

**The London School of Economics and Political
Science**

*The Value of Other People's Health: Individual models
and motives for helping.*

Christopher James Desmond

A thesis submitted to the Development Studies
Institute of the London School of Economics for the
degree of Doctor of Philosophy, London, February
2008.

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without the prior written consent of the author.

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

C. Desmond

Abstract

This thesis was motivated by a wish to understand the slow response to HIV/AIDS and so seeks to address the questions of why and when people help others. The questions are important, as typically it is others who decide who among those most in need receive help. To narrow the focus to motives of help, the discussions consider the helping of distant others, as typically such actions do not involve material gain to the helper.

Related literature exists on charitable behaviour but tends towards a marketing approach. This limits its use in providing specific input regarding motives. The economics literature is far more specific and a review identifies four groups of models. These models have a number of weaknesses, so an alternative is suggested: the balance model, and a responsibility formulation, is proposed for this and existing models, which, it is argued, assist in addressing the questions.

Following on from the identification and development of alternative theoretical frameworks, these options are taken through a process of attrition. They are contrasted with evidence and theory from the psychology literature - first on helping and then on harming. This exercise suggests the relative strength of the balance model and the maintenance of a responsible self-image model. Both deal well with explaining how help differs according to context and how individuals might avoid or deliberately misinterpret information.

The balance model is used to examine individual helping behaviour, by way of an economic experiment. The model is then expanded to consider the social context, which allows for the consideration of the HIV response.

The theoretical discussions and the experimental results suggest that individuals feel a responsibility to help. They can, however, try to avoid information which may prompt such feelings and manipulate themselves to reduce the pressure such feelings place on them to help.

Acknowledgements

While only one name appears on the cover, a long project such as this requires the support of many people and thanks are due. Firstly, I would like to thank and acknowledge Prof. Tony Barnett for his support in getting me to the LSE and his excellent supervision over the course of my studies. Tony allowed me great freedom, but with sufficient guidance to prevent me from getting lost. The many long sessions of supervision and the very quick turn around time to drafts, with insightful comments, have been greatly appreciated. Many of the ideas outlined in this thesis result from, or were refined by, discussions with Tony. I would also like very much to thank my second supervisor Dr. Erik Ester who was very generous with his time and helped greatly in the conceptualisation of the argument.

I would not have been able to undertake my studies at the LSE without funding. I must thank the Economic and Social Research Council for their financial support throughout the period, the Ernest Oppenheimer Memorial Trust for the support towards my fees for the second and third year and to the Human Sciences Research Council for their support in the first year.

In order to undertake my studies it was necessary to leave my job for a period and I would like to thank the Human Sciences Research Council and in particular Prof. Linda Richter for creating the space for me to do so, then for helping me return and finally for supporting and encouraging me to finish.

I would like to acknowledge the support of those who helped me in the research: Andy Gibbs who assisted in the running of the experiments, Janet George for assistance with the references and my good friend Lincoln King who programmed the experiments presented here. It is necessary also to note the contribution of two friends prior to the commencement of this work. Much of my initial thinking on the question of self-interest grew out of an argument involving Kurt Sass and Sheldon Magardie some years ago concerning whether you can do anything that is not in

some way for yourself. I should add to this the contribution of many people since, with whom I have discussed my arguments, they were more a part of the research than anything else I have done.

Finally, I would like to thank and acknowledge my family and in particular my parents. They have encouraged and supported me throughout my education and without them I would never have reached this level – this is truer for me than for most. In many ways this is their PhD and they should acknowledge me for writing it. A special additional acknowledgement goes to my Dad who reads and edits everything I write. Without his patience and support I would surely be on a different career path and would never have undertaken this PhD.

