory of Justice. Oxford: Clarendon Press.

RESNIK D. B. (1998). The ethics of science - an introduction. London: Routledge.

REST J.R. (1976). New Approaches in the Assessment of Moral Judgement. In T. LICKONA (editor): Moral Development and Behavior: Theory, Research, and Social Issues. New York: Holt, Rinehart and Winston, pp. 198-218.

REST J.R. (1982). A Psychologist Looks at the Teaching of Ethics. Hastings Center Report 12:29-36.

REST J.R. (1983). Morality. In J.H. FLAVELL, E.M. MARKMAN, (editors): Cognitive Development. 4th ed. New York: John Wiley & Sons, pp. 556-629.

REST J.R. (1986). DIT Manual. 3rd edition, 8/90 Revision. Minneapolis: Center for the study of ethical development, University of Minnesota.

REST J.R. (1986b). Programs and interventions. In J.R. REST (editor): Moral Development: Advances in Research and Theory. New York: Praeger.

REST J.R. (1988). Can Ethics Be Taught in Professional Schools? The Psychological Research. Easier Said Than Done 1:22-26.

REST J.R. (1994). Background: Theory and Research. In J.R. REST, D. NARVAEZ, (editors): Moral Development in the Professions: Psychology and Applied Ethics. Hillsdale, New Jersey: Lawrence Erlbaum Associates, pp. 1-26.

REST J.R., DEEMER D. (1986). Life Experiences and Developmental Pathways. In J.R. REST (editor): Moral Development: Advances in Research and Theory. New York: Praeger, pp. 28-58.

REST J.R., EDWARDS L., THOMA S. (1997). Designing and Validating a Measure of Moral Judgement: Stage Preference and Stage Consistency Approaches. Journal of Educational Psychology 89:5-28.

REST J.R., NARVAEZ D. (1994). Summary: What's Possible? In J.R. REST, D. NARVAEZ, (editors): Moral Development in the Professions: Psychology and Applied Ethics. Hillsdale, New Jersey: Lawrence Erlbaum Associates, pp. 212-24.

REST J.R., NARVAEZ D., THOMA S.J., BEBEAU M.J. (1999b). DIT2: Devising and Testing a New Instrument of Moral Judgement. Journal of Educational Psychology, 91:644-659.

REST J.R., NARVAEZ D., BEBEAU M.J., THOMA S.J. (2000). A Neo-Kohlbergian approach to moral judgement: An overview of Defining Issues Test research. Educational Psychology Review, in press.

REST J.R., THOMA S., NARVAEZ D. (1999a). Moral Judgement: Stages and schemas. Presented at the 1999 Annual Conference of the Association for Moral Education, Minneapolis.

RHOLES W.S., BAILEY S., MCMILLAN L. (1982). Experiences That Motivate Moral Development: The Role of Cognitive Dissonance. Journal of Experimental Social Psychology 18:524-536.

ROSS R. (1994). Skilful Discussion. In P.M. SENGE, A. KLEINER, C. ROBERTS, R.B. ROSS, B.J. SMITH, (editors): The Fifth Discipline Fieldbook: Strategies and Tools for Building a Learning Organisation. New York: Currency Doubleday, pp. 385-391.

ROTBLAT, S. (1999). A Hippocratic Oath for Scientists. Science 286:1475.

ROWE C. (1995). Ethics in Ancient Greece. In P. SINGER (editor): A Companion to Ethics. Cambridge, Massachussets: Blackwell, pp. 121-132.

SCHLAEFLI A., REST J.R., THOMA S.J. (1985). Does Moral Education Improve Moral Judgement? A Meta-Analysis of Intervention Studies Using the Defining Issues Test. Review of Educational Research 55:319-352.

SELEPENG D. (1999) Evaluation of Attitude Change. Unpublished manuscript, presented at a staff meeting at University of Glasgow.

SELF D.J., ELLISON E.M. (1998). Teaching Engineering Ethics: Assessment of Its Influence on Moral Reasoning Skills. Journal of Engineering Education 87:29-34.

SELF D.J., WOLINSKY D.C., BALDWIN D.C.J. (1989). The Effect of Teaching Medical Ethics to Medical Students' Moral Reasoning. Academic Medicine 64:755-759.

SHAPIRO D.A. (1999). Choosing the Right Thing to Do. San Francisco: Berrett-Koehler Publishers.

SILVER M., WILKERSON L. (1991). Effects of Tutors with Subject Expertise on the Problem-based Tutorial Process. Academic Medicine 66:298-300.

SMITH J.A., BOYD K.M. (1991). Live in the Balance: The Ethics of Using Animals in Biomedical Research (A Report of a Working Party of the Institute of Medical Ethics). Oxford: Oxford University Press.

SNAREY J.R. (1985). Cross-Cultural Universality of Social-Moral Development: A Critical Review of Kohlbergian Research. Psychological Bulletin 97:202-232.

THOMA S. (1984). Do Moral Education Programs Facilitate Moral Judgement? A Meta-Analysis of Studies Using the Defining Issues Test. Moral Education Forum 9:20-24.

THOMA S.J., REST J.R. (1986). Moral Judgement, Behavior, Decision Making, and Attitudes. In J.R. REST (editor): Moral Development: Advances in Research and Theory. New York: Praeger, pp. 133-175.

THOMA S.J., REST J.R. (1999). The Relationship between Moral Decision-making and Patterns of Consolidation and Transition in Moral Judgement Development. Developmental Psychology 35:323-334.

THOMA S.J., REST J.R., DAVISON M.L. (1991). Describing and Testing a Moderator of the Moral Judgement and Action Relationship. Journal of Personality and Social Psychology 61:659-669.

WEISBERG M., DUFFIN J. (1995). Evoking The Moral Imagination: Using Stories to Teach Ethics and Professionalism to Nursing, Medical, and Law Students. Change 27:20-27.

WESTBERG J, JASON H. (1996). Fostering Learning in Small Groups: A Practical Guide. New York: Springer Publishing Company.

WILMUT I. (1998). Pharmaceuticals from Milk: Producing Pharmaceuticals in Sheep Milk. In D. BRUCE, A. BRUCE (editors): Engineering Genesis: The Ethics of Genetic Engineering in Non-Human Species. London: Earthscan, pp. 60-62.

WILSON J. (1990). A New Introduction to Moral Education. London: Cassell.

WILSON M. (1998). Vaccination Made Easy: Proteins from Plants, Using Modified Plant Viruses. In D. BRUCE, A. BRUCE, (editors): Engineering Genesis: The Ethics of Genetic Engineering in Non-Human Species. London: Earthscan, pp. 50-54.

WOODS R.G., BARROW R.S.C. (1975). An Introduction to Philosophy of Education. London: Methuen.

WORLD MEDICAL ASSOCIATION (1996). Declaration of Helsinki. In B.A. BRODY (1998): The Ethics of Biomedical Research: An international perspective. New York: Oxford University Press.

WRINGE C. (1988). Understanding Educational Aims. London: Unwin Hyman.

ZAR J.H. (1999). Biostatistical Analysis. 4th ed. New Jersey: Prentice Hall.

Appendix I: University Survey

During September/October 1998 UK Universities offering a bioscience degree were sent an e-mail with the following questions:

Does your university include bioethics components in the science curriculum? Yes/No

Or do you have special courses available in bioethics for students of biomedical or life sciences? Yes/No

If the answer is no, thank you for your response,
if the answer is yes, please look at the questions below.

Is the bioethics teaching included in compulsory studies? Yes/No
Or are the courses elective? Yes/No
Or both Yes/No

At which stage of the studies are the bioethics courses or components available?
1st year undergraduate
2nd year undergraduate
3rd year undergraduate
4th year undergraduate
post-graduate

Is the bioethics teaching provided in form of:
lectures, (Yes/No)
tutorials, (Yes/No)
self-study modules (Yes/No)?

How are the bioethics courses or the components of bioethics within other courses assessed?

Is there a set of objectives clarified for the bioethics teaching?

If yes, what are they?

The e-mails were sent to the heads of departments/faculties and if there was no response to the first e-mail, the second one was sent to the same recipient.

The following tables detail the Universities that did/did not reply and the form of ethics teaching in the Universities that included ethics in their curriculum.

Replied	**Type of ethics, if any**
University of Aberdeen	Compulsory, 4th Year, examined
University of Bath	Optional, 3rd&4th Year
Queens University Belfast	
University of Birmingham	
University of Bradford	
University of Bristol	Ethical use of animals course
University of Cambridge	
Cardiff University, Wales	
University of Central Lancashire	Optional self-study, 2nd Year
University of Derby	
University of Dundee	
University of Durham	
University of East Anglia	
University of East London	
University of Edinburgh	
University of Essex	Inclusion of ethics in lectures, exam questions
University of Exeter	Inclusion of ethics in lectures, exam questions
Heriot Watt University	
University of Hertfordshire	
University of Hull	
Liverpool John Moores	
University of Keele	
Kingston University	
University of Leicester	
University of Liverpool	
University of Manchester	
Manchester Metropolitan	
University of Oxford	
University of Plymouth	Tutorial discussions, 3rd Year
University of Reading	Optional course
University of Sheffield	
University of St. Andrews	Inclusion of ethics in lectures, exam questions
University of Stirling	
University of Strathclyde	
University of Wales, Swansea	Compulsory lectures, 2nd year on ethical issues
University of York	

No reply

University of Leeds
University of Wales, Bangor
De Montfort University
University of Greenwich
Liverpool Hope University
University of Middlesex
Napier University
Nottingham Trent University

University of Nottingham
Open University
University of Southampton
University of Sussex
University of Westminster
University of Portsmouth
Lancaster University
Cardiff University
University of Newcastle Upon Tyne

Appendix II: One Theory

To a certain extent at least, the first perspective of Gilligan and the pre-conventional stage of Kohlberg can be considered similar, both pre-occupied with the self and the idea that moral issues have validity only through the self. Again, the second perspective of Gilligan's moral development and the conventional stage in Kohlberg share the concept of accepting social norms as the guidance through moral problems, even though the perception of the content of these conventional moral rules is different. Gilligan is describing a traditional agenda of femininity, of care, harmony, compassion and self sacrifice, while Kohlberg depicts a level of social organisation, a system of rights, rules, respect and fairness. The primary moral imperative for Gilligan is thus non-violence and care and for Kohlberg it is justice.

Also in the transition from the second perspective to the third or from Kohlberg's conventional stage to post-conventional, both sexes move away from absolutes, though the absolutes themselves differ. In Gilligan's description of development it is the absolute of care, defined initially as not hurting others. The recognition of the need for personal integrity gives rise to the claim for equality embodied in the concept of rights. This changes the understanding of relationships and transforms the definition of care. According to the theory of Kohlberg, the absolutes to move away from are truth and fairness, defined by concepts of equality and reciprocity. These are called into question by experiences of multiple truths and the existence of difference between self and others. This leads to a relativising of equality in the direction of equity and gives rise to an ethic of generosity and care. In both theories the existence of two contexts for moral decision makes judgement by definition contextually relative and leads to a new understanding of responsibility and care (Gilligan, 1993; Kohlberg 1976).

But the differences are not only in content, but also in the nature of moral dilemmas, determinants of moral obligation, and views of the self as moral agent. For Gilligan the moral dilemmas are threats to harmony and relationships, while for Kohlberg conflicting rights are the source of moral dilemmas. The determinants of moral obligation for Kohlberg are principles of justice, while for Gilligan they are relationships. In Gilligan's view, moral agents are connected and attached within a net of relationships while in Kohlberg's view the agent is separate and individual, not tied to other people (Brabeck, 1993).

Moral development as depicted both by Gilligan and Kohlberg seems to entail a similar integration of rights and responsibilities. According to the development process as presented by Gilligan, the integration of rights and responsibilities takes place through an understanding of the psychological logic of relationships. As for Kohlberg, recognition through experience of the need for more active responsibility in taking care corrects the potential indifference of a morality of non-interference and turns attention from the logic to the consequences of choice. In Gilligan's theory subjects come to see the violence inherent in inequality, while Kohlberg's subjects come to see the limitations of a conception of justice blinded to the differences in human life (Gilligan, 1993).

Possibility of a consensus

Rest has attempted to merge these disparate but related concerns presented by Gilligan and Kohlberg into a comprehensive moral theory. He has described four components of morality: 1) interpretation of a situation as moral and the appropriate affective response (outrage at a wrong committed, sorrow at a pain inflicted); 2) judgement about what constitutes the moral ideal or the just outcome; 3) decision about a course of action; and 4) an appropriate behavioural response. Rest places Kohlberg's theory in component 2, judgement of the ideal, while Gilligan speaks of components 1 and 3 (Rest, 1983).

When Gilligan's and Kohlberg's theories are taken together, the moral person is seen as one whose moral choices reflect reasons and deliberate judgement that ensure justice be accorded to each person while maintaining a passionate concern for the well-being and care of each individual. Justice and care are then joined; the demands of universal principles and specific moral choices are bridged, and the need for autonomy and for interconnection are united in an enlarged and more adequate conception of morality (Brabeck, 1993).

At least from the point of view of Gilligan, the merger of these two moralities, - Kohlberg's based on impartiality and justice, and the one she puts forward of partiality and care, are not incompatible. Gilligan, according to Blum (1993), holds that there is an appropriate place for impartiality, and universal principle within morality, and that a final mature morality involves a complex interaction and dialogue between the concerns of impartiality and those of personal relationship and care.

There is no logical reason why both care and justice considerations cannot be introduced, where relevant, into one and the same reasoning episode. It is possible to imagine a person who is both caring and just and who, in addition, has finely toned sensitivities for perceiving moral salience and seeing particular problems as problems of great diversity. This is not to deny that in some cases, construing a particular problem from both perspectives will block moral clarity about what should be done, nor is it to deny that it is important to have, even imperfect, decision-making procedures to resolve such conflicts (Flanagan and Jackson, 1993).

The concepts of justice and care can therefore be considered as mutually compatible. People who treat each other justly can also care about each other. Justice is relevant to personal relationships, because considerations of justice determine appropriate ways to treat friends and intimates. Justice as it bear on relationships among friends and family, or other close personal ties, might not involve duties which are universalisable , but this does not entail the irrelevance of justice among friends and intimates. Justice is present in two ways. One, justice plays a part in distinguishing what people can justly expect in close relationships. The maintenance of a relationship requires effort by participants. One intimate may bear a much greater burden in sustaining a relationship than the other(s) and may derive less support, concern and so forth than she deserves for her efforts. Two, the opportunity for harm within a close relationship (physical injury, sexual assault, psychological blackmail etc.) is significant and if such harming takes place, it requires rectification of some

sort. These are considerations of justice about the limits of actions within close relationships. In a similar manner care is present in the public domain. It shows itself, perhaps, in foreign aid, welfare programmes, famine and disaster relief, or other social programmes designed to relieve suffering and to attend to human needs (Friedman, 1993).

It is also possible to consider that the division of the care and justice-oriented approaches to moral problems coincide with the distinctions between perfect and imperfect duties. Perfect duties are negative duties - that is, duties of omission; 'do not kill', 'do not cheat' etc. Imperfect duties are in comparison positive duties, duties of commission, duties to act. These are not necessarily specific guides to action, but maxims like 'practice charity' or 'love thy neighbour'. Perfect duties should be followed in all situations. This is possible because they do not require one to act, at least in non-conflictual cases. Imperfect duties, on the other hand, cannot be observed completely; it is impossible, for example, to practice charity all the time and regarding everybody. Imperfect duties are therefore contextual and require judgement of the situation in their application. In comparison, in a non-conflictual case, the following of perfect duties presupposes scarcely any knowledge of situational specifics. As all that is required is not to act in a specific way at any time or location and with regard to everybody, all one needs to know is some general empirical facts (giving someone an overdose of morphine will kill them) or at best some narrowly limited specific facts (in case of a child or otherwise weak individual, a lesser amount of morphine will lead to death). But this is true only if there are no conflicting moral duties and, in case of conflict, one believes there are rules absolutely without exceptions. Such an extreme position is not supported by many and therefore even concerning perfect duties one needs to engage in justifiability of exceptions to rules which are widely accepted (Nunner-Winkler 1993).