Table of Contents

Declaration	2
Abstract	3
Acknowledgements	5
Table of Contents	7
List of Tables	11
List of Figures	11
Chapter 1: Introduction	12
1.1 The question	12
1.2 Background literature and limitations	16
1.3 Alternative conceptions	22
1.4 Cutting the number of alternative models	29
1.5 Application to practical questions	31
1.6 Considering the social context	32
1.7 Thesis structure	34
Chapter 2: Economic theories of helping	37
2.1 Introduction	37
2.2 The economics of altruism and moral behaviour	39
2.3 Summary and responsibility	74

Chapter 3: Comparison with psychology literature on helping	77
3.1 Introduction	77
3.2 Comparing economic models with the psychology literature on helping	79
3.3 Summary and conclusions	115
Chapter 3 Appendix: Game theory example	119
A3.1 Introduction	119
A3.2 Purely self-interested	119
A3.3 Incorporation models	120
A3.4 Private benefit of action	125
A3.5 Private benefit of maintenance	127
A3.6 Multiple-self models	129
A3.7 The balance model	130
Chapter 4: Comparison with psychology literature on harming	132
4.1 Introduction	132
4.2 Denial	135
4.3 Harming others	143
4.4 Helping in extreme situations	150
4.5 Conclusions	152
Chapter 5: Questions and experiment methodology	156
5.1 Introduction	156
5.2 Health priorities literature	158
5.3 Research design	165
5.3.1 The sampling	166
5.3.2 The instrument	167
5.3.2.1 Basic socio-demographic questions	168
5.3.2.2 Allocation decisions	169
5.3.2.3 Follow-up questions	177
5.3.2.4 Pay-off matrix	178

5.3.3 The process	180
5.3.4 The pilot	182
5.3.5 The basis for content hypotheses tested	183
5.4 The analysis	187
5.4.1 Assessment of donations and associated feelings	189
5.4.1.1 Differences within cases	189
5.4.1.2 Differences across cases	193
5.4.2 Analysis of case ten responses	195
5.5 Summary	195
Chapter 5 Appendix: Technical issues and details	196
A5.1 Technical issues	196
A5.1.1 Model outline: ordered regression models	196
A5.1.2 Model specifications and specifications of formal hypotheses	200
A5.1.2.1 Differences within cases	200
A5.1.2.2 Differences across cases	201
A5.2 Experiment design	203
A5.2.1 Summary of verbal introduction	203
A5.2.2 Handout: Explaining a pay-off matrix	204
A5.2.3 Introductory text and the interface	206
Chapter 6: Experiment results	213
6.1 Introduction	213
6.2 Description of data	213
6.3 Examining within-case variation	217
6.3.1 Emerging across-case issues	224
6.4 Determination of donations and feelings across cases	225
6.5 Case ten	229
6.6 Discussion	232
Chapter 6 Appendix: Detailed results	241

Chapter 7: The individual in context	254
7.1 Introduction	254
7.2 Information preference interaction and the role of the media	257
7.3 Media selection and presentation	266
7.4 Opinion and policy link	274
7.5 Expanding the model to consider context	279
7.6 Explanatory power	283
7.7 Conclusions	291
Chapter 8: Extensions and implications	292
8.1 Introduction	292
8.2 Constructions of responsibility	293
8.3 Integration of social norms	298
8.4 Payment for responsibility	306
8.5 Different implications	308
8.6 Conclusions	318
Chapter 9: Summary and final question	319
9.1 Introduction	319
9.2 The process	319
9.3 The argument	326
9.4 The implications	328
9.5 Final question	330
References	332

List of Tables

Table 3.1: Predictions of helping behaviour 1	101
Table 3.2: Predictions of helping behaviour 2	103
Table 3.3: Predictions of helping behaviour 3	104
Table 5.1: Set and case differences	172
Table 5.2: Follow-up questions	178
Table 6.1: Respondent characteristics	214
Table 6.2: Responses	216
Table 6.3: Donation regression	226
Table 6.4: Sadness regression	228
Table 6.5: Pressure regression	229
Table 6.6: Case 10 responses	230
Table 6.7: Significance tests of case 10 responses	231
Table A6.1: Donation differences across sets by case	243
Table A6.2: Case version differences in donations and follow-up questions	245
Table A6.3: Summary of significant differences across versions of each case	250
Table A6.4: Donation regressions	251
Table A6.5: Sadness regressions	252
Table A6.6: Pressure regressions	253
Table 8.1: Combinations of motivation and responses	314
Table 8.2: Possible combinations of motivation and responses	314

List of Figures

Figure A5.1: The ordered probit model	198
Figure 7.1: Expanded model, simple version	279
Figure 7.2: Expanded model, full version	282
Figure 8.1: Individual relationship to social norm	301
Figure 8.2: Continuum of choice	316

Chapter 1: Introduction

1.1 The question

Countless newspaper articles, research reports, theses and academic papers have opened by detailing the magnitude of the HIV and AIDS crisis. They present prevalence, deaths and orphan numbers all in the millions. The magnitude of the suffering is lost, as the world appears unable or unwilling to comprehend what these figures mean. The HIV and AIDS epidemic is of a frightening scale and the figures are no longer only predictive, but are increasingly describing a historic, current and ongoing tragedy.

Reaction has been slow in coming and focuses almost entirely on symptomatic interventions. The response to HIV and AIDS is taking shape as if the epidemic came from nowhere. The socio-economic roots of the crisis are largely ignored. The virus is slowly being treated but the social ills, such as family-separating labour migration, which have allowed it to thrive are barely mentioned; it is almost as if those in power do not wish to admit the role such ills play.

The response has taken a long time to come, not because those who have been dying did not think treatment was a priority or because those left behind felt that they were not in need of assistance, but because those with influence did not prioritise and support a large-scale response. This practice of others deciding on the priorities for those in need does not only apply to HIV and AIDS, but it is this instance that prompted this work.

Triage is conducted by medical staff not patients; health priorities are set by the healthy not the sick. The burden of disease in the world falls largely on poor countries and, even more so, on their poorest citizens (Murray and Lopez, 1997). Resources are limited and the burden is great, difficult choices have to

be made. While in some States democratic processes and other political structures may give expression to the views of those most affected, practical planning and priority setting are typically beyond their control.

It is not, therefore, those who suffer the greatest ill health who decide what suffering, or potential suffering takes precedence when forming responses. While politicians, bureaucrats, researchers and influential members of the public have power to influence health policy, it is unlikely that they themselves are the ones who suffer most. Similarly, the poor of the world do not decide on the appropriate division of aid, or which charities are established or how much funding these receive. It is the priorities of others, or others' perceptions of the priorities of the suffering, that are reflected in policies, distributions of aid and establishment and funding of charities.

Efforts have been made to measure suffering, to find 'objective' means of prioritising across competing causes. These measures themselves, however, reflect what others have considered as the important aspects of suffering, or what they believe those suffering consider as important. They do still try to identify the priorities of those affected and this is indeed important, but given the important role played by those who are not affected, time needs also to be spent on understanding what they prioritise and why. It is not being suggested that this time should be spent because the opinions of others should be considered for any moral or political reason, but because they have so much impact on the welfare of those who are in need.

Local policy makers, international players and members of the public from wealthy countries are presented with research, media images of the suffering of others, or maybe they have seen it for themselves. They respond to some and not to others, at times with great gusto and sometimes hardly at all. Given the importance of these individuals in shaping priorities, understanding when and why they help or advocate for help is critically important.

The processes whereby individuals set priorities for others raise many interesting and difficult questions. How does the nurse in a poor rural village decide which child receives treatment when supplies are short? How do local health policy makers prioritise limited funds? On what basis do researchers develop measures of suffering? How do development agencies in the wealthy world allocate aid to those in need? Why do the public in wealthy countries call for more aid in response to one crisis while remaining relatively or completely silent on another and how do they decide which charity to give to if they decide to give at all? In more general terms, why do people take responsibility for responding to some needs while avoiding it for others?