The difference between perfect and imperfect duties is, hence, that imperfect duties are embedded in their context and contextual assessment and justification of action is necessary for all moral decisions. Perfect duties are less contextually bound, as rules can be followed without contextual appreciation or judgement, while perfect duties need not be considered to be totally context-free and absolute. Including contextual appreciation in fulfilling perfect duties is a less dominant feature than in imperfect duties. First, there are rules, and before their application some contextual assessment might be necessary or at least desirable: but rules apply unless otherwise proven. In imperfect duties, contextual assessment is necessary in all situations as none of the imperfect duties can be proposed without contextual information.

The ethics of care and responsibility can be described as having a primary orientation to imperfect duties and the ethic of rights and justice to perfect duties. At the same time, both kind of duties can be part of the same morality. Therefore there need not be two different moralities, but a dissonance in the preference of which type of duties take precedence in conflicting situations. The order of priority need not be rigid, but precisely a mode of preference.

Appendix III - Perry's Scheme

Perry's scheme is a description of evolution in students' interpretation of their lives. The evolution consists of a progression in certain forms in which the students construe their experience. These forms characterise the structures which the students explicitly or implicitly impute to the world, especially those structures in which they construe the nature and origins of knowledge of value and of responsibility.

Methodology

In the 1950's and 1960's Perry conducted a longitudinal study involving lengthy interviews of students from various classes at Harvard and Ratcliffe. From analysis of these interviews, a 'Scheme' of intellectual and ethical development for higher education students emerged. The starting point was to illustrate the variety in students' response to the impact of intellectual and moral relativism, which the research team had encountered in their role as student counsellors. The original idea was to measure personality differences in relation to learning. A selection of students presenting a wide range of views in a Checklist of Educational Views (CLEV) from the ultimately dualistic to mature relativism, were interviewed at the start of their first year and again at the end of the year. Some students changed their scores markedly between the two points of measurement. In the original study 98 tape-recorded interviews were collected, including 17 complete four-year records. The interviews were open-ended - "what do you think has influenced you most during the year?". Based on these preliminary results a larger study including 67 complete four-year records was conducted. From these interviews the researchers started to detect a common sequence of challenges to which each student addressed himself in his own particular way. The way the students addressed the challenges in both academic and social life seemed to represent a coherent development of forms in which they experienced values and in the form in which they construed the world. The formulation of theory from the interviews could only be done by inference, as the students rarely spoke explicitly at the abstract level of developmental theory. The interviews were presented to independent observers and they were asked to identify which stage of the hypothesised developmental scheme best represented the dominant form of structuring of the world in each report. Agreement between independent observers was the measure of validity of existence of the developmental scheme. From these interviews a scheme of moral and intellectual development was constructed.

The main line of development has nine positions which can be grouped into three categories.

Dualism

Basic Challenge: Impact of multiplicity

Position 1 - basic duality: The student sees the world in polar terms of we-right-good vs. other-wrong-bad. All issues are construed in terms of sweeping and unconsidered differentiation between in-group vs out-group. The right answers to everything exist

in the absolute and they are known to the authority whose role is to mediate (teach) them. Knowledge and goodness are perceived as quantitative increase of distinct rightness to be collected by hard work and obedience. One's own role and level of personal responsibility consists of simple obedience. Moral decision making consists of commitment to memory, through hard work of an array of discrete items - correct responses, answers, and procedures, as assigned by an authority. The existence of absolute answers is taken for granted. In very primitive forms the truth of the authority is considered absolute. This is the level of epistemological innocence. Very few university students hold or persist in position 1: it is, rather, a scheme representing how small children divide their world between family and the vague inchoate outside.

Position 2 - Multiplicity pre-legitimate: The student perceives diversity of opinion and uncertainty and accounts for them as unwarranted confusion and poorly qualified authorities or as exercises set by the authority 'so we can learn and find the answer ourselves'. This is the first step in the journey from epistemological innocence. The students at this level of development often take a stand in opposition to what they perceive as the 'vague theorising' of academic authorities and take a bold step in personal individuation, but then quickly find out that they have painted themselves into a corner. Alternatively the students who perceive diversity only as the authority presenting complexities as a mere exercise, are taking a far less radical step in personal development, but find themselves in a more flexible position in the future. All the same diversity is perceived only as a teaching method for the students to learn the 'truth' for themselves. Diversity is still alien and the authority can be perceived to be a poor authority who is failing in the role of mediator of the truth.

Position 3 - multiplicity subordinate: The student accepts diversity and uncertainty as legitimate but still temporary in areas where the authority has not yet found the answers. The epistemology has room for legitimate human uncertainty, but this does not affect the nature of truth itself, only human relation to it. This accommodation loosens the tie between authority and the absolute - uncertainty is now unavoidable in the present. This proposes a procedural problem - how in an education institution where the students' every answer is evaluated are the answers judged if even the authority does not yet know the right answer yet: is not any answer as good as another? Rightness and hard work vanish as standards and good expression (the importance of quantity) seems to be left as the only criteria of judgement. The student solves this tension by focusing on an effort to reappraise what is that 'they want' and will now listen with more open ears to what the instructors say they are up to.

Contextual relativism

Basic challenge: the instability of the self in a diffuse relativism

Position 4 - Multiplicity correlate and relativism subordinate. In their efforts to develop some rational which would account for the anomaly of being judged over issues where truth is not known, the students develop one of two alternative methods. These alternatives are developmental equivalents in that they represent an ultimate extension or accommodation of the old fundamentally dualistic structure before its

capitulation to the vision of generalised contextual relativism in position 5. A) *Multiplicity correlate*. The student perceives legitimate uncertainty to be extensive and raises it to the status of an unstructured epistemological realm of its own in which 'anyone has a right to his own opinion' and 'they have no right to call me wrong'. Moreover as students find the area to be ever larger than expected, and the day of revelation of the truth ever more remote, they claim for multiplicity a domain of its own, an epistemological equal, over against the authorities realm where right and wrong still prevail. This approach is typical for those taking an oppositional role in position 2. B) *Relativism subordinate*. The student discovers qualitative contextual relativistic reasoning as a special case of 'what they want' within the authority's realm. This is a more common pathway from position 3 to the vision of relativism in position 5. The paradox of this position is that students are learning to think independently, because they believe that is what the authority wants - independent-like thought gets good grades. Genuine independence of thought is an issue to be met later. The requirement that an answer be reasonable raises the possibility that some questions may have *some* legitimate answers instead of *one* or *any number* of equally legitimate answers. The difficulty of making a dualistic determination as to whether a given opinion or answer is reasonable or not will then inevitably lead to the discovery of degrees of reasonableness which is typical of position 5.

Position 5 - Relativism correlate, competing or diffuse: This is a drastic revolution. Up till now students have been able to assimilate the new to the fundamental dualistic structure: this is no longer possible. The student perceives all knowledge and values (including authority's) as contextual and relativistic and subordinates dualistic right-wrong functions to the status of special case, in context. This is the major shift, for relativism to be promoted from its status of special case to the status of context. Students discover not only a multiplicity of point of view about such matters as literature, history, or politics, but a pattern within each points of view, an interdependency of parts within the whole, which gave each 'point of view' its special character, its coherence, its integrity. This is a quiet revolution of the relativistic thinking of position 4 slowly becoming habitual.

Position 6 - Commitment foreseen: The student apprehends the necessity of orienting himself in a relativistic world through some form of personal commitment. When one is confronted with an infinite universe of potential contexts for truth and care, one is threatened with loss of identity - a humanly unbearable disorientation. Other than by rejecting relativism, one can solve the threat to identity in three ways: 1) to deny any meaning beyond one's immediate, passive responses, 2) to exploit the situation deliberately by becoming an active opportunist of the relativist reality, or 3) taking a step towards commitment - becoming an agent who chooses the aspects in his/her life and takes responsibility for them. The first two are defensive, while the third option provides a way forward. Position 6 is the moment of realisation of the responsibility of the individual in relating him/herself to the relativistic world. In position 6 commitment is foreseen as the resolution of the problems of relativism, but it has not yet been experienced.

Commitment in Relativism

Basic challenge: responsibilities of commitment. No major reconstructing is apparent in the last stage of development: the drama of maturation has been stabilised. The development is therefore more qualitative than structural - one is already aware of the need to make commitment. Positions 7, 8, and 9 are therefore less distinct and more an expression of degrees of maturation

Position 7: The student makes initial commitment in some area - this can be a career or political orientation for example. The student is quite taken up by the content of the commitment.

Position 8: The student experiences the implications of commitment and explores the subjective stylistic issues of responsibility - deciding between the ways to fulfil one's commitment.

Position 9: The student experiences the affirmation of identity among multiple responsibilities and realises commitment as an ongoing, unfolding activity through which he expresses his life style. Very few of the students in Perry's sample had reached position 9 - it is rather a way to round out the scheme by extrapolation, as much as position 1 did in the other direction.

In any of the positions in the developmental scheme, a person may suspend, nullify or even reverse the process of growth: 1) a student may pause for a year or more, often quite aware of the step that lies ahead as if waiting or gathering the force (temporising). 2) The student may entrench him/herself, in anger or hatred of otherness, in the me-they or we-other dualism of the early positions (retreat). 3) Students may settle for exploiting the detachment offered by some middle position in the scale, in the deeper avoidance of personal responsibility known as alienation (escape). These are elements of a growth that is rarely linear and more usually wavelike. Growth usually occurs in surges and between these surges students might pause to explore the implications of the new position, or they might wait for the resurgence of strength to meet the next challenge, or even detach themselves completely and/or retreat to an old position. Every moment between surges involves a risk of the forces of growth being denied.

Perry's scheme is dominantly used as a description of intellectual development (Finster, 1989) or measuring attitude change in students to teaching and learning (Katung *et al.*, 1999), but was originally seen by Perry as also encompassing moral development.

Opinions on Social Issues

0	0	0	0	0	0	0
1	1	1	1	1	1	1
2	2	2	2	2	2	2
3	3	3	3	3	3	3
4	4	4	4	4	4	4
5	5	5	5	5	5	5
6	6	6	6	6	6	6
7	7	7	7	7	7	7
8	8	8	8	8	8	8
9	9	9	9	9	9	9

Use ONLY a pencil or a black pen. Do NOT fold or bend this form.
For each question choose ONE answer and mark your choice with a stroke like this : ⊸

This questionnaire is aimed at understanding how people think about social problems.
Different people have different opinions about questions of right and wrong.
Therefore there are no right answers to the questions in this questionnaire in the
way there is a right answer to a math problem. We would like you to tell us what YOU
think about several issues. The answers will be fed to a computer to find averages
and trends for the whole group, and your answers will not be identified or appear in
your academic record.
The questionnaire has three parts. Work through them in the order they are here.
You have 30 minutes to complete the questionnaire.

Surname:	Forename(s):

PART I

*Read the following scenario and then list below no more than five issues/questions YOU believe
should be considered when making the decision.*

A research group is planning a project to create a cow that would produce milk containing
a protein that could be used to treat patients with cystic fibrosis, which is a very
debilitating human disease leading to premature death. Other pharmaceutical methods to
produce this protein have not been successful or they have been very expensive.
The plan is to introduce a new gene to the cow from another animal.
The new gene will be introduced by nuclear transfer, a tecgnique also used in cloning.
The group hopes to develop its research findings into a commercial product

	Yes	No	I am not sure

Do you think the research should go ahead?

Please list below no more than five issues/questions YOU think should be considered
when deciding whether to start the research:

1.

2.

3.

4.

5.

Which one of these reasons is the most influential in your decision-making   1   2   3   4   5
(please state the number of your response)

Henriikka Clarkeburn, June 2000

▬ PART II (Instructions)   — —

In this part you will be asked to give your opinions about several stories. Here is an
example, to show you what we will be asking you to do:

Frank Jones has been thinking about buying a car. He is married with two children and
earns an average income. The car he buys will be his family's only car. It will be used
mostly to get to work and to drive around town, but also sometimes for vacation trips.
In trying to decide what car to buy, Frank realises that there were a lot of questions
to consider. If you were Frank, how important would each of the following questions
be in deciding to buy a car?

(On the left hand side fill in ONE of the boxes for EACH question, as shown below)

▬   Great Much Some Little No

1. Whether the car dealer was in the same block as where Frank lives
(in this sample the person answering the questions did not
think this was important in making the decision)

▬   Great Much Some Little No

2. Would a used car be more economical in the long run than a new car
(the tick here indicates that this is a very important
issue to the person answering the question).

▬   Great Much Some Little No

3. Would a large roomy car be better than a compact car?

▬   Great Much Some Little No

4. Whether the cubic inch displacement was at least 200
(If you are unsure about the meaning of the statement, as this person was,
then mark it 'no importance')

From the list of questions above, select the most important one of the whole group
and put the number of the question on the top line below. Do likewise for the second,
third and fourth most important choices.

From the list above, select the four most important statements:

		1	2	3	4
▬	Most important		▬		
▬	Second most important			▬	
▬	Third most important	▬			
▬	Fourth most important				▬

PART II

\- \- \-

Story 1: Heinz and the Drug

In Europe a woman was near death from a special kind of cancer. There was one drug that the doctors thought might save her. It was a form of radium that a pharmacist in the same town had recently discovered. The drug was expensive to make, but the pharmacist was charging ten times what the drug costs to make. He paid £200 for the radium and charged £2000 for a small dose of the drug. The sick woman's husband, Heinz, went to everyone he knew to borrow the money, but he could only get together about £1000, which is half of the cost. He told the pharmacist that his wife was dying and asked him to sell it cheaper or let him pay later. But the pharmacist said "No, I discovered the drug and I am going to make money from it." So Heinz got desperate and began to think about breaking into the man's store to steal the drug for his wife.

Should Heinz steal the Drug?   Should steal it   Can't decide   Should not steal it

How IMPORTANT are the following statements in making the decision:

Great   Much   Some   Little   No

1. Whether a community's laws are going to be upheld.

Great   Much   Some   Little   No

2. Isn't it only natural for a loving husband to care so much for his wife that he'd steal.

Great   Much   Some   Little   No

3. Is Heinz willing to risk getting shot as a burglar or going to jail for the chance that stealing the drug might help.

Great   Much   Some   Little   No

4. Whether Heinz is a professional wrestler, or has considerable influence with professional wrestlers.

Great   Much   Some   Little   No

5. Whether Heinz is stealing for himself or doing this solely to help someone else.

Great   Much   Some   Little   No

6. Whether the pharmacist's rights to his inventions have to be respected.

Great   Much   Some   Little   No

7. Whether the essence of living is more encompassing than the termination of dying, socially and individually.

Great   Much   Some   Little   No

8. What values are going to be the basis for governing how people act towards each other.

Great   Much   Some   Little   No

9. Whether the pharmacist is going to be allowed to hide behind worthless law which only protects the rich anyhow.

Great   Much   Some   Little   No

10. Whether the law in this case is getting in the way of the most basic claim of any member of society.

Great   Much   Some   Little   No

11. Whether the pharmacist deserves to be robbed for being so greedy and cruel.

Great   Much   Some   Little   No

12. Would stealing in such a case bring about more total good for the whole society or not.

From the list above, select the four most important statements:

	1	2	3	4	5	6	7	8	9	10	11	12
Most important	1	2	3	4	5	6	7	8	9	10	11	12
Second most important	1	2	3	4	5	6	7	8	9	10	11	12
Third most important	1	2	3	4	5	6	7	8	9	10	11	12
Fourth most important	1	2	3	4	5	6	7	8	9	10	11	12

Story 2: Escaped Prisoner ▬ ▬ ▬ ▬

A man had been sentenced to prison for 10 years. After one year, however, he escaped
from prison, moved to a new area of the country, and took on the name of Thompson.
For 8 years he worked hard, and gradually saved enough money to buy his own business.
He was fair to his customers, gave his employees top wages, and gave most of his profits
to charity. Then one day, Mrs. Jones, an old neighbour, recognised him as the man who
had escaped prison 8 years before, and whom the police had been looking for.

Should Mrs Jones report Mr Thompson to   Should report him   Can't decide   Should not report him
the police and have him sent back to prison?