Are those who can influence responses motivated only to try and alleviate as much suffering as possible or are other motives at play? Is the answer to these questions that people do their best to identify the greatest need and respond as they think is appropriate?

At all levels, the answer to this final question would seem to be no. Prioritisations, and the associated valuations of responding to the needs of others, are likely to be based on a host of motivations. At the local level, power plays, political processes, family ties and many other factors may have an influence. At the international level, public relations and strategic concerns, among other variables, could be considered. The answers to the above questions are clearly complex and messy.

From this mess, this thesis aims to disentangle and examine in detail one important aspect: the valuation of different people. Arguably there are two aspects to the differential valuations being discussed here: the differential valuation of the individual and of actions to alleviate the suffering of that individual. It is the latter which is of primary concern here, although many arguments would have it that it is the former that shapes the latter.

It is obvious that people value helping some people more than others. Family and friends are clear examples; people may happily make efforts for family members that they would not make for strangers. There is a moral question here relating to the justifiability of differentiating between the values one places on the welfare of others. With regard to placing a higher value on helping family and friends there are few who would say it is not justified (Chatterjee, 2004). The debate, however, becomes more difficult when higher value is extended to include not only friends and family, but fellow citizens of a country, or members of the same religion or ethnic group (Chatterjee, 2004). The purpose of this thesis is, however, not to focus on the moral debate as to the justifiability of differential valuations of people or action. It is rather to accept their existence and attempt to understand their nature and the role they play in influencing patterns of support to those in need.

To investigate the role played by the differential valuation of others' welfare it is helpful to narrow the investigation. At the local, national and international levels there is likely to be a multitude of factors and motivations shaping prioritisations. To examine the specific impact of one motivation, however important, would be extremely difficult. In order to isolate the impact of this motivation as best as possible, the area where it is likely to play the most prominent role will be investigated. It is argued that that this area is in the formation of public opinion and in the selection of charitable causes by citizens of the wealthy world, specifically in responses to human health crises in the developing world. This avoids many of the strategic and economic concerns which shape the behaviour of international actors, while still remaining free of the complexities of domestic priority setting. Such narrowing also provides space for interesting theoretical discussions, as motivations for helping relating to reciprocity and family are difficult when the other is at a distance. That these cannot be the reasons forces attention towards other possible motivations which may prompt individuals to help.

This thesis, therefore, focuses on the narrowed topic of the role played by the differential value placed on the welfare of others or rather on actions to affect that welfare, specifically with regard to the health of others, in the formation of public opinion and the determination of individual charitable giving in response to human crises which are distant from them. Distance is considered here in the physical sense, although this may influence perceptions of social distance. While this narrow focus is necessary for the investigation, the discussions will consider what role such differential valuations might play more broadly. Factors that shape individual behaviour when considering distant others may well play a role when considering family or members of one's own community. It is not being suggested that the arguments here are only relevant when distance is an issue; rather it is argued that it is easier to isolate this one particular area of interest in this setting. After its isolation here, the possible implications that differential valuation might have in more complex contexts can be discussed.

The focus on health is primarily a result of the starting point being HIV and AIDS. Examining health specifically is useful in that it concentrates the discussion while not being too limiting, as so many factors can be argued to impact on health. What is more, health provides a very clear example of where others are setting priorities for those in need and where that need is often very clear and urgent - characteristics which may prove important in shaping responses. Health is also both an important determinant and an aspect of development.

1.2 Background literature and limitations

While not always health specific, substantial research has been conducted relating to charitable giving and the formation of public opinion. Research has been done to measure the scale of giving to charity and the somewhat similar exercise of measuring giving by government to developmental causes (see for

example Pharoah, 1996 for individual giving and for international assistance OECD, 2005). Measuring the total is, however, of little relevance to this work; what is more interesting are the efforts to understand why people give, how they determine how much to give and on what basis they decide what to give to.

Highlighting some of the factors considered in the research in this area is informative and will draw attention to its limitations with regard to the questions of this thesis. The first step in research in this field has been to investigate the characteristics of who gives, for example: the educational profiles, income levels and gender of donors and non-donors. This is typically achieved through the analysis of household expenditure data, where demographic characteristics can be linked to giving occurrence and level (such as Jones and Posnett, 1991; Jones and Marriott, 1994), or through the examination of donor groups such as alumni (Clotfelter, 2001).

Not only have the characteristics of donors been examined, but also the differences across characteristics of self-reported motivations to donate. For example, Silver (1980) discusses the different motivations of low-income groups compared to high-income groups. Kottasz investigated the stated motivations among more affluent individuals with a particular focus on the differences between men and women (2004).

Associated with motivation, have been considerations of selection between different charities: for example, the role played by a charity's reputation and perceived efficiency in determining selection (Bennett and Gabriel, 2003), or the relative importance of different causes to individuals, an issue we shall return to shortly.

Motivations to give have been associated with the possibility of individual benefit. Factors that have been examined and discussed in the literature

include a sense of belonging, career advancement, tax advantages, political gains, and peer pressure (American Association of Fundraising Council, 1994).

Increasingly, work in this area has come from a marketing standpoint, drawn in as the number of charities increases and donations stagnate (Sargeant, 1999). Schlegelmilch and Love, for example, examined how individuals with different socio-economic and demographic characteristics differ in terms of their responses to alternative fund-raising approaches (1997). In effect, they are attempting to assist in targeting by building donor profiles associated with the success of different approaches.

Sargeant provides a review of much of the literature on motivations behind charitable donations and develops from it a model of giving, aiming to understand how the decision to donate is made and what key variables influence that decision (1999). The review considers contributions to this debate from economics, clinical psychology, social psychology, anthropology and sociology and the more recent marketing literature. His paper draws together inputs from these various disciplines into what is essentially a marketing model, the purpose of which is to assist in targeting and framing of funding efforts.

Sargeant's model traces through the likely impact of charity appeals, dealing first with the potential donors' perceptual reaction to these appeals, which he argues are based on such things as the strength of the stimulus and the portrayal of the cause. Following this, the model suggests that a decision whether, and what, to give is made based on processing determinants, which include factors such as past experience and judgement criteria. Sargeant adds another dimension to the model by arguing that the literature reviewed suggests that the perceptual reaction and the processing determinants are both influenced by individual extrinsic and intrinsic determinants. Extrinsic determinants include age, gender, social class and income. Intrinsic

determinants include self-esteem issues, feelings of guilt and pity, questions of social justice and responses of empathy, fear and sympathy.

By developing frameworks, such as the above, which incorporate various research results, it becomes easier to investigate broader issues. Bennett, for example, investigated the basis of selection between charities involving human and animal welfare (2003). Micklewright and Wright made use of Sargeant's model to frame a discussion of a number of the factors raised in the literature regarding motivations and charity selection in regard to development aid (2003). They outline how these factors may have come to determine the relatively low level of development aid compared to programmes aimed at addressing domestic problems. They seek to use this understanding to suggest approaches that can assist in building support for development work.

While research like that described above is insightful and useful, particularly to charitable organisations, it is largely limited in its use by its marketing design. For much of the work that has been done, it would appear that the purpose was to assist charity fundraisers in the design of their appeals so as to better frame and target their approaches. Applying marketing models, such as Sargeant's, to the questions posed previously would likely yield a similar result to Micklewright and Wright's application to development aid. It would be useful, as it may well highlight many of the possible factors which are related to seemingly different valuations; and for those wishing to increase funds directed towards certain causes it may result in suggestions for more efficient fund-raising methods. It would not, however, provide a means to understand why these factors are related; so it is limited in how much it can say about who will be helped and in what circumstances. The approach identifies correlations and so to some extent is descriptive rather than predictive as it does little to identify causal relationships. Producing models that allow for a host of factors to be included means that the need to give thought to underlying motivation is bypassed by including everything.