How IMPORTANT are the following statements in making the decision:

Great Much Some Little No

1. Hasn't Mr. Thompson been good enough for such a long time to prove
he isn't a bad person?

Great Much Some Little No

2. Every time someone escapes punishment for a crime, doesn't that just
encourage more crime?

Great Much Some Little No

3. Wouldn't we be better off without prisoners and the oppression of our
legal system?

Great Much Some Little No

4. Has Mr. Thompson really paid his debt to the society?

Great Much Some Little No

5. Would society be failing what Mr. Thompson should fairly expect?

Great Much Some Little No

6. What benefits would prisons be apart from society, especially for
a charitable man?

Great Much Some Little No

7. How could anyone be so cruel and heartless as to send Mr. Thompson
to prison?

Great Much Some Little No

8. Would it be fair to all prisoners who had to serve their full sentences if
Mr Thompson was let off?

Great Much Some Little No

9. Was Mrs. Jones a good friend of Mr. Thompson?

Great Much Some Little No

10. Wouldn't it be a citizen's duty to report an escaped criminal, regardless of
the circumstances?

Great Much Some Little No

11. How would the will of the people and the public good be served?

Great Much Some Little No

12. Would going to prison do any good for Mr. Thompson or protect anybody?

From the list above, select the four most important statements:

	1	2	3	4	5	6	7	8	9	10	11	12
Most important												
Second most important	1	2	3	4	5	6	7	8	9	10	11	12
Third most important	1	2	3	4	5	6	7	8	9	10	11	12
Fourth most important	1	2	3	4	5	6	7	8	9	10	11	12

Henriikka Clarkeburn, June 2000

Story 3: Newspaper   — — — —   ▩

Fred, a senior in high school, wanted to publish a newspaper for students so that he could express many of his opinions. He wanted to speak out against the Vietnam war and against some of the school rules, like the rule of forbidding boys to wear long hair.

When Fred started his newspaper, he asked the principal for permission. The principal said it would be all right if before every publication Fred would turn over all his articles for the principal's approval. Fred agreed and turned several articles in for approval. The principal approved all of them and Fred published two issues of the paper in the next two weeks.

But the principal had not expected that Fred's newspaper would receive so much attention. Students were so excited by the paper that they began to organise protests against the hair regulation and other school rules. Angry parents objected to Fred's opinions. They phoned the principal, telling him that the newspaper was unpatriotic and should not be published. As a result of rising excitement, the principal ordered Fred to stop publishing. He gave as a reason that Fred's activies were disruptive to the operation of the school.

Should the principal stop the newspaper?   Should stop it   Can't decide   Should not stop it   ▩

How IMPORTANT are the following statements in making the decision:

Great Much Some Little No
  1. Is the principal more responsible to the students or the parents?   ▩

Great Much Some Little No
  2. Did the principal give his word that the newspaper could be published for a   ▩
  long time, or did he just promise to approve the newspaper one issue at a time?

Great Much Some Little No
  3. Would the students start protesting even more if the principal stopped   ▩
  the newspaper?

Great Much Some Little No
  4. When welfare of the school is threatened, does the principal have the right   ▩
  to give orders to students?

Great Much Some Little No
  5. Does the principal have the freedom of speech to say 'no' in this case?   ▩

Great Much Some Little No
  6. If the principal stopped the newspaper, would he be preventing full discussion   ▩
  of important problems?

Great Much Some Little No
  7. Whether the principal's order would make Fred lose faith in the principal.   ▩

Great Much Some Little No
  8. Whether Fred was really loyal to his school and to his country.   ▩

Great Much Some Little No
  9. What effect would stopping the newspaper have on the student's education   ▩
  in critical thinking and judgement?

Great Much Some Little No
  10. Whether Fred was in any way violating the rights of others in publishing   ▩
  his own opinions.

Great Much Some Little No
  11. Whether the principal should be influenced by some angry parents when   ▩
  it is the principal who knows best what is going on in the school.

Great Much Some Little No
  12. Whether Fred was using the newspaper to stir up hatred and discontent.   ▩

From the list above, select the four most important statements:

	1	2	3	4	5	6	7	8	9	10	11	12	
Most important	1	2	3	4	5	6	7	8	9	10	11	12	▩
Second most important	1	2	3	4	5	6	7	8	9	10	11	12	▩
Third most important	1	2	3	4	5	6	7	8	9	10	11	12	▩
Fourth most important	1	2	3	4	5	6	7	8	9	10	11	12	▩

PART III   — — — — — —

In the following questions you are provided with a pair of opposing statements with
five boxes in between. For example:

	D	M	N	M	D	
I must have background music when I study						I cannot stand any background noise when I am studying

By shading either of the outer boxes you indicate that you agree strongly with the closest
statement. Shading the second boxes means you favour the statement, but less strongly.
The middle box would mean that you do not have a strong preference or you are unsure.
Again there are no right answers, only YOUR view matters.

Mark the appropriate responses - note:
D = Definitely my opinion
M = More or less what I believe
N = Neither of the statements represents my view

Left statement	D	M	N	M	D	Right statement
1. It is not my place to make moral choices.	D	M	N	M	D	1. When I have a moral problem I try to think the answer through myself.
2. When we make moral decisions, the best we can do is to decide what is better or worse in different situations.	D	M	N	M	D	2. I believe we can always make a judgement whether actions are right or wrong and these rules do not change.
3. I don't think teachers should assess my moral arguments if they do not know the right answers themselves yet.	D	M	N	M	D	3. It is important that teachers assessing moral arguments look for logical structure and good reasoning rather then a particular answer.
4. Personal values need to be re-considered from time to time.	D	M	N	M	D	4. Moral values are the same forever.
5. People cannot choose their values, because values are either right or wrong.	D	M	N	M	D	5. I am committed to a set of values I have chosen for myself.
6. I don't enjoy discussing moral problems, unless the teacher can give the right answer in the end.	D	M	N	M	D	6. I am comfortable with discussing my values in the class even when we cannot agree on one right answer in the end.
7. It is almost impossible to answer moral questions without providing arguments to support them.	D	M	N	M	D	7. Answers to moral questions should be short and simple.
8. There are very few absolutely right answers in the world and answers to moral questions are not amongst them.	D	M	N	M	D	8. Moral questions have absolutely right answers just like scientific ones.
9. I don't think discussing moral problems is beneficial for me unless a right answer can be found in the end.	D	M	N	M	D	9. Discussing values with other people gives me a beneficial opportunity to reflect on my own values, even when there is no agreement in the end.
10. I need to commit myself to a set of values even when I am uncertain whether they will always be the right values to have.	D	M	N	M	D	10. I do not doubt that my values are the right values to have.

Henriikka Clarkeburn, June 2000

Appendix V - Computing C -scores from DIT Data

From Rest *et al.* (1999a)

	story	pro con		Stage
1. *SS total* =	$\sum\limits_{i}$	$\sum\limits_{j}$		$\sum X^2(ijm)$ over m

2. $CF = (\sum Xijm)^2 \div$ (total number of items in test)

3. *SS deviation* = SS total - CF

	stage	story		pro-con
4. *SS stage* =	$\sum$	$\sum$		$\sum Xijm^2 \div$(number of items in stage) - CF

5. $C = $ (SS stage $\div$SS(deviation) $\times$ 100

(1) To calculate for each participant the *SS - total*, do the following: (a) For each item that represents stage 2 or stage 3 (22 items in 6 story DIT), square the rating for the item and sum. Then multiply by 20/22 (to adjust to equal number of items for each group of 20). (b) For each item that represents stage 4 (19 items), square the ratings; sum these squares; then multiply by 19/20. (c) For each item that represents stage 5 or stage 6 (21 items), square the rating; sum these squares, then multiply by 20/21. (d) the *SS - total* is the sum of (a) + (b) + (c).

(2) To get *CF*: (a)take each item that represents stage 2 or stage 3, add the ratings of the 22 items together, adjust this total by multiplying by 20/22; (b) take each item that represents stage 4, add the ratings of the 19 items together, adjust this total by multiplying by 19/20. (c) take each item that represents stage 5 or 6, add the ratings of the 21 items together, adjust this total by multiplying by 20/21. (d) add together sums (2a) + (2b) + (2c), square this total, then divide by 60 (the number of items in the entire 6-story DIT).

(3) To get *SS-deviation*, substract CF (2) from *SS-total* (1).

(4) To get *SS-stage* do the following: (a) take each stage-group sum derived in (2a), (2b), and (2c), and square each; (b) add the squares together; (c) divide by 10 ; and then subtract *CF* (2).

(5) The *C-score* for a participants is the *SS-stage* divided by *SS-deviation*, then multiply by 100.

Appendix VI - Perry Judges

First round (all from University of Glasgow unless stated otherwise):
Prof Alex Johnstone, Science Education
Dr. Norman Reid, Science Education
Dr Lisa Schwartz, General Practice
Dr Jen Harvey, University of Edinburgh
Ms Cecilia Edwards, University of Dundee
Dr Craig Gray, Teaching and Learning Services
Ditshupo Selepeng, Centre of Science Education
Prof. Robin Downie, Department of Philosophy

Second round:
Dr. Bob Matthew, Teaching and Learning Service
Dr. Erica McAteer, Teaching and Learning Service
Dr. Rob Hoyle, Department of Chemistry
Dr. James Wilson, Department of Management Studies

Appendix VII - Preliminary Perry Questionnaire

In the following questions you are provided with a pair of opposing statements with five boxes in between. For example:

I must have background music when I study		I cannot stand any background noise when I am studying

By shading either of the outer boxes you indicate that you agree strongly with either statement. Shading the second boxes means you favour the statement, but less strongly. The middle box would mean that you do not have a strong preference or you are unsure.
Again there are no right answers, *only YOUR view matters*.

Mark the appropriate responses - note: D = Definitely my opinion, M = More or less what I believe, and I = I can't decide

<center>D M I M D</center>

1. I don't think discussing moral problems is beneficial for me unless a right answer can be found in the end.					1. Discussing values with other people gives me a beneficial opportunity to reflect on my own values, even when there is no agreement in the end.
2. When we make moral decisions, the best we can do is to decide what is right as far as we can tell in different situations.					2. I believe we can always make a judgement whether actions are right or wrong and these rules do not change.
3. I don't think teachers should assess my moral arguments if they do not know the right answers yet.					3. It is important that teachers assessing moral arguments look for logical structure and good reasoning rather than a particular answer.
4. Personal moral values need to be re-considered from time to time.					4. Personal moral values are the same forever.
5. There are very few absolutely right answers in the world and answers to moral questions are not one of them.					5. Moral questions have absolutely right answers just like scientific ones.
6. I don't enjoy discussing moral problems, unless the teacher can give the right answer in the end.					6. I enjoy discussing my values in the class even when we cannot agree on one right answer in the end.
7. You cannot have a good moral answer without arguments to support it, because moral answers are never straight forward.					7. A good moral answer is short and simple, because you know the right answer.
8. People cannot choose their values, because values are either right or wrong.					8. I am committed to a set of values I have chosen for myself.
9. It is not my place to make moral choices, because right answers have been found already by others.					9. When I have a moral problem I try to think the answer through myself
10. I need to commit myself to a set of values even when I am uncertain whether they will always be the right values to have.					10. I do not doubt that my values are the right values to have.

Appendix VIII - Moral Sensitivity - preliminary results

Story 1 - research laboratory

	no issues	1 issue	2 issues	3 issues	4 issues	5 issues
L1	20%	13%	20%	27%	13%	7%
L3	0%	22%	11%	28%	17%	22%

L1 average = 2.1   L3 average = 2.5

Employment of support staff	13	Who decides on the research	2
Benefits to asthma sufferers	13	Benefits of research in general	2
Patent rights	8	Advances science?	2
Other alternatives	7	New location	1
Employment of academic staff	6	Convenience of the move	1
Who gets the credit	5	Facilities in the new laboratory	1
Rights of the company to exploit the research results	3	Animal testing	-

Story 2 - pharmaceutical milk

	No issues	1 issue	2 issues	3 issues	4 issues	5 issues
L1	25%	5%	25%	30%	15%	0%
L3	0%	5%	14%	14%	14%	53%

L1 average = 2.1   L3 average = 4.0

Cost and benefits	30	Ethics of mixing genes/cloning	8
Animal Welfare	20	Environmental impact of cross-breeding	6
Risks to humans	19	Public opinion	5
Scientific viability	10	Advancement of science	2

Story 3 - modified plant virus

	No issues	1 issue	2 issues	3 issues	4 issues	5 issues
L1	13%	9%	4%	31%	17%	26%
L3	9%	3%	19%	6%	13%	50%

L1 average = 3.1   L3 average = 3.4

Is it worth it?	36	Animal welfare	6
Environmental risk	31	Issues of research safety	5
Food safety	30	Public perception	5
Developing country	25	Issues of genetic engineering	4
Cost and funding	21	Land use	1
		Commercial driving force	1
		Ethics of abandoning the technology	-

Appendix IX - Plant Virus - sample responses

Cost and Benefit/Is it worth it?			Ethics of mixing genes/cloning		
• Is it feasible to create an enhanced crop? • Benefit-cost? • Potential benefits? • Is it worth it? • Can you fund the entire experiment?	• What are the dietary benefits to the population? • Is it necessary to improve the crop? • How great is the dietary improvement?	• How large is the increased quality of crops vs. any harmful effects? • Is it necessary/justified to increase protein content? • What are the alternative methods for achieving the same effect?	• Is it ethical?	• Is genetic modification acceptable?	• We have to think where this can lead to?
Environmental risk			**Animal welfare**		
• Are there side-effects? • What are the effects on the environment?	• Will the crop interfere with organic farming? • Who knows about the long-term effects? • Can the virus become a disease when spreading to other plants?				
Risks to humans/ Food safety			**Public perception** • What is the public reaction		
• What are the risks of eating the plant? • Is it safe? • Any oncogenic effects?	• What is the level of certainty of side effects? • How is the safety tested?			• Will majority of people accept the final product?	
Developing countries			**Advancement of science** • Is it possible		
• Will it grow in areas that have difficulty supporting vegetation? • Use of plant in developing countries?	• Can developing countries afford this plant? • Will it have a beneficial dietary value in developing countries? • The willingness of developing countries to participate	• It is ethically nice to develop, but if developing countries cannot pay enough to recoup the R7D costs, who will subsidise? • GM food may not solve the problems that are rooted in poverty and lack of education.		• Safety of research protocols (to contain the virus in the laboratory) • How are the field trials organised to secure that the virus does not escape?	
Non-ethical statements (not scored): • Is it profitable? • What is the overall expense? • What is the mutation rate of the virus • Are the viruses going to be broken in the body?			• Are the plants tested the same as used for food? • Pathogenity of the virus described? • Would people buy the product? • Number of posts created? • Cost		

I RISKS			
IA HUMAN HEALTH	**IB ANIMAL**	**IC SUPERVISION**	**ID TESTING and LABELLING**
LEVEL 0: questions of risk for which an answer can be given on purely factual basis - i.e. no moral considerations required.			
• What is the physiological influence of the drug to CF patients? • Should proteins manufactured in an animal transfer to humans?	• how will the gene affect cow's original genes? • Where will the gene come from? • How will the milk affect the calves?	• Is it legal? • Where will the cows be kept?	• Whether the taste of the milk will alter. • How will the milk be purified?
LEVEL 1: First brief recognition of risk, which might serve as a stepping stone for higher level considerations.			
• Any threat to humans? • Will it be safe for humans to drink this milk? • Will anyone get hurt?	• How many animals are involved? • Is nuclear transfer safe? • Side-effects on the cow? • The effects of inter-breeding between normal cow and engineered one?	• Chance of modified cow interbreeding with normal cows • Would the cows be living in a normal farm or would they be segregated to sterile environment?	• Is the product tested? • How will you distinguish normal/altered milk?
LEVEL 2: Better understanding of risks, the considerations are more factual than moral, though moral elements are now present. Reference to long-term safety and harm characterise the human health risks. Emergence of concern for control and supervision also characterise the level, but there is yet no concern about who are the decision-makers and how do we balance risks and benefits. Responses also sometimes include strong, but unqualified, value-statements.			
• Possibility of more harm if cow diseases transfer to humans? • Will it get into human food chain? • Long-term effects - what if it causes death to the patient a few years later?	• Will the cow suffer from producing the milk? • is the quality of cow's life adversely affected? • Animals should not be subjected to any pain or distress? • What are the long-term effects of creating a transgenic animal?	• Will there be controls? • What are the research protocols to guarantee non-breeding between engineered and normal cows? • Can we monitor the welfare of the animals at all times? • There should be no way this gene could enter the wider population.	• Does the product require animal testing? • How can the product be tested for human consumption without any risk to (young) subjects? • How will the product be tested?
LEVEL 3: The responses now include serious considerations about the role of decision-makers and what should influence the acceptance of different levels of risk. Justification for risk in using animals is explicitly sought.			
• How big of a risk to adverse health implications are we ready to accept when this product is used to treat CF patients?	• How much animal suffering can we justify for commercial profit?	• Who should supervise the project- it should be an outside body? • How could the research results be abused by others?	• Is there an opportunity to refuse to drink the modified milk - labelling is necessary!