Given the narrow marketing nature of research on charitable giving, it would appear to be beneficial to bring in insights and approaches from elsewhere. Economics is at times accused of being simplistic in its approach to explaining human behaviour because of its reliance on simple models of behaviour, typically with a single underlying motivation. While this may be so, examining the above issues through an economics lens forces the examiners to refine their arguments. It is not acceptable to say that this and this will influence behaviour in such a way, it is necessary to suggest why. Economics is not the only discipline to require such. Psychology requires similar detail, but leans towards more comprehensive explanations rather than the simple ones sought in economics.

The above conception of the questions dealt with in this thesis, asking what charities and causes are responded to, leads to the type of research mentioned. Using a more economics and psychology approach, the questions can be reformulated so as to address the heart of the issue.

Reframing the questions separates the thesis into practical and theoretical components. What do people take into account when considering whom and when to help when the 'whom' is distant from them? This question forms the core of the thesis's practical side, but to address the question more fully it is helpful to ask the question why people help those distant from them. There are a number of theories on this and the discussion of these, and the proposal of another, are the primary concern of the theory dealt with from here on.

Models of helping behaviour, or altruistic behaviour, as it is more commonly referred to in economics, have been developed in economics and numerous psychology studies, particularly from social psychology, have examined the issue. A number of different motivations for altruistic behaviour have been identified, including social acclaim, reciprocity, self-image and a host of others.

None of these models is presented with an argument that they capture 'the' motivation for help, but rather 'a' motivation. There are many different circumstances in which individuals act altruistically and they are likely to do so for many different reasons. The interest here is in which of the models can be applied to explaining the helping of distant others.

Actions that appear to be for the benefit of others have presented a challenge to economic models based on very narrow definitions of self-interest, but many have responded to this challenge and a range of alternatives have been proposed. A review of the economics literature in relation to this area of study identifies a number of existing models (Chapter 2). These models can be grouped in the following way: incorporation, private benefit of action, private benefit of maintenance and multiple-self models.

Incorporation models essentially argue that the welfare of others enters the welfare function of the decision maker: the decision-maker's welfare then becomes linked to that of others, so they help others as part of the process of maximising their own welfare.

Private benefit of action models are a group of models which are based on assumptions regarding the private utility of certain actions. Individuals are argued to get some benefit from actions that help others and so help in order that they receive this benefit.

Private benefit of maintenance models are very similar to private benefit of actions models but the difference is important. While in action models the benefit is from the action itself, in maintenance models the benefit is in what that action helps maintain. For example, both models could be based on arguments relating to the importance of self-image. An example of a private benefit of action model could be a model which is based on individuals gaining a benefit from helping others because it boosts their self-image and they value

this boost. An example of a private benefit of maintenance model could be a model where the individual values holding a positive self-image. They will act to help others in defence of this image, but it is not the action that they value, it is the maintenance of the self-image. If they could maintain it in some less costly way than helping then they would.

Multiple-self models stem from a difficulty, which many authors have, with seeing all actions of helping, and indeed all moral behaviour, as essentially self-interested. These models propose that individuals have more than one utility and that helping behaviour is driven by a motivation other than self-interest.

These existing models are undoubtedly useful in explaining aspects of altruistic behaviour in the context of interest; there is, however, room for improvement and additions. The theoretical component of this thesis seeks to do three things: suggest a formulation which can be applied to existing models to better explain helping of distant others; suggest an alternative model and again apply the formulation; then critically compare all the models. The exercise is essentially one of development and attrition. A formulation and alternative model will be developed expanding the number of options under consideration. All of these options will then be critically evaluated, gradually reducing the number of surviving alternatives to a manageable number so that they might be applied in the practical sections of this thesis.

1.3 Alternative conceptions

The first part of the development portion of the thesis is the development of a potentially useful formulation. It is argued here that a formulation for understanding helping of distant others, which is of use, is a responsibility formulation. This formulation essentially suggests that the question of responsibility is central to the decision-making process. Responsibility is a

useful construction as it takes into consideration a host of contextual factors which appear to come into play when deciding whether or not to help.