II COST AND BENEFIT + RESEARCH ISSUES		
IIA MEDICAL BENEFITS	IIB OPPORTUNITY COST and RESEARCH	IIC COMMERCIAL INVOLVEMENT AND ACCESS
LEVEL O: Considerations of cost and benefit that are either purely commercial or otherwise answerable without any moral considerations and questions about the scientific procedure answerable with scientific knowledge		
• How much milk do the CF sufferers need to drink? • What is the reason to think this protein is therapeutic? • How many CF sufferers are allergic to dairy products?	• Cost / How expensive is the research going to be? • Commercial viability • /Why have previous research failed?	• Is there already a patent for CF cow? • How large is the market for new CF drug? • How quickly could this become a commercial product?
LEVEL 1: Introduction of some cost-benefit considerations, though still answerable by factual information, some moral considerations could be included as well		
• Will it benefit CF sufferers? • Will it greatly relieve suffering? • How many people will benefit?	• cost-benefit / Available resources / Alternative methods • Has enough research been done? • Effectiveness of technique? • Chances of success?	• Is there enough money to go ahead? • Funding
LEVEL 2: More profound considerations of benefits in terms of improvement of quality of life for CF sufferers and animals or comparison of costs and benefits by asking 'Is it worth it?' Any consideration of the whether the product should be commercial or access is likely to be at least level 2.		
• There is a need to reduce human suffering! • Will the drug improve the quality of life of CF sufferers? • How long will it take for patients to benefit?	• Should funds be used to develop the original technique or the new one? • How hopeful are researchers. Is it one in a million chance? Is it worth it? • Is it a must? • Do benefits outweigh harms? • Would there be other benefits?	• Who will pay for it? • How expensive/ affordable will the treatment be? • copyright • How would the cost/distribution be handled? • What is risked for commercial gain? • Who will benefit the most? • Is economics the only reason cows are used? • Is it for human or commercial gain?
LEVEL 3: The responses now seek justification of costs in comparison to benefits, are concerned with the role of commercial companies in relation to copyright or access to the drug and whether use of resources for this research are justifiable in light of opportunity costs (i.e. understanding of resource scarcity)		
• Will benefits to patients be worthwhile enough to justify altering the genetic composition of a cow? • Could it raise false hopes of cure? • Who decides which human disease is most debilitating and therefore deserves attention and investment?	• Are funds being transferred from other beneficial causes? • Is the research justifiable in terms of time& money& sacrifice? • Is this the path that will benefit the CF patients most?	• Should it be a commercial product? • Should a company be allowed to copyright life-saving treatments? • Is there any point doing this if people cannot afford the treatment in the end? Is it for human or commercial gain

III ISSUES OF BASIC VALUES	
IIIA GENETIC RESEARCH	IIIB ANIMAL RIGHTS
LEVEL 0: All considerations of basic values include moral elements, so there are no level 0 responses	
LEVEL 1: Responses indicate distinct recognition that there are moral issues, but offer no further considerations.	
• Messing with genes. • I don't like the idea of cloning	• Is the new gene from an ethical source?
LEVEL 2: Responses elaborate on the level 1 type responses, but do not involve issues of justification.	
• Is it right to play God and mess around with genes? • We have to consider the whole Gm thing • Moral issues regarding the creation of transgenic animals • Consequences for further developments • Does this group have good human rights/other rights record? • Ethics of cloning and use of cloned materials for treatment of humans.	• Should we clone animals? • Animal rights • Is it ethical to use a protein from another animal to treat human disease?
LEVEL 3: Responses show mature understanding of the questions of genetic engineering and seek answers that would justify choices, namely control, should it be limited and what type of animal experimentation is in general acceptable.	
• Further growth of cloning idea - is this acceptable? • Who will control how far this goes?	• Ethics and rights to use animals for human use • Do we have a right to create a lifeform? • Is it fair to use animals in research that dangers their health? • Is it better that human quality of life is improved, when cows are damaged? • What value is given to suffering and life of cows?

IV PUBLIC OPINION
LEVEL 0: All public opinion entries contain at least a possibility of moral concerns, and therefore no level 0 items are scored
LEVEL 1: Basic recognition that public opinion exists
• Political opinion • Public opinion
LEVEL 2: Recognition of public opinion as something that may influence research decision-making. Also recognition that the opinion of some groups may be more important than others.
• How the CF sufferers and their families feel about this? • Public opinion (after education) • How will this finding affect society? • It would probably cause a lot of negative media attention after Dolly the sheep. Has this been considered? • Opposition from public/environmentalists • Should we ask a larger group what they think? • Will the public be informed or will this be kept hush-hush?
LEVEL 3: An understanding that public opinion is partly dependent on information it is given and that the public is allowed to oppose even beneficial treatment on moral grounds.
• Whether or not the use of genetic engineering in the process will be accepted by the public/CF sufferers? • To make sure the sufferers know that the drug was produced via transgenic animal.

Appendix XI - Group Personalities

Rabow *et al.* (1994) list seven non-functional group roles and provide some tips on how to reduce their negative effect on group dynamics: similar problem types can be also found in Westberg and Jason (1996) and the following list comprises their suggestions for remedying the problems:

1. The silent one: to be a true member a person must listen and verbally participate. This is not to say that one cannot talk less than another, but that participation is mandatory for full group membership. Remedy: try to identify the reasons for a group member being silent, whether it is lack of knowledge or confusion when it can be remedied with short re-caps on the subject, or slowness to express oneself when the group can grow more comfortable with some quiet moments for all to gather their thoughts, or shyness when a leader can encourage the shy person to speak when they can detect eagerness and support their comments.

2. The over-participant: talking a lot is not necessarily over-participating, but talk that does not help the group move towards its goal (often involves use of detrimental statement styles), that leads to irrelevant directions, that takes too long in relation to the value of the idea to the group, or which makes others very uncomfortable should be considered counterproductive. Remedy: the leader or other members of the group should recognise the valid points and then invite other opinions to assist the discussion. It is important to listen to the comments actively and non-judgementally, just as it is important to have the courage to stop people kindly when it is to the benefit of the group.

3. The wanderer: a person who has an idea, but who has not thought it through and thus rambles along in trying to formulate the thought. Remedy: the person could be asked leading questions that should help to clarify their thoughts - 'Are you saying that ....?'

4. The tangent person: goes off the point with irrelevant associations that are only vaguely related to the discussion topic. Remedy: A question to follow immediately to bring the discussion back onto the subject by either asking the person to clarify their vision on how their comment relates to the discussion or not to comment at all and swiftly return to the subject.

5. The storyteller: stories, personal anecdotes and experiences can be valuable for a group, but when the stories get too long or there are too many of them, the time used is no longer beneficial. Remedy: clear and positive comment to draw attention to time restrictions 'I wish we had the time, but...'.

6. The insecure talker: a person who often interjects with seemingly irrelevant comments may have feelings of insecurity in the group. Remedy: praise the contribution of an insecure talker and ask them to prepare something special for the next meeting to reduce their need to get their place in the sun by irrelevant comments.

7. The lone dissenter: an obstinate person expressing a minority voice in a forceful and stubborn manner. The voice of dissent need not be a problem, if the points can be accepted with respect and they help the group to appreciate a different point of view, but when the dissenting comments become dominant and increasingly frequent, they produce a hindrance for the group's other learning goals. Remedy: Comments like 'that really sets us thinking about this

in a new way' show appreciation and the dissenter may feel that his/her cause is acknowledged and their need to interrupt is thus reduced. Also asking the group member to clarify their point of view with supporting reasons may prove beneficial.

As important as trying to reduce the effects of problematic group roles, the facilitator needs to recognise and support beneficial ones. Ideally the facilitator should adopt these roles as necessary for a smooth functioning of the group and when they seem to be lacking among the group members. Rabow *et al.* (1994) describe seven positive group roles:

1. Encourager who praises, agrees, and accepts others' ideas
2. Harmoniser, who mediates and relieves tension
3. Compromiser, who comes half way, yields status and admits error
4. Expediter, who encourages and facilitates participation of others
5. Observer, who records group processes, feeds back to the group when needed.
6. Evaluator, who suggests new ways of looking at things (it might be better if we....) or re-directs the groups activities (we seem to bogged down here, it might be beneficial to try the next step and come back to the definitions if we need to.)
7. Follower, who actively listens and accepts decisions.

An alternative to these is Belbin's (1993) set of team roles. She has identified eight basic team roles, each serving a purpose in a well functioning team and each role having both negative and positive attributes. In a synergistic team individual strengths complement each other and individual weaknesses can be both tolerated and compensated for, provided there is someone else in the team with the relevant strength.

The eight team roles described by Belbin are:

1. The Chairman: has a strong sense of overall objectives. Is able to keep an open mind and values contributions form any source. Generally of average mental ability and creativity. Good at controlling and co-ordinating resources. Democratic and encourages participation, but willing to take responsibility for decisions. Sometimes seen as reserved and detached, the Chairman's ability to remain objective is valuable when directing the efforts and activities of others towards an overall goal or objective. Allowable weaknesses of the Chairman are: uncompetitiveness, unambitious, amateurish, lazy, and not forceful. A person with low anxiety levels and fairly extrovert personality.
2. The Shaper: has strong sense of drive and urgency, and an outgoing, sociable, and dynamic personality. Readiness to challenge ineffectiveness, complacency, self-deception and a general lack of progress. Prone to provocation, irritation and impatience. May be seen as a bully by some, and may sulk if not getting own way. Strong preference to lead 'from the front' with an inner need to control decisions and actions personally. Can be quite aggressive and wants to see own ideas implemented, and quickly. Hates rules and regulations. A 'natural' leader in some ways, and can command respect and generate enthusiasm and energy in others. Can be sceptical of others and yet be over-

sensitive to criticism of own ideas. A person with high anxiety levels and an extrovert personality.

3. The Plant: the Plant's name stems from an ability to scatter around lots of ideas (seeds), many of which may lead to success while many may not. The Plant is the source of a team's creativity, with a fertile and intelligent mind, with plenty of original ways of looking at things. The Plant is concerned with the challenge of the new and can be obsessive. Often seen as having a head in the clouds, can be unaware of the need for sensitivity towards others. The Plant may not have much time for protocol, or the 'proper way of doing things', nor be terribly concerned with the practical implications of own schemes. Tends to be self-sufficient and can be difficult and uncomfortable colleague. However, can also be caught up in a wave of enthusiasm and can be swept along with general team euphoria. There is a childlike element present in the Plant - likes to be flattered and does not like own ideas criticised. A person with s dominant yet introverted personality.

4. The Monitor-Evaluator: The Monitor-Evaluator is a highly intelligent team member whose principal team asset is an ability to process large amounts of information in an analytical, objective way. Possesses good judgement and shows hard-headed, shrewd approach to issues and ideas. Is cautious, has perspective and is highly critical of flawed thinking in others. Is the most likely person in the team to spot a fatal error in a scheme that everyone else has missed. The Monitor-Evaluator is, as a result, likely to be seen by the rest of the team as a 'wet blanket', dampening down enthusiasm. Tends to lack ability to inspire and motivate others, and is rarely the source of new ideas. Often seen as over-critical and negative, the role is nevertheless crucial to successful team outcomes. A person with low anxiety levels and introvert personality.

5. The Resource Investigator: is the team's ambassador in its dealings with the world outside the team. Has the capacity for making highly effective contact with people and for exploring anything new. Will respond to new situations as exciting challenges, but can also lose interest quite quickly if progress is slow or once the initial fascination has passed. Variety and people are the essential diet of the Resource Investigator, who also has the ability in turn to stimulate and motivate others. An extrovert with low anxiety levels.

6. The Company Worker: is a team member who, above all, will be able to foresee how the team's ideas and plans will work out in practice. The Company Worker tends to identify strongly with the organisation and has a knack of knowing what practical issues will need to be faced. The Company Worker is naturally conservative and has to be convinced that an idea is a good one not just because it is new, but because it is of genuine worth. Has real organising ability and plenty of common sense. Hard working and strongly self-disciplined but can sometimes lack flexibility. An essentially stable personality.

7. The Team Worker: is the fabric which helps to bind a team together. Promotes team spirit. Makes people laugh, is sensitive to other's feelings and to overall team mood. Is aware of the strengths and weaknesses of others and responds appropriately to people's differing needs. Can be indecisive in a crisis and may lack the necessary toughness in certain situations, but ability to 'read' others and to recognise own differing abilities promotes high morale and a good sense

of co-operation. For this reason is often a team leader. The Team Worker is an extrovert with low level of anxiety.

8. The Completer-Finisher: has a strong capacity for follow-through coupled with a striving for perfection which causes much anxiety. Will worry over small items, but, overall, accomplishes tasks well and on time. The nervous energy which is invested in the team's final product results in a high standard. Tends not to be a good leader - is fussy and can get bogged down in detail which may lower team morale. A introvert with high anxiety levels.

Appendix XII - L3 Structured Discussion Programme

Together with course co-ordinators the following ethical themes were chosen as the core of ethics education for each L3 Honours course:

Aquatic Bioscience and Zoology	Animals in scientific research Ecological decision-making Scientific misconduct and integrity
Biomedical Science	Animals in scientific research Scientific Integrity
Biotechnology and Botany	GMOs, publicity and scientific integrity Social implications of GMOs Scientific misconduct and integrity
Genetics and Molecular Biology	Animals in scientific research Embryo research / Genetic screening Scientific misconduct and integrity
Immunology	Animals in scientific research Scientific misconduct and integrity
Pharmacology	Animals in scientific research Drug testing (in the developing world) Scientific misconduct and integrity
Physiology	Animals in scientific research Ethical issues in pharmaceutical industry Scientific misconduct and integrity

Appendix XIII - PBL

Microbiology and Parasitology

Microbiology students participated in two ethics PBLs (one in each term), while parasitology students worked at the same time on scientific PBLs. Each PBL consisted of 5 contact hours and 10 hours of independent work.

The Ethics PBLs were designed to give students an opportunity to explore ethical issues in context with scientific decision-making. The two PBL exercises covered two areas of vaccine development: 1) the choice of research topics (TB or Meningitis B), and 2) the design of an ethically sound testing protocol involving both animals and humans. The 'problems' can be found in Appendix XXI - Microbiology PBL.

The PBLs were assessed by students completing learning logs Appendix XXIII - Introduction to Learning Logs.

Physiology and Sport Science

The 180 student L3 Physiology and Sport Science course was divided in two groups, one working on an ethics PBL and the other on a scientific PBL. The ethics group were involved in a problem dealing with drugs in sport: 1) investigating the reasons why they are forbidden, and 2) deciding on a suitable punishment in two hypothetical cases. The PBL consisted of 3 contact hours and 10 hours of independent work. Appendix XXII - Sport Science PBL.

The PBL was assessed in both groups by students completing learning logs. Appendix XXIII - Introduction to Learning Logs.

Appendix XIV - Animals in Scientific Research

Animal research and animal welfare - where should we draw the line?

In 1998 in Great Britain 2.66 million scientific procedures were performed on living animals. Of these 60% used mice and 22% rats. Fish and birds were each used in just under 5% of procedures, rabbits in 1.4%, dogs in less than 0.3%, monkeys in less than 0.2% and cats in about 0.05%. Fifty-two per cent of all procedures in 1998 were carried out for medical or veterinary research, drug development or pharmaceutical safety testing; thirty-four per cent as part of fundamental studies in the biomedical or biological sciences; and about six per cent for the safety testing of non-pharmaceutical products for worker, consumer or environmental protection. The total animal usage has declined steadily since the mid-1970s.

An increasing number of people do not accept this use of animals in the laboratory. Their reasons for opposing differ, as do the extent of animals rights they put forward. The choice of using animals in research is in many instances a legal requirement, or elementary to scientific research, but it is also a moral choice. To be able to formulate and defend your own view on animal testing, as well as respect the views of others, it is essential to understand the ethical arguments involved.

For this session on animal research and animal welfare you are asked to read two short papers (both attached) on ethical issues relating to animal research. The first paper, 'Why Ethics Matters', is a general discussion on the ethical theory involved in any moral choice and it will give you some tools to read the second, 'Animals in Scientific Research', which discusses the use of animals in research with a more focused attention.

Preparation for the session

- When reading the two articles, write down all the words that are unfamiliar, so that they can be defined in the class.
- For the second article prepare a short answers to the questions asked in the paper.
- Be prepared to discuss your comments and views during the session.

If you wish to learn more about animal welfare issues, the following are good Internet sites to get started with (the links can also be found on http://www.gla.ac.uk/Acad/IBLS/DEEB/ethics/link6.html (follow study packages and then Animal rights and animal welfare) - this will save you from typing all the URLs):

- **Animal Welfare ethics - a long list of links** http://www.ethics.ubc.ca/resources/animal/
- **Animal Rights Resource Site** http://arrs.envirolink.org
- **Animal Rights; Ethics** http://ethics.acusd.edu/animal.html
- **Fund for the Replacement of Animals in Medical Experiments (FRAME)** http://www.frame-uk.demon.co.uk/
- **Animals (Scientific procedures) Inspectorate** http://www.homeoffice.gov.uk/animact/aspileaf.htm
- **New Scientist web-site on animal experiments** http://www.newscientist.com/nsplus/insight/animalexperiments/animalexperiments.html
- **Interesting arctile 'Util-izing animals' by LaFollette and Shanks on utilitarian theory** http://www.etsu-tn.edu/philos/faculty/hugh/utilize.htm
- **Ministry of Agriculture, Food and Fishery web site on animal welfare** http://www.maff.gov.uk/animalh/welfare/default.htm

Henriikka Clarkeburn, 1999

Why Do Ethics Matter?

Henriikka Clarkeburn

Our lives as private persons, and responsibilities as professionals, require us to make moral decisions each day of our lives. In our private lives, we make decisions such as telling a white lie to save a friend from pain, or saving up to buy a large purchase rather than giving any money to charity. In our professional lives as scientists we need to make decisions on which problems to study, which methods to apply, and how to present our data. Very few of these questions have straight forward answers, and we are often unsure whether we have made the right choice.

Ethics is an academic discipline devoted to finding and evaluating solutions to ethical problems. The ethical principles found in these academic studies can give valuable guidelines to everybody, when making their personal and professional moral choices.

Occasionally, those working in life sciences are suspicious that 'soft' disciplines, such as moral philosophy, lack the type of academic rigor displayed in their own fields. Some people believe that ethical opinions are mere preferences akin to expressing a taste for a flavour of ice cream, or a type of music. However, very few philosophers would agree with such a strongly subjective view of ethics. We can make rational decisions about our ethical positions in a way we cannot about ice cream. If a friend expresses a preference for strawberry, we are not compelled to argue the merits of chocolate. This would not be the case, if friends expressed an intent to commit murder - then we would be compelled to persuade them to change their mind. However, ethics is not as strongly objective as many scientific principles are. Scientists around the world, or at any time throughout history, who seek to measure the density of pure gold will find, within the accuracy of the instruments, the same result. Yet there is no comparable experiment which we could perform to assess the morality of a cultural practice, such as polygamy, which is acceptable is come cultures and taboo in others. Ethics falls in between these two extremes. Moral positions are not a matter of taste or habit, nor immutable physical constants which can be objectively determined irrespective of time and culture.

While there is not always an agreement on the best solution to a moral problem, and not all philosophers advance identical ethical theories, this fact should not be attributed to any inherent weakness in the discipline. It is not all that uncommon that two biomedical scientists disagree on the implications of a particular data set. It is also quite common for two scientists to approach a problem with two different hypotheses in mind. Likewise, given an ethical dilemma, you can find ethicists who reach differing conclusions as to the best course of action. The difference of opinion may be attributable to the fact that each ethicist has tried to solve the dilemma by using a different theory, or alternatively, they have used the same theory, but attributed different weights to different aspects of the theory. In addition there might be disagreements over the empirical facts of the case (for example, whether animal feels pain, and how much, during a particular experimental procedure).

However, it is equally important to realise that while many ethical dilemmas do not have a 'right' answer, there are answers which are clearly wrong. Who would seriously suggest that moral choices should be made by tossing a coin, or that abortions are moral on Mondays and immoral on Tuesdays? Ethical positions can be evaluated and compared by using techniques which are not all that foreign to those used in science. Ethical theories are judged on the basis of their rationality, their consistency, and even their usefulness.

Ethical theories, like any other theories, are expected to be internally consistent. Similarly, theories which are unclear and incomplete are obviously less valuable than theories that do not suffer from these flaws. Simplicity is also an advantage, because all things being equal, it is preferable to employ a simple theory over one which is difficult and complex to apply. We should also require that an ethical theory would provide us with guidance in those moral problems where intuition, our gut feeling of right and wrong, fails to provide us with a clear answer. Most real-life moral dilemmas are considered problematic precisely because convincing arguments can be presented to support each side of the issue. These types of situations are where we most require the guidance of a moral theory.

Additionally, ethical theories should generally agree with our moral intuition. Who would wish to adopt an ethic, which, although consistent and logical, would support murder for profit? However, it is more difficult to decide about a moral theory which runs counter to our moral intuitions in an area less clear-cut than murder. How are we to decide whether it is the theory, or our intuition, that is out of line? One way is to subject the moral theory to several moral problems, to try it out in real-life moral dilemmas and if it provides good, intuitively

acceptable answers in many of the situations, than it might be that our intuition is the one in a need of re-consideration rather than our moral theory. While, if the moral theory fails to provide intuitively acceptable answers to the moral dilemmas which it is applied to, then we need to look for alternative theories or ways to improve the one we are using. Because no moral theory is perfect, we need to keep re-evaluating our adopted ethic time and time again, to make sure that our actions and opinions are coherent with the values we hold dear.

The ethical theories which give us a framework to make our moral choices are in general divided into two major categories: consequentalist theories, which concentrate on the moral consequences of actions (hence the name) and deontological, which look at the value of actions separately from their consequences. Thus to determine whether an act is moral or immoral, a consequentalist needs to evaluate whether the consequences of that act are good or bad, while a deontologist determines the morality of an action solely on the moral principles that were used in deciding on the action. Some of the common moral disagreements arise from people approaching the moral dilemma from either a consequentalist or deontological point of view. In order to make up your mind on a moral problem, it is to your advantage to understand how the different positions are constructed, to give you an opportunity to evaluate their worth. The most common of consequentalist theories is utilitarianism and of deontological theories, Kantian categorical imperative.

Utilitarianism

One cornerstone of how we treat other people (and animals?) is the thought that hurting unnecessarily is wrong, whether we hurt them physically or their feelings. We might supplement this by considering it morally valuable to try to increase happiness and reduce suffering, in general to improve the lives of those around us. Utilitarianism is a systematic philosophical way of capturing these ideas. The core of utilitarianism is to consider the interests of all, not just self. The best moral action is the one which maximises the general good and minimises the general evil.

For example, you may be considering lying about the results from an experiment you have set up for your project. Your motive for lying is to get better marks, which could lead to better chances of doing a post-graduate degree or a better job and salary after you have graduated. However, utilitarianism requires you to consider the impact of your decision to lie on other people. You must consider how presenting the false results may affect someone else's work, who is to build upon your work next year? What if the results lead to other experiments, which could have clinical outcomes and your false results could lead to direct harm to patients or animals? What if you get caught and this leads to a reduction in students' freedom to do their projects in subsequent years, as they are considered untrustworthy because of your lie? What if the employers find out that people complete their university degrees based on false data: they might lose their trust in the degree, which could lead to decreased employment opportunities for you and your class-mates?

If you consider all these effects of lying, not just the positive effects for you, it will become clear that the net outcome is a bad one. According to utilitarian theory, the act of deceit in this situation is not good and you ought not carry it out. But consider a different situation, where a mentally disturbed friend of one of your class-mates shows up in a lab you are working in, waving a scalpel screaming to kill your friend for 'ruining his life' and asks you to tell where your friend is at the moment. Although, you know where your friend is, should you tell the disturbed man in your lab? After performing a similar utilitarian calculus as above, the answer is mostly likely no. The net of good and bad consequences that flow out of this deceit are markedly different from the previous example. One of the core elements of utilitarianism is exactly this, you must evaluate each choice independently and decide its moral worth based on the case specific particulars.

But like all moral theories, utilitarianism is not without problems. First, we will often find it difficult to decide on whether the net balance of an action is good or bad. We may not have included all those involved; we may not have considered all aspects of benefits and harm that it may cause, and not all may agree on our evaluation on what outcomes are good and which ones are bad.

Second, utilitarianism can allow actions which are against our moral intuition. For example, torture of one person could be justified as long as it made enough people happier as a result.

Kant's categorical imperative

Kant, a German philosopher in the 18th century, had a problem with the changing moral status of the same action (e.g. lying), which is accepted in utilitarian theory. He thus formulated a moral theory which requires us to 'act only on that maxim through which you can at the same time will that it should become a universal law'. In other words, according to Kant, in order to determine if a particular act is moral, we must first ask ourselves if we could wish that the rule governing our action be made a universal law - that is, if we wish everyone to use the same rule to decide on a course of action. An action is thus immoral, If we cannot truthfully desire everyone else being permitted to perform the action that we are considering. To Kant, the uttermost importance is why we do what we do, what rules and theories we use in making our choices, and not necessarily what results from our action. So it is your intentions, not the consequences of your actions that make them moral or immoral.

Consider again lying about experiments you have performed. Before doing this the categorical imperative requires you to first ask yourself whether or not you can honestly wish that your deed becomes a universalisable rule. This rule would allow scientists to submit fraudulent data as genuine. Clearly such a rule would destroy the credibility of science and halt scientific progress considerably. No one could legitimately wish such a rule to be universalised and thus the action of lying about your project would be immoral as well. This rule would then apply to all situations of lying, making them immoral from the outset.

Many of our political rights are justified by reference to universalisable rules. For example, we do not wish to qualify basic human rights at each point of contention by reference to the consequences of each action. There seems to be something intuitively appealing about having some rights, which we feel are not going to be broken, that give us a baseline for interaction with others.

The difficulty with Kant's categorical imperative, as with all other deontological theories, is to find a balance on how widely, or narrowly, one defines universalisable laws. If we deduce from the above example that deceit, in all its forms, is immoral, then we are committed to tell the scalpel-waving lunatic where our friend is. On the other hand, if we believe that the universalisable law refers only to specific scientific conditions, we are faced with the laws which are not really universal, but particular, and thus the moral rules would change from situation to situation, which would give an opportunity to change the description of morality to suit personal aspirations. This would stop being moral laws altogether.

Conclusion

There are several valid ways of looking at moral problems. While both basic types of theories, utilitarian and deontological, have their fierce opponents and supporters, it seems that solving real moral problems frequently requires both types of considerations to be used before a solution can be reached. Understanding the basic methods of analysing and justifying moral decisions gives you tools to make, if not always better, than at least more conscious and mature moral decisions.

Animals In Scientific Research

by Henriikka Clarkeburn

In recent years much attention has been focused on the use of animals in research. Both as scientists and consumers you are faced with personal decisions regarding this issue. Whether you are for or against the use of animals in research, it is in your interest to understand the arguments used on both sides of the issue so that you can both form the most coherent view for yourself and defend it when questioned. This short paper maps out the main arguments: we hope that it will provoke you into thinking what your own views are.

The use of animals in research -debate has two paramount concerns: 1) that animals are being used for the wrong purposes i.e. questions of animal rights, and 2) that animals are mistreated or abused, whether or not the purposes are justifiable i.e. questions of animal welfare. The first type of concern can be linked with improper use of animals in general - for food, clothing, and experiments, while it can also be solely focused on the scientific purposes for laboratory animals - teaching, commercial testing, or medical research. The second concern can similarly span outside the scientific arena, considering the proper use of animals in farming, zoos and experiments alike. For reasons of consistency, it is often important to test the animal rights and welfare arguments presented for animal research by applying them to other types of animal use, but for the purpose of this paper, we will consider primarily how we define the correct use of animals in research.

ANIMAL RIGHTS APPROACH

The animal rights argument is often supported by asking whether there are any relevant criteria for differentiating between humans and non-human animals? If we cannot think of any such relevant criteria, than we must include animals in our moral concern equally with humans and give them at least the most basic moral rights - right to life as the most important one. In the absence of relevant criteria, if we still differentiate between humans and non-human animals our actions are analogous to racial and sexual bias, as we are distributing rights based on morally irrelevant criteria not unlike skin colour or gender. This is a deontological argument based on some inherent value which all animals, humans included, possess, and which grants them moral consideration.

Can you think of any such criteria that would logically describe all humans and humans only? Do you think that the criteria often suggested - rationality, autonomy, linguistic capacity - actually irrefutably differentiate between all humans and non-human animals?

This kind of argument is essentially negative. It can demonstrate the absence of a significant difference between humans and other animals. We can also form a positive argument to support animal rights. This claims that moral standing is derived from the ability to feel pleasure and pain, or to be sentient. As Peter Singer[2] puts it: "If a being suffers there can be no moral justification for refusing to take that suffering into consideration ... If a being is not capable of suffering, or of experiencing enjoyment and happiness, there is nothing to take into account." Any moral agent must consider the pain and pleasure that results from his or her actions. This is the minimum requirement of morality. The capacity for experiencing pain and pleasure is the primary moral similarity between humans and non-human animals. Sentience, then, is the non-arbitrary, non-speciesist basis of moral value. This argument is dominantly utilitarian and it is thus a consequentalist doctrine in which pain and pleasure are the main determinants of moral value.

This second kind of argument is also used by those concerned for animal welfare. The presumption is that the pain experienced by an animal is morally significant and the use of animals should take into consideration the pain and pleasures of animals involved. The animal

---

[2] Singer, Peter (1977). Animal Liberation: A New Ethics for Our Treatment of Animals, p.8. New York, Avon Books.

welfare concern often accepts the use of animals in research, but advocates for procedures that minimise pain and discomfort experienced by animals in a laboratory.

Why do you think avoidance of pain is a moral paradigm? If you imagined humans beings were used in research in the fashion non-human animals are, what moral problems would you see?

Both of these reasons for supporting animal rights and welfare, either in an attempt to stop the use of animals in research altogether or to increase the welfare requirements for laboratory animals, have been criticised and suggestions made of their fallibility as solid grounds for the case of animals.

The deontological concern for animal rights, i.e. that the denial of animal rights without a specific moral criterion is analogous to racism and sexism, has been claimed, for example by Katz[3], to rest on the marginal cases of humanity. These Katz defines to be the severely retarded, the insane, the comatose , newborns with severe birth defects, and fetuses. Katz suggests that it is empirically false to say that these 'marginal human beings' are treated as normal or typical humans from the moral point of view. If these 'marginal humans' are not treated with equal moral concern, Katz continues, than the entire speciesm argument collapses as there is no speciesist criterion used for different moral concerns, but rather morally relevant criteria of rationality, autonomy and the like. He goes on to say "the cases obviously differ, but all in all, these [marginal] humans are clearly deemed to have less moral value because of their reduced capacities... This is a factual moral truth, however depressing it might be, that the hierarchy of moral value exemplified in the human treatment of animals is echoed and repeated in the human treatment of other humans."

Are you convinced by Katz's refutation of the deontological argument for animal rights? If not why not? Can you think of any other ways of criticising the deontological argument?

The utilitarian argument for animal rights, which puts sentience in the centre of the moral argument is often found problematic in two ways. First, how far down the scale of animal life can one safely assume the experience of pain and pleasure? This question has two elements: first, which animals can experience pain? and second, how can we determine whether they are experiencing pain or pleasure? It is suggested that insects have a requisite nervous system[4], which would suggest that insects have a serious claim for moral consideration. Do you think this possibility suggests that the utilitarian basis for animal rights can be pushed too far, offering a *reductio ad absurdum* of the position i.e. extending the principle until we are faced with an absurd conclusion? Alternatively we can choose to include only higher animals in our moral concern and avoid the *reductio ad absurdum* situation. Both are problematic. If we include all animals in the realm of moral consideration we would find living difficult, as accidental squashing of a fly on the windscreen while driving on the motorway would become a morally reprehensible action. If we alternatively draw the line to include only vertebrates in our moral consideration, we are just shifting the speciesm line further down the scale, but setting it at an equally irrational point.