This formulation can be applied to many of the existing models but arguably with most effect to self-image models, particularly those falling into the private benefit of maintenance group. As mentioned above, self-image models of this type are based on the assumption that individuals value feeling that they have certain qualities: for example, in the Dana et al version they value feeling that they are fair (2007). Valuing feeling that you are fair is very different from valuing fair actions. In this thesis it will be argued that a model based on the assumption that people value feeling that they are acting responsibly will better explain behaviour in relation to helping distant others than existing alternatives. Although this model is argued to fare better, it is still firmly within the broader group of behavioural models based on the general assertion that, one way or another, actions are based on self-interest: that is, the models are based on motives which could all be characterised as 'wants'.

It is this assertion that everything is self-interested that has led to the above-mentioned multiple-self models. These models attempt to capture the differences in motives, such as 'ought' compared to motives such as 'want'. These models, however, struggle to develop a decision theory that accounts for how the multiple utilities or motivations interact without collapsing into a higher order mono-utility. Although this failing is important, it is still important to appreciate the difficulty with the self-interested assertion apparent in this group of models.

While it may not be relevant in day-to-day decision making, it is useful to examine this general assertion at the extremes. At the extremes the question becomes: do people have 'haves'? Alternatively, do people have 'cannots'? Or is every action essentially self-interested? While perhaps a little differently phrased, this is a very old debate, but very relevant here.

On your way home from work you pass by a child drowning in a lake in an otherwise empty park. No doubt you want to save the child, but is the motivation of the same type as that which pushes you to want steak for dinner? Could the situation be so extreme that you feel that you have to help? Could it be that there is some internal pressure to help so great that it cannot be ignored? Taken from another angle, is there an amount of money great enough that you would accept it as payment for killing an innocent child or allowing them to be raped? Or is the thought so revolting that you could never bring yourself to do such a thing or even consider it?

There are those who do leave people to die and who do rape and kill innocents, but we shall return to them later (Chapter 4). The point of discussing the extremes here is to get a sense that certain motives are different; that arguing that you save the child only to avoid denting your self-image may be missing something. There are ways in which self-interested models can be contrived to explain why you might help no matter what the cost or why you might refuse to undertake an act no matter what the benefits, but these models fail to capture the idea that different types of motives shape these actions.

Embarrassment is a useful comparison. People do not do what they say they really want because they are embarrassed to do so and may go for years regretting their inaction or trying to get past their embarrassment, as if it were a constraint. Again a self-interested model could be constructed so as to explain this, but doing so seems to miss an important aspect of the decision: the playing off of one feeling (want) against another (embarrassment). The goal should not only be to describe action, but as far as possible to understand it, as this may have implications for its interpretation.

In mainstream economics the decision process is framed in relation to utility. When considering a choice between food and drink, the net benefits of both are

predicted and the choice is made in favour of the prediction which scores higher in relation to subjective utility. An alternative way of thinking about this is that feeling hungry or thirsty leads to different feelings that place a pressure on the individual to eat or drink. The individual then probably heeds the greatest pressure after considering the cost of addressing it. At first glance this would seem an identical argument but, while it is very similar, there is more to it.

Imagine that thirst and hunger have a range from extreme to not at all. Imagine further that you are in a situation where your thirst is approaching the extreme of its range and your hunger is at roughly mid-point. In such a situation, when presented with the option of drink or food you would obviously take drink. Until your thirst is far enough from its extreme that attention can be given to weighing up of options, you can think only of ways to quench it. Now add to this a whole host of feelings, both physically and mentally prompted, and you have the image of an individual trying to push each feeling along its range to the comfortable lower end while avoiding any climbs towards high pressure extremes; it's a balancing act. It is a balancing act because the feelings are not perfectly substitutable. Eating when you are afraid may help a little, but efforts to relieve a negative pressure by attending to another will likely have limited success.

In this model of decision making, when all of these feelings are far enough from their extremes they can be traded off against one another to maintain maximum overall well-being. This is close to, but not the same as, a mono-utility model. The individual seeks balance and balance may not always be the same as utility maximisation. If some or even one of the motives and feelings are not self-interested then, even if balancing them is self-interested, the outcome of this balancing may not be the same as utility maximisation (Chapters 2 and 8).

The difference in the outcome is even starker when the extremes of non-self-interested motives are considered. If it is argued that as any feeling approaches

its extreme so attention to it approaches total, then effectively only choices which pull that particular feeling towards normality are considered when the feeling is sufficiently close to its extreme. Using utility theory, it could be argued that this would simply be a way of understanding marginal utility and diminishing marginal returns: that you gain more benefit from improvements in feelings that are at extremes than from improvements in feelings that are not so far along their range.

It is probably safe to suggest that most people would agree we have feelings of right and wrong, of good and bad. The question is, is it safe to suggest that some actions feel so wrong that we can't pursue them or so right that we have to respond to them? Moreover, can such a situation be conceived of as self-interested?

Take a statement such as 'I cannot commit murder, it's not in me, but in protection of my family I will do whatever necessary'. To allow for murder, a rationale on the same right-wrong scale is needed, individuals justify the act to themselves. It is this non-substitutability which differentiates the argument from the traditional utilitarian approach. But at the same time an important possibility is raised. Feelings such as thirst, hunger or pain are physical in origin and are difficult to ignore or manipulate. Feelings such as right or wrong or embarrassment are not physical in origin and may well be open to mental manipulation. So, while extreme feelings of wrong may constrain action, individuals may convince themselves that the action is 'not so bad' and in doing so relieve the constraint.

It may seem that this discussion has come a long way from understanding the basis of altruistic behaviour. The above, however, is the beginnings of the argument for an alternative model: the balance model. The model which will be developed over the course of this thesis argues that individuals have a range of feelings that motivate action. These feelings are imperfect substitutes for each

other, requiring the individual to address each as they arise to maintain balance between non-substitutable feelings. Some of these feelings relate to relationships with others and may not always be entirely self-interested. Their existence could be conceptualised as moral constraints on behaviour. These constraints, however, as well as other mentally-generated feelings are open to manipulation. If individuals do not have moral constraints then they are worse than Sen's rational fool (1977); they are the rational psychopath – purely self-interested with no moral constraints or conscience.

The balance model can also be placed within a responsibility formulation. It can be argued that people feel a differential responsibility to help, depending on who needs help and in what context that help is needed. Further, it can be argued that this feeling can, to some extent at least, be manipulated by the individual. This feeling is one of many that individuals trade-off against each other when evaluating alternative courses of action, so long as none are close to their extreme. When they are close to their extreme, they become difficult, if not close to impossible, to ignore as attention to them approaches total.

This feeling of responsibility could be argued to be an example of a non-self-interested motivation. This, however, is primarily of theoretical interest as, in terms of predictions, it makes little difference if responsibility is considered a non-self-interested internal pressure or if individuals are assumed to value maintaining a responsible self-image. Both of these versions of the responsibility formulation can arguably be used to explain behaviour better than the alternatives currently available. Both of these models are based on the value being placed on internal states not on external actions or consequences. Actions and consequences may influence the internal state, but they are not valued in and of themselves. The valuing of internal states is what opens the way to self-manipulation. It is this possibility for self-manipulation and the associated attitudes towards information which sets these models apart from the majority of those currently used. The introduction of self-manipulation and

the valuing of internal states require a consideration of the psychological processes that would be involved. Arguments relating to self-manipulation appear to suggest the existence of an actor and a watcher, along the lines of a split ego. Within the psychology literature there is a variety of alternative explanations of how self-manipulation may occur, some of which do not rely on splitting. It is not important here to resolve which of these arguments is the strongest, but rather it is important to identify