In your view is this just a theoretical quibble? Why would the issue of consistency matter in this situation? If you were going to draw a line, where would you draw it?

The second problem with the sentience criterion is the contextual significance of pain, as Katz phrases it. Katz suggests that the utilitarian argument contends that pain is an intrinsic evil. An inconsistency lies with the natural existence of pain in nature. Katz states that in its concrete natural existence pain has an instrumental function in organisms and if understood in context pain is not an evil at all, but it is an essential part of a successful organic life. Thus the abstract

---

[3] Katz, Eric (1997). Defending the use of animals by business. In Nature as a Subject: human obligation and natural community. pp. 79-90. Lanham: Rowman&Littlefield.
[4] Lockwood, Jeffrey A. (1988). Not to Harm a Fly: Our Ethical Obligations to Insects. Between Species 4:3:204-211.

denial of pain proposes a practically impossible denial of pain, which weakens the pain arguments significantly.

*Do you think this argument is convincing? When do you think pain is not evil?*

ANIMAL WELFARE APPROACH

The animal welfare approach to animal use in research is most often based on utilitarian arguments, marking the significance of animal suffering, but giving it an unequal weight with human suffering. This argument often by-passes the above discussed considerations, which is both its deficiency and benefit. The presumption of lesser value of animals weakens the theory considerably because it cannot be easily defended and thus leaves the theory open to attacks of inconsistency, among others. But the avoidance of deep ethical considerations is also a benefit, because it allows people to move to discuss the actual animal use situations rather than being stuck on theoretical considerations of animal rights and moral status.

The animal welfare approach is often an application of a pain/benefit calculus. The aim is to get the best possible benefits for minimal amount of pain. This involves both minimising the number of animals in laboratory and the pain they have to endure, and maximising the benefits expected from the results. In other words animal experiments would be morally acceptable only when we have reduced the animal pain to the minimum and when we can say our aims are morally significant. Most people would agree that cosmetics testing on animals does not qualify as a morally significant aim capable of justifying animal testing, while cancer research does.

Do you find this way of looking at animal research convincing? Where would you draw the line between morally justifiable research and morally unjustifiable research? Who in your view, should make that moral judgement?

Appendix XV - Ecological decision-making - DDT/Malaria
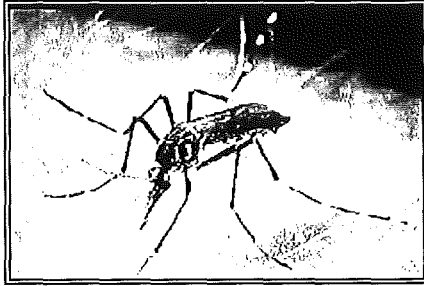
DDT and ecological decision-making

Decisions in ecology often require us to balance environmental harm with human welfare, and to consider costs of each alternative. These decisions are rarely easy. When they become complicated, it is important to use the best possible decision-making methods and tools to make sure that we make the best possible decision based on the knowledge we have at the time.

The use of DDT as a malaria control is a case in point. We are faced with the dilemma that DDT is an inexpensive and effective malaria control but that DDT is a pesticide seriously harming marine and bird-life.

Before our seminar, you should read the attached article by Curtis (1994) on the subject. While reading prepare a short summary of Curtis's key points and make notes which will help you to discuss the following issues during the seminar:
- Descriptions of any unknown words in the text
- What are the major elements to consider when deciding on a DDT ban?
- How do these elements of the DDT control issue apply to other ecological problems?
- How would you decide on this issue?
- How well do you think Curtis covered the subject?

If you wish to learn more about malaria, the following web-sites are a good place to start:
- http://www.iea.org.uk/env/malaria.htm
- http://www.who.int/inf-fs/en/fact094.html
- http://www.wehi.edu.au/MalDB-www/who.html
- http://www.malaria.org/

• CURTIS C.F. (1994). Should DDT continue to be recommended for malaria vector control? [review]. Medical and Veterinary Entomology 8:2:107-12.

Henriikka Clarkeburn, 1999

Appendix XVI- Genetic screening

Moral implications of genetic screening and the possible elimination of disease

Great hopes have been placed on genetic knowledge to provide a tool to create a disease free society, where all medical problems would be either removed or treated by genetic manipulation of one sort or another. The most realistic dreams for the genetic eradication of disease relate to inherited conditions resulting from a localised mutation which has direct and known impact on the carrier's health. An example of such conditions is cystic fibrosis. The dreams have extended further to cover diseases which result from a combination of genetic and environmental factors, including cancer and coronary heart disease.

There are at least two difficulties in accomplishing a disease free society by application of genetic knowledge:

-   It is not obvious how diseases could be eliminated by genetic manipulation alone.
-   An attempt to eradicate genetic diseases has social, ethical and financial implications, which may outweigh the benefits of the programme.

Before next weeks discussion you should read the attached article by Bruce Ponder (Science 278(5340) pp. 1050-1054), which focuses on genetic testing for cancer. After you have read the article prepare a short summary containing answers to the following questions (you will be asked to present your summary in the class):

-   What is the author's main message?
-   What are the major themes in an attempt to eradicate cancer by application of genetic information

Further, while reading, prepare notes that will allow you to discuss the following issues in the class:

-   List of terms or concepts you are not sure of.
-   What other social and ethical problems can you think of relating to genetic testing?
-   What is your personal view of trying to eradicate diseases by genetic testing and/or manipulation?

If you wish to find further information relating to genetic research relating to cancer, the following articles are a good place to start with. All of them can be found both on the internet (use Ovid Biomed search tool and 'Core Biomedical Collection' - you'll need an Athens password, which you can get from the Library Inquiries desk) and in the library's Life Sciences Periodicals:

-   DICKENSON D.L. (1999). Can children and young people consent to be tested for adult onset genetic disorders? British Medical Journal 318:7190:1063-5.
-   HABER D.A., FEARON E.R. (1998). The promise of cancer genetics [The Promise of Cancer Research and Treatment]. The Lancet 351:2S:1SII-2SII.
-   HOLTZMAN N.A., SHAPIRO D. (1998). The new genetics: Genetic testing and public policy. British Medical Journal 316:7134:852-6.
-   KODISH E., WIESNER G.L., MEHLMAN M., J.D., MURRAY T. (1998). Testing for Cancer Risk: How to Reconcile the Conflicts. JAMA 279:3:179-81.
-   LOW L., KING S., WILKIE T. (1998). Genetic discrimination in life insurance: empirical evidence from a cross sectional survey of genetic support groups in the United Kingdom. British Medical Journal 317:7173:1632-5.
-   WELCH H.G. (1998). Uncertainties in Genetic Testing for Chronic Disease. JAMA 280:17:1525-7.
-   WILCKE J.T. (1998). Personal paper: Late onset genetic disease: where ignorance is bliss, is it folly to inform relatives? British Medical Journal 317:7160:744-7.
-   WARNING OVER 'NAZI' GENETIC SCREENING: http://news.bbc.co.uk/hi/english/health/newsid_415000/415136.stm

Henriikka Clarkeburn, 1999

Appendix XVII- Embryo research

The benefits and hazards of embryo research
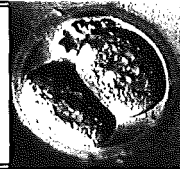
Human embryo research promises to find ways to improve infertility treatment, to increase our knowledge, and possibly ability to cure, inherited diseases, and an opportunity to learn more about human development. Despite the potential in embryo research, it is not allowed in many countries (France forbids it completely and US will not fund it from public resources). Several ethical issues have led to the controversy over embryo research, for example:

- Is an embryo a human being and should it thus be entitled to the same respect as people in general, which would mean that destruction of embryos is equivalent to murder?
- Is there something inherently valuable in the potential of developing into a person, that requires us to pay special attention and respect to embryos, though not the same as already existing people?
- Will embryo research lead to technology that may not be acceptable; human cloning, genetic engineering, and pre-implantation diagnosis?

This seminar is your opportunity to learn more about these issues, to discuss them and develop your personal view on embryo research.

Before the session you should read 'Pre-embryo Research: Medical Aspects and Ethical Considerations' by Eisenberg and Schenker (copy attached). While reading, you should make notes that will prepare you to answer the following questions:

- Were there any terms or concepts you were not sure that you understood correctly?
- What are the main ethical points made by Eisenberg and Schenker?
- What in your mind are the major issues in embryo research?
- What other moral and medical problems do you think relate to embryo research?
- What is your personal opinion about embryo research?

To learn more about the issues of embryo research, you may want to start with the following articles:

ANDREWS L., ELSTER N. (1998). International regulation of human embryo research - Embryo research in US. Human Reproduction 13:1:1-4.

ANNAS G.J., CAPLAN A., ELIAS S. (1996). The politics of human-embryo research - Avoiding ethical gridlock. New England Journal of Medicine 334:20:1329-32.

BERNAT E. (1999). The human embryo as object for scientists: Legal-ethical and legal-comparative reflections. Journal of Fertility and Reproduction 9:1:7-21.

CONCAR D. (1996). Into the mind unborn. New Scientist 19 October:40-5.

HUNT G. (1999). Abortion: Why bioethics can have no answer - a personal perspective. Nursing Ethics 6:1:47-57.

SCHENKER J.G. (1998). International regulation of human embryo research - FIGO statements and world experience. Human Reproduction 13:8:2047-9.

WATT H. (1996). Potential and the early human. Journal of Medical Ethics 22:4:222-6.

Henriikka Clarkeburn, 1999

Appendix XVIII - GM crops

Genetically modified crops: Ethical and social issues

"GMO's, or genetically modified organisms, have hardly been out of the news during the last couple of years. And with each new headline has come another round of debate over whether biotechnology is the way forward. Many scientists and politicians are confident that genetic engineering will produce wonder plants that resist pests and diseases, flourish despite drought and capture their own nutrients. Indeed, some consider biotechnology to be the only way in which to ensure that everyone in the world has enough to eat. Many others, however, are not so sure. Those that oppose genetically-engineered material do so for a wide variety of reasons, some are scientific, others concern power, choice and control, while many simply object on moral or ethical grounds. There is also deep rooted concern over who will have access to this technology and benefit from it." (New Agriculturist On-Line)

This and the following seminar are your opportunity to learn more about the social and ethical issues of GM crops, and as a result you should have more tools to both develop your own opinion about them and to understand the debate in the media.

The seminars are structured around The Nuffield Council on Bioethics Report on GM food, which is one of the most comprehensive of its kind produced in UK. You will find a copy of the report's introduction attached, and you may read and down-load the entire report from <http://www.nuffield.org/bioethics/publication/modifiedcrops/index.html>. You will also find three case studies in this package.

Before the first session on 11th October you should read the report introduction up to paragraph 1.20 and the case studies 'GM pollen warning' and 'US to label GM foods'. While reading you should make notes, which will allow you to discuss the following questions during the seminar:

- Were there any unfamiliar words in the text?
- What are the five ethical principles relevant to the evaluation of GM technology?
- How do you feel the 'precautionary principle' should be interpreted in the case of 'GM pollen warning'?
- What would be your approach to protecting both consumer and producer rights over labelling GM products (see case study 'US to label GM foods')

For the second seminar on the 1st November, you should read the rest of the Nuffield Report, in particular the paragraphs 1.20-1.31, and the case study 'Food for All'. Some of the questions to consider before the session are:

- How is justice important in GM technology?
- How do you understand the division of benefits and harms in introducing GM crops?
- How would you start solving the justice problems between developing and developed countries regarding GM crops?
- How convincing do you find the arguments of GM being 'unnatural'?

<div align="right">

Henriikka Clarkeburn, 1999
(H.Clarkeburn@bio.gla.ac.uk)

</div>

GM pollen warning
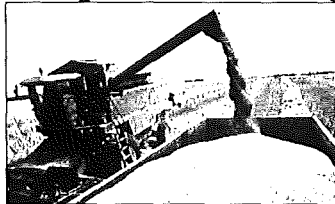
Controls to protect crops from cross-pollination by genetically-modified (GM) plants may be seriously inadequate, new research claims. Dr Jean Emberlin, Director of the National Pollen Research Unit, has produced evidence to show that pollen from maize an be dispersed over much greater distances than has been accepted by government scientists.

At present, a 200-metre "exclusion zone" is set up around a GM maize crop undergoing trials and is considered a sufficient barrier to prevent cross-contamination of ordinary maize crops or sweet corn. But, Dr Emberlin, whose research was commissioned by the Soil Association, says bees or strong winds will take the pollen such further.

She believes the government should now stop the controversial large-scale cultivation of GM crops, which is planned to start in a matter of weeks. Environment Minister Michael Meacher told BBC Radio that he accepted there could be a contamination risk of around 1% at 200m under moderate speed wind conditions. But he said the 200-barrier should be sufficient to ensure the purity of nearby crops.

It was based on many years of research and recognised as adequate by the European Commission, the Origination for Economic Co-operation and Development, and the US authorities.

"If it is necessary and right to revise the criteria by which we assess the likelihood of cross-pollination, we will do so," he said. Once again, he stressed the government's line on GM crops: "We are not going to allow commercial planting of GM crops until we are sure that we have the evidence to guarantee that there will be no damage to the environment - or indeed to human health."

Political pressure
Concerns over genetically-modified food in the last month have put the government under pressure to halt development until further tests are carried out.

The Soil Association, which promotes organic food and is opposed to genetic engineering, asked Dr Emberlin to undertake this latest research after the government's refusal last summer to order the destruction of a GM maize crop bordering an organic farm in Devon.

The government's Advisory Committee on Releases to the Environment (Acre) maintained there was little or no risk of cross-pollination.

But Craig Samms, of the Soil Association, says the new report does indicate a real risk to farmers growing GM free crops as well as the consumers who believe they are buying organic produce.

"Nobody knows what the risk is to people because this technology has been introduced without any of the normal checks you would expect.

"Genetic engineering in medicine is subjected to strict containment rules," Mr Samms added, "But here you have a technology that is just being planted in the countryside with no real idea of what the implications are".

'Disingenuous'

"The lack of acknowledgement of potential pollen spread concerns me," Dr Emberlin said. "Once the pollen is out there it is very difficult to redress the situation. I don't think it would be wise to go ahead with large-scale planting of GM crops without knowing more about the possible repercussions."

A statement from the Department of Environment released before Mr Meacher's radio interview rejected as "disingenuous" the suggestion that bees are a major factor in maize pollination in the UK.

"The issue of bees carrying maize pollen is a smokescreen to cast doubt on the competence and quality of Acre's advice," the statement read.

Liberal Democrat food spokesman Paul Tyler and environment spokesman Norman Baker welcomed the report.

"This report gives the lie to the dismissive attitude of both Conservative and Labour ministers. It is now clear that the risk is far greater than they have told us," they said in a joint statement.

Pete Riley, senior food campaigner at Friends of the Earth, said: "This new report highlights once again the appalling advice that the government has received on GM crops."

The report is published two weeks after a US biotechnology company, Monsanto, was fined £17,000 for breaking GM crop test site safety rules. The firm failed to maintain a six metre-wide barrier around a plot of genetically-modified oil seed rape in Lincolnshire.

Jean Emberlin's report can be found at:
http://www.soilassociation.org/SA/SAWeb.nsf/?Open (follow links 'library' and 'research papers')

US to label GM foods

US farmers have taken to GM crops in a big way. The US Government has said that it will probably agree to label genetically-modified (GM) food. At the moment, American law does not require this. However, European governments have threatened to continue their ban on the import of certain US GM products if the Americans do not accept such labelling.

Up to now, the Clinton administration has opposed GM labelling, agreeing with the American food industry that it unfairly stigmatises what they regard as perfectly safe products.

But the US Agriculture Secretary, Dan Glickman, said that relations with Europe on the GM issue could deteriorate into an all-out trade war and labelling was a way in which such a crisis might be avoided.

Speaking before an audience of environmentalists, lobbyists and lawmakers, Mr Glickman extolled the benefits of biotechnology. He said the technology would lead to increased yields and a decrease in the use of pesticides.

According to Mr Glickman, several European countries were letting their fears override these potential benefits. And he urged them to sort out their internal differences as soon as possible.

"Quite frankly, the food safety and regulatory regimes in Europe are so split, and so divided amongst the different countries, that I am extremely concerned that failure to work out these bio-tech issues in a sensible way could do deep damage in our next trade round, and affects both agriculture and non-agricultural issues," he said.

"Both sides of the Atlantic must tone down rhetoric, roll up their sleeves, and work towards conflict resolution, based on open trade, sound science, and consumer involvement. And I think this can be done if the will is there."

American farmers are producing more and more genetically-engineered products, with 44% of American soya beans and 36% of corn coming from GM seeds.

FOOD FOR ALL

Crops that resist drought and disease promise to transform the lives of poor farmers -if only they could afford them

Five million Brazilians faced starvation this year. This time it was a drought related to El Niño that halved grain crops in the north-east of the country, but next year it will be something else. Famine is perennial in Brazil.

In September Monsanto, the world's largest supplier of genetically modified seeds, announced it would invest $550 million in Brazil to build a factory producing its herbicide Roundup. Shortly afterwards the Brazilian government made Monsanto's Roundup-resistant soya beans the country's first legally approved, genetically engineered crop. The soya beans will boost profits for the big landowners who grow them to feed beef cattle for export. But most rural Brazilians are subsistence farmers who do not grow soya. No help will trickle down from Monsanto's beans to the starving millions.

The story exemplifies the limited contribution genetically modified crops have made so far to eradicating world hunger. It is not that biotech companies are uninterested in the developing world. Far from it: Brazil and other newly industrialising countries are in fact prime targets, with their growing demand for agricultural products, little opposition to biotechnology, and farmers who have risen above hard graft subsistence, but have not yet become customers of the world's seed and agrochemicals conglomerates.

But who will benefit from genetically modified crops in these countries? The companies speak of feeding the starving millions, while conserving the environment. They say that the new technology will have greater benefits in the Third World than anywhere else. "Biotechnology is a key factor in the fight against famine," claims the literature from EuropaBio, the association of European biotechnology companies. "Biotechnology will help increase the yield on limited land." Critics maintain that there is little evidence of this. Instead, they say most of the engineered crops developed or in the pipeline will benefit rich farmers, not the needy. Worse still, they fear the biotech industry's increasing domination of crop research will hurt, not help, the poor.

Agriculture does need a new technological saviour. Most of the world's food calories come from grain. A simple redistribution of what we grow now, even if it were possible, will not feed the 10 billion humans expected by 2030. Traditional methods of improving crops seem to have gone about as far as they can. "The fact that we start from the results of more than 5000 years of selective breeding makes further staggering yield increases unlikely," says Lloyd Evans of the CSIRO Division of Plant Industry in Canberra, Australia. "The biggest opportunity for increasing grain yields is to produce varieties more precisely adapted to local conditions."

Yet few of these crops have emerged so far. Those that are on or near the market aim to increase farmers' profits by cutting expensive "inputs", such as pesticides. This is little help to farmers who can afford no inputs to begin with, not even the reduced levels needed for these crops, and no help if they cannot afford the patented seed. Steven Briggs, head of the Novartis Agricultural Discovery Institute in San Diego, which sequences plant genomes, points to several innovations in the pipeline which might help: fodder crops that contain more calories, so more meat can be produced per hectare of corn or soya; crops that destroy toxins produced by moulds, such as fumonisin, which

cause massive crop losses after harvest; and disease-resistant crops, such as sweet potatoes and cassava, staples of the poor, which fend off viruses.

Crops that thrive despite drought and salty soils could also let farmers expand production into marginal lands. And the nutritional content of staples could be improved. If maize, for example, can be made to produce more of the amino acids it naturally lacks, the 80 million people who live almost exclusively on maize would get more protein. Ganesh Kishore, head of nutrition at Monsanto, says: "We can make it into a complete balanced meal."

Briggs agrees that there are contradictions inherent in bringing high-tech remedies to low-tech farmers. Breeding crops for subsistence, he says, is "emergency aid, not a path to economic growth". Pol Bamelis, from the German giant Bayer and chair of the German and European biotechnology associations, says that the industry "cannot help the fact that there are rich and poor in the world".

Buy out

Biotech companies think genetic engineering will be in the best position to help once farmers everywhere switch from small-scale subsistence to large- scale mechanisation. But many activists fear just that process. The high price of the technology could allow the few farmers who can afford it to out-compete their poorer neighbours and eventually buy them out, driving people from the land, says Hope Shand, of the Rural Advancement Fund International in Canada.

Monsanto also argues that helping poor farmers would reap another kind of benefit: richer peasants who no longer need to destroy forests to get more land. But this could be simplistic. Steve Vosti, of the International Food Policy Research Institute in Washington DC, has studied poor farmers and deforestation in Amazonia. He says any technology that increases a farmer's profits, or reduces the labour needed per hectare, will cause the farmer to cut down trees to get more land. It is not clear whether the kind of farmer who needs to fell forests to get land, or who eats little but maize meal, will ever be able to afford genetically modified crops. But even if only rich farmers benefit, says Vosti, their expansion would tend to push poorer farmers into forest margins.

And there are other disadvantages for the poorest farmers. "New biotechnologies threaten to aggravate problems of genetic uniformity, and increase the dependence of farmers on transnational corporations," says Shand. Even in the industrialised world, people are worried about genetic uniformity arising from the widespread introduction of genetically modified crops. In Missouri this summer, half the soya plants on some farms died of Fusarium mould, after three-quarters of the land was planted with Roundup-resistant varieties which turned out not to resist mould.

The handful of modified varieties offered by biotech companies will inevitably be more genetically uniform, hence more susceptible to unforeseen stress, than the plethora of classically bred varieties grown now. That problem could be worse in the tropics, where there is more existing crop diversity together with stresses that seed breeders based in the North may not have anticipated. Tropical countries will also have less money to pay multinationals for the rights to incorporate proprietary genes into several local varieties.

The last problem stems from the big companies' growing control of both markets and plant genes. Crop scientists must continually breed new crop varieties to meet the ever-evolving threats of pests and disease. In the Third World, this is mainly done by government-funded institutions, and the Consultative Group for International Agricultural Research. But public sector breeders are losing funding, while companies

such as Monsanto are rapidly becoming the only source for improved varieties. It already, for example, sells half the maize seed in Argentina.

Losing access

The public breeders are also losing access to plant genes. Last May the CGIAR completed a detailed study of the problems posed by the fact that the genes it needs to do its work are increasingly available only at a price, because companies hold the patents. India recently declined to pay Monsanto $8 million for the use, by its state-owned crop laboratories, of Monsanto's Bt insecticide gene. Those labs will not be able to provide Indian farmers with cheap, locally bred insect-resistant crops. Farmers who can afford to will have to buy whatever Monsanto has to offer.

Even if Third World breeders get access to patented genes, they may be forced to protect them in ways that put them out of reach of the poor. Terminator, a gene owned by Monsanto, keeps a plant from producing viable seed. So farmers cannot save seed planted, genetically modified varieties for the next harvest. It also keeps farmers from crossing patented strains with other crops to create new varieties. "Public sector breeders could be under great pressure to use Terminator to protect patented genes in the breeds they produce, in exchange for access to those genes," says Shand.

The overall effect could be that breeders will not be able to create new varieties to meet evolving threats unless they pay for the genes, and couple them with technologies to prevent the saving of seed. That means fewer, more expensive varieties, plus increased costs for the 1·4 billion poorest farmers who grow 80 per cent of subsistence crops from saved seed. As big northern companies expand their control of crop genes, their choice may be to buy seed, or die.
Debbie Mack

Appendix XIX - Clinical trials in developing countries

Clinical Trials in Developing Countries

One of the great challenges of medical research is to conduct clinical trials in developing countries, for the benefit of the population in those countries. Many features of developing countries - poverty, endemic disease, low level of investment in health care systems, and culturally distinct conceptions of health and disease - affect both the ease of performing clinical trials in these countries, and the selection of trials that would benefit the population in these countries (Varmus and Satcher, 1997).

Three basic principles need to be considered in the design of a clinical trial to be carried out in a developing country: the balance of benefits (an acceptable balance of burdens and benefits of research to the research subjects), justice (equitable access to clinical trials), and autonomy (guaranteeing that research subjects are in a position to give informed consent, and do so).

This seminar is designed to give you, by studying examples of ethical and unethical research set-ups, an opportunity to investigate some of the research standards accepted by the biomedical profession, and to develop your abilities to judge and design clinical trials according to ethical standards.

Before the session you should read the attached paper by Lurie and Wolfe (1997) and the Declaration of Helsinki. You should write a short, one paragraph, abstract which describes the main points of the Lurie and Wolfe paper. You should also make notes, which will allow you to discuss the following questions during the seminar:

- Meaning of any unfamiliar words in the text
- What are the major themes in designing a clinical trial?
- How these themes relate to research in both developed and developing countries?
- What is your personal view on clinical trials in developing countries?

The following papers relate directly to the subject and will help you to understand the issues better. Make use of them!

ANGELL M. (1988). Ethical imperialism? Ethics in international collaborative clinical research. New England Journal of Medicine 319:1081-3.

ANGELL M. (1997). The ethics of clinical research in the third world [Editorial]. The New England Journal of Medicine 337:12:847-9.

ANNAS G.J., GRODIN M.A. (1998). Human rights and maternal-fetal HIV transmission prevention trials in Africa. American Journal of Public Health 88:4:560-3.

BARRY M., MOLYNEUX M. (1992). Ethical dilemmas in malaria drug and vaccine trials: a bioethical perspective. Journal of Medical Ethics 18:4:189-92.

FADEN R., KASS N. (1998). Editorial: HIV Research, Ethics, and the Developing World. American Journal of Public Health 88:4:548-50.

VARMUS H., SATCHER D. (1997). Ethical complexities of conducting research in developing countries. New England Journal of Medicine 337:14:1003-5.

Henriikka Clarkeburn, 1999

CLINICAL TRIALS FOR A MALARIA VACCINE - Student hand-out

This is a brief trial protocol proposal based on research on a malaria vaccine. The novel vaccine is based on T-cell activity when malaria is in liver stage of infection. This vaccine protects the vaccinated individual against new malaria vectors.

The research is funded my the European Commission, biomedical framework.

Research protocol:

Placebo-control trial to be carried out in Botswana. Subjects recruited from rural villages where malaria is known to be prevalent. All recruited subjects are >15 years of age and seemingly healthy. The research is carried out by a European research team over two years.

Steps in Botswana:

1. Screening for malaria in the subjects, only healthy individuals are included in the trial.
2. Subjects are randomly divided into treatment with the novel vaccine and non-treatment groups.
3. The research subjects are screened for malaria every three months. Those with malaria are excluded from the trial.
4. The vaccine is re-administered after one year to those in the treatment group. The control group will receive placebo vaccine.
5. The research is considered successful if malaria incidence is reduced by 50% in the test group.
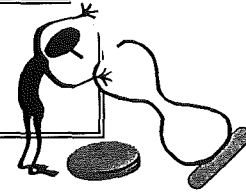

Your task:

Comment on the ethical and scientific issues of this trial.

Prepare an improved research protocol, if you found faults in the proposed one.

Appendix XX- Scientific misconduct and integrity

Scientific Misconduct and Integrity

The scientific enterprise is built on the foundation of trust. Scientists trust that the results reported by others are valid. Society trusts that the results reflect an honest attempt by the scientists to describe the world accurately and without bias. The level of trust that has characterised the relationship between science and society will endure only if the scientific community is able, and willing, to exemplify and transmit the values associated with the ethical conduct of science. The values of science, values of scientists, and the values of consumers of science need to interact in order to maintain the trust that allows science to flourish.

Ethical issues in science are all around us: is animal experimentation acceptable; who should get the credit for joint research efforts; should we try to alter the human genome; and when does commercial funding imply a conflict of interests, or incompatibility of values. Finding acceptable answers to these questions requires understanding of scientific, legal, economical, and ethical issues. This seminar is designed to give you some basic tools for dealing with the ethical elements of these, and other, questions in science.

Before the session, you should read the attached article: 'Scientific misconduct' and make notes that will prepare you for discussion of the following:

- Any words or concepts in the article you were not sure of.
- For the case studies prepare short answers to the questions.
- Think of a situation where you have been, or could imagine being, tempted by scientific misconduct.
- Think of a research proposal that you would have to turn down because it is incompatible with your personal values.

For further reading, you may want to start with the following:

BUZZELLI D.E. (1993). A definition of misconduct in science: a view from NSF. Science 259:584-585-647-648.

CHANTLER C., CHANTLER S. (1998). Dealing with research misconduct in the United Kingdom. Deception: difficulties and initiatives. British Medical Journal 316:7146:1731-2.

GOODSTEIN D. (1992). What do we mean when we use the term 'science fraud'? The Scientist 6:5:11

ROY L.P. (1999). Review of research protocols. The Lancet 353:9151:428

SCHACHMAN H.K. (1993). What is misconduct in science? Science 261:148-149-183.

SMITH R. (1996). Time to face up to research misconduct [editorial]. British Medical Journal 312:7034:789-90.

Henriikka Clarkeburn, 1999

MISCONDUCT AND INTEGRITY IN SCIENCE

by Henriikka Clarkeburn

It would be strange if there was no misconduct or fraud in science, for why would scientific research be the only human activity immune to misbehaviour, arrogance, greed, and selfishness? But misconduct in science has severe repercussions, because science by nature is an enterprise built on a foundation of trust. Within the scientific community scientists trust that results reported by others are valid. In the relationship between scientific community and society, trust reflects on results to be an honest attempt to describe the world accurately and without bias. Further, there is trust that scientists will co-operate with society in deciding its aims and methods, instead of withdrawing into a detached community within the larger society. This trust has contributed to the unparalleled scientific productivity seen during the past century. Without this trust science could not have gained its position as one of the powers in public decision-making, and it cannot maintain that position if the trust is broken. For these reasons, for science to flourish and to contribute to the improvement of society, it is paramount that the scientific community works to maintain and strengthen this trust. (Alberts and Shine, 1994)

There are two related, but separate, issues in maintaining confidence in science: 1) 'How far beyond what we know, should science reach and what are the methods of research we accept in finding new information?'; and 2) 'How to disseminate research results without breaching this trust?'. The first refers to actions of scientific integrity, the latter to scientific misconduct.

SCIENTIFIC MISCONDUCT

When most people talk about scientific fraud and/or misconduct they refer to obligatory rules that are not opinions or attitudes. This means they are formal guidelines of scientific practice. These rules are traditionally classified under three separate headings: falsification of scientific results, fabrication of research data, and plagiarism. There is common acceptance within the scientific community that scientific misconduct is unethical and that sanctions should be imposed on those failing to comply with the rules (Korenman *et al.*, 1998). Errors occur unavoidably in scientific practice, and this should not be confused with scientific misconduct.

Falsification
Falsification_of scientific observations is an intentional alteration of data or the presentation of observations in a manner which alters the end result. This is most commonly done by altering and/or selecting the research results in a way that is scientifically unjustifiable. The temptation to falsify results is encountered by every scientist, for one of the following three reasons: 1) desire for merit or career advancement, 2) direct financial benefits, and 3) collegial pressure. Results pave the way for an academic career, sponsors of research may wish to influence the publication of results (possibly suggesting omitting results financially strategic to them), or we may not wish to publish data that is not in accordance with the research of our superiors. (see case 1)

Not all cases of falsification are obvious intentional alterations of data. A far more subtle influence is felt by our values and interests, that are not purely scientific. Our own religious and moral convictions alter the ways we view our own results and those produced by others, and influence our acceptance and interpretation of them, as does our bias towards results that confirm the theory, rather than dispute it. Values cannot be taken out of science, because it is impossible to function as a human being without some basic convictions about how and why the world functions. Open scientific practice,

collaboration with colleagues and susceptibility to critique and suggestions, as well as healthy scepticism help in keeping the influence of personal values and perceptions at bay within scientific work.

Fabrication

Fabricated observations are invented observations not based on methods presented in the research report. Fabrications also include entirely imaginary results based on no actual observations. For fabricated results, a scientist is hard-pushed to give any scientific justification, while in the case of falsification, some scientific reasons for the action may be found, though often they are not convincing. Researchers are tempted to fabricate results when they believe they know what the results would be and wish to avoid laborious research to prove the 'obvious'. Another reason is related to external pressures to get publications and/or reports out on time, or to gain financial benefits by presenting work, possibly intended, but not yet carried out. Temptation to fabricate research results is obvious, but classifies as fraud.
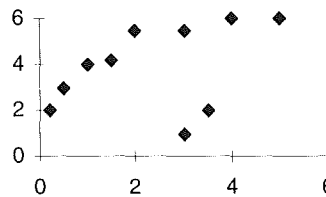
Case 1: Consider the following data set

Katherine and John have been working on a new experimental cell culture. The graph shows their data set from the experiments. A newly proposed theory predicts the measurements to follow the curve in the graph. During the experiments Katherine and John had problems with the temperature controls. When looking at the results, two observations are not in accordance with the expected results. Katherine and John discuss the theory with another group doing similar experiments with the same cell culture, and found out that they had got results confirming the theory. When writing up this research, Katherine and John are tempted to leave the two rogue measurements out of the published graph and from the statistical analysis, as they are 'obviously wrong', most likely because of malfunctioning temperature controls during the experiment. It is clear that these two rogue measurements could be due to the temperature control problems, but there is no way of knowing for sure. How should the two suspected measurements be handled? What would be a suitable punishment for falsification, if Katherine and John choose to do so? (modified from an example found in 'On being a scientist')

Plagiarism

The National Research Ethics Council of Finland (1998) defines plagiarism and misappropriation as: "the adoption of the original research idea, a research plan or research observations of another researcher (misappropriation); or the presentation, either as a whole or in part, of a research plan, a manuscript, article or other text created by another researcher as if it originated from the researcher in question (plagiarism)." Me copying that definition without due credit would have been plagiarism.

The cases of plagiarism are often, but not always, obvious. Difficulties in defining plagiarism and misappropriation can be found when research ideas, methods, and results are informally discussed, and possibly developed further in co-operation. This leads to widening the focus of plagiarism and misappropriation to the problem of allocation of credit in general. It is expected that the principle of fairness and the role of personal recognition works within the reward system of scientific practice. Credit of contribution to a scientific work is in the standard scientific paper given in three places: 1) in the list of authors - all those who contributed directly to the scientific work that lead to the paper should be included as authors; 2) in the acknowledgement of the contributions of

others - those who have supported your research should be explicitly named, and 3) in the list of references or citations - all previous research you used to design and carry out your research, all the data your refer to in support of your research design and results and all additional information you provide based on works of others must be included in your list of citations and references. (see case 2)

SCIENTIFIC INTEGRITY

Where scientific misconduct in clearly rejected as proper scientific practice, issues of scientific integrity are subject to a wide range of opinions. There are two main areas in scientific integrity; 1) methods and 2) objectives of scientific enterprise. Far more than scientific misconduct, issues of scientific integrity relate to personal values of what is

Case 2: Allocation of credit

Over the past two years Professor Hewitt has built a DNA sequencing machine, which allows fast detection of mutations and she is now working on a large sequencing project with her three graduate students. Jocelyn is in charge of operating the machine and analysing the data under Hewitt's supervision. One day Jocelyn notices 'a bit of scruff' in the data. She remembered seeing the same deviation earlier and, by measuring the period of its occurrence, determined that it must come from mutations induced by the machines power surge. Together Professor Hewitt and Jocelyn analyse the data and find similar occurrences throughout the data. This leads Hewitt, Jocelyn and the other graduate students to the discovery of a new mutative source - a major scientific break-through. Professor Hewitt proceeds to report the discovery in a prestigious scientific journal with herself as the sole author. Jocelyn is not happy with Professor Hewitt's decision. What can Jocelyn do in this situation and how should the contribution of Jocelyn and the other two graduate students be recognised?

appropriate in science. These are not solely a matter of personal opinion on what and how. The scientific community holds certain values, upon which the trust between scientists and society is based. To conduct research outside the accepted scientific practices may jeopardise not only your personal career as a scientist, but contribute to erosion of public trust in science.

Scientific methods and integrity
Other scientists and society care about how science is done. This is very apparent in the current animal welfare debate where the public has reacted strongly against animal research, even when it is conducted according to the stringent laws regarding laboratory use of animals, and when the pain and discomfort of these animals is minimised. Similar controversial methods are involved in embryo research, any medical research with human subjects, and involving wild life (zoos, harm of observation in the wild). Some methods of research may be legal, but that does not necessarily indicate that they are morally acceptable. Legality does not guarantee morality, nor does morality guarantee legality. Moral acceptability changes often quicker than legislation, which means the following legal requirements may not be enough for scientific integrity. Science has independent goals and methods, but its functioning is dependent on the desire of the society to support it financially and to allow research to be carried out within the legal parameters set by the society.

Objectives of Science and Scientific Integrity
It has long been debated whether scientists are responsible for the applications and use of the knowledge they produce as a result of their research efforts, and if they are responsible, to what extent? Discovery of nuclear power is a case in point. Nuclear researchers claimed that their work was purely scientific to discover the properties of the nucleus; and that the application of these discoveries to nuclear warfare was only

contingently related to their research efforts, morally no different to the efforts of those, say, providing catering for those who made the decisions. More recently similar issues of responsibility have been discussed in relation to genetic research; genetically modified organisms, cloning, sequencing the human genome etc. The difference is that this time the quest for responsibility has begun before any adverse effects of these scientific discoveries have materialised. And the discussion is not just about safety or misuse of the knowledge, but also a more fundamental concern whether there is a limit beyond which science should not advance. That there may be areas where research is not desired, where more knowledge is considered to bring either undesirable consequences or that there is a fundamental reason why humans should not attempt to obtain certain knowledge. (see Case 3)

For your personal integrity as a scientist, you need to consider the scientific methods you carry out, not only within the legal and institutional requirements, but also as a social and moral issue. You should ask yourself at least two questions: 'Is what I do in the laboratory in accordance with the actions of a 'good person', and 'How likely is my research to contribute positively to the world?'

Bibliography
ALBERTS B., SHINE K. (1994). Scientists and the integrity of research. Science 266:5191:1660-1.
COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY. (1995). On being a scientist: Responsible conduct in research. Washington D.C. National Academy Press.
KORENMAN S.G., BERK R., WENGER N.S., LEW V. (1998). Evaluation of the Research Norms of Scientists and Administrators Responsible for Academic Research Integrity. JAMA 279:1:41-7.
NEWSON A., WILLIAMSON R. (1999). Should we undertake genetic research on intelligence? Bioethics 13:3/4:327-42.
THE NATIONAL RESEARCH ETHICS COUNCIL OF FINLAND. Guidelines for prevention, handling and investigation of misconduct and fraud in scientific research. 1998; Helsinki: Ministry of Education.

Case 3: Genetic research about intelligence (details from Newson and Williamson, 1999).

Professor McCarthy is writing a research proposal to study the genetics of intelligence. Her research aim is to localise several Quantitative Trait Loci (QTLs) for high intelligence by allelic association that tests whether one of the number of alleles of a particular DNA sequence (marker) occurs at a higher frequency in children with high intelligence than in children with low intelligence. A difference in allelic frequency of a DNA marker between the two populations could indicate the presence of an intelligence gene in the region of the marker. Professor McCarthy believes this research will lead to the identification of genes for intelligence and to an understanding of the significance of each gene to intelligence. She also believes that genes do not 'determine' intelligence, but that the environment will have a significant influence on the actual intelligence of a person. In writing her proposal, Professor McCarthy engages in heated debates with colleagues and friends about the usefulness and acceptability of her research aims. Those doubting and objecting to her research refer to possible negative social implications if intelligence genes were found - of people being branded from birth (or even before birth?) to be of certain intellectual capacity, which may lead to discrimination in various degrees. They have also raised a view that research funds should rather be directed towards genetic research of pathogenic entities, not personality traits. Those excited about the research idea, including Professor McCarthy herself, believe that identification of intelligence genes would allow for testing children early on to find out who might need extra support in school and possibly even creating opportunities to enhance the intelligence of future generations. Should the research proposal go ahead?

Appendix XXI - Microbiology PBL

PBL - Vaccine Development

Here is your problem for the microbiology Problem-Based Learning (PBL) course:

"As scientists in microbiology you are asked in consultation to prepare a recommendation for a large charity on whether to support research in developing a vaccine against *Meningococcal Meningitis (serotype B)* or *Tuberculosis.* You are asked to present scientific, sociological, economical, and ethical reasons to support your decision."

According to the nature of PBL, you are free to choose any methods of learning you wish. But to help you to get started, I have written a learning suggestion here, which you are more than free to change in any way you like:

Session 1:
* Start with a brain-storming session to list all the possible things you could study/research/think about before making your decision
* Brainstorm for sources of information
* Consider a division of labour - this is group work, so not everyone needs to do everything!
* Leave with a clear idea of what you are going to do between now and the next session

Session 2
* Share information you have collected since the last session with your group
* Based on this information, decide what should be done next
* Make an action plan for the work to be done before the next session

Session 3
* Share information from your independent study
* Work together towards a shared decision on the problem
* Formulate a final output for the PBL session.

I am here to help you, to answer questions you may have and sometimes to ask you questions that I think might help you in your learning. That means I am **not** here to lecture or organise your learning. You are free to run the sessions yourselves and you are also responsible for your own learning. But don't forget, I am here help you!

Henriikka

h.clarkeburn@bio.gla.ac.uk

*6013

Here is your 'problem' for the second PBL:

"Your tuberculosis vaccine project has progressed for several years, thanks to the funding from a charity. You now have a proto-type vaccine in vitro and your research team is ready to start animal and human testing. For a clinical trial committee you need to submit a trial protocol for both animal and human testing plans. The committee is particularly concerned about the increasing public objection to animal testing and you are asked to give explicit supportive reasons for animal testing you are planning to carry out."

Now that you have experienced PBL before, you are in a good position to organise your own learning even better than last term. I hope you find ways to improve your group work and find the sessions successful.

I am here to help you, but remember you need to ask for help!

Henriikka

h.clarkeburn@bio.gla.ac.uk
*6013

Learning logs for term 2

The assessment for term 2 is based on learning logs again. To remind you, each of the 10 entries should have four elements:

- What you did (1-2 line)
- What you learnt (2-4 lines)
- Was it important and why (10-20 lines)
- How does this shape your next learning unit (1-3 lines)

There is a 2000 word maximum for your logs, though I believe you should be able to say all that is relevant in 1500 words. CRYSTALLISE YOUR THOUGHTS!

You are expected to do better for the next log, because now you have had experience and now better what is involved.

Appendix XXII - Sport Science PBL

The following two swimmers from the same swimming team produced a positive test for illegal substances during out of competition testing.

1. Female age 21, best achievement 10th in National Championships last year, at the moment recovering from shoulder injury, uncertain about the date of next competition. Illegal substance Dextropropoxyphene (narcotic analgesics)
2. Junior (14) male swimmer, in training for his first national championships. Illegal substance Nandralone (anabolic androgenic steroid AAS).

The UK Anti-Doping Directorate and UK Sports Council are proposing a 2 year suspension for the female swimmer and a 4 year suspension for the male swimmer.

As a response to these test results and proposed suspensions, the team leader John Williams released the following statement:

"According to the Olympic rules, no drug may be taken that is ergogenic (performance enhancing). In my understanding this rule is applied in a selective manner. There is no ban for using legitimate drugs, for example antibiotics for infection, which undoubtedly enhances the performance of some athletes. This is considered to restore 'normal' performance, and the problem is said to be in substance use to enhance the performance beyond 'normal' maximum. But again, not all methods of such enhancement are illegal, such as blood doping, known to be ergogenic, is not tested for, neither are techniques of carbohydrate loading, or megavitamins included in the illegal methods or substances even though they have positive impact on performance. Enforcement of inconsistent rules is not morally acceptable.

Much of the anti-doping literature emphasises the hazards of these drugs to the athlete. This is likewise contradictory as many sports, boxing, or rugby for example, by their nature are dangerous to the athlete's health. Furthermore, athletes often put themselves under risk of on-going physical disability by excessive training or training and competing while injured and sometimes this is facilitated by legitimate drugs, such as non-steroid painkillers. If athletes are considered mature and competent enough to make decisions about their training and participation is sport, surely they should be granted the same freedom to choose the medication the wish to use?

Everyone who follows the doping debate is familiar with the argument that use of ergogenic drugs in sport gives unfair advantage to some competitors. Again

the fairness argument follows the line of restoring normal performance (legal, thus not unfair) and exceeding 'normal' maxims (illegal, thus unfair). There are at least four ways to look at this issues. First we can ask 'why should sport be fair?', especially in a society that is not fair. Is there some special element in sport, that sets it apart from the rest of the society and requires it to be a model area among variety of human activities? Second, if ergogenic drugs were not condemned, and it would be illegal to use them, the question of unfairness would no longer exist. Surely the ability to use the drugs in a best possible way would be in no significant way different from the current differences between training methods. Third, would the image and role of sport change if egrogenic drugs were made legal? Sport is entertainment and better results could attract bigger audiences and as such enhance the image of sport by enlarging the circle of people participating in it. And finally, athletes are professionals, they earn their living with their sporting performance. Why shouldn't they be allowed to earn their income with all the possible methods available?

When it comes to the particular cases of positive doping-tests in my swimming team, I would like to make the following comments. Laura, who tested positive for Dextropropoxyphene, was given the drug by myself as pain relief for her serious shoulder injury. As was found in the Samantha Riley's case earlier, Dextropropoxyphene does not enhance the performance of swimmers and Laura is not planning to compete for the next couple of months at least. I was unaware that the painkillers I gave to Laura contained Dextropropoxyphene and thus she is no way responsible for her positive test result. As for Rob's positive test result for Nandralone, I was not aware of Rob's current substance use, while I was aware of his contacts with some steroid suppliers. Rob is a very promising young swimmer, who has not yet competed in any National competitions. I believe it is not acceptable that a young athlete of no previous national competition appearances is included in the out of competition testing protocol. In my view it results in a violation of his privacy and right to conduct his swimming pursuits in a manner he chooses, at least as long as they still are only for his private enjoyment.

I thus put forward a motion of removing the suspension on Laura and reduce the suspension of Rob to three months, based on the moral considerations and case particulars I have put forward here."

You, as specialist in Sport Science, are asked to collate a response to Mr Williams' statement. The UK Sports Council has asked for a report that would recommend suspensions for the two athletes and answer in detail the comments made by Williams.

Appendix XXIII - Introduction to Learning Logs

LEARNING LOGS
brief introduction

During this course you will keep a learning log. This is an important part of learning and it will contribute to your **continual assessment**. Learning logs allow you to reflect on how you are progressing with your study and they encourage you and help you think in a 'deep' manner. When you complete a log you are doing more than writing a diary. A diary records events and your commentary of events. **The learning log is an account of your thinking.**

Each unit of study (it could be reading, discussion, or just inner thinking) should be included as an entry into your log. Your first task for every entry is to set yourself some starting criteria. They are best put in terms of questions:

What have I learned from this learning unit?

To answer this, consider your learning unit, and think what has been:
* important
* meaningful
* good
* decisive
* uncomfortable

Have a conversation with yourself in the log. Write honestly/frankly about issues that mean something to you.

In order for the log to be assessed, you will need to spell out the criteria for you comments. In other words, if you write that something is meaningful, then you must also note:
        'How do I know this is meaningful?'
so that the marker will be able to understand where your thoughts come from and make sense of your log.

You are free to structure your log as you choose, but in every entry you should include information on the following four elements:
* Briefly state what you factually did: what reading are your thoughts based upon (provide a full reference!), who did you talk to, or what spurred your thinking.
* Explain what did you learn and why it was important to learn this.
* Describe how the learning will shape your next learning unit.
* Note what issues you feel this learning unit did not answer satisfactorily.

For this course you are asked to have **minimum 10 entries** in your log. The 11th/last entry should be a conclusion for the entire log, which is best written as a statement of what you learnt, what you still feel unsure of, and how well you think you did in this course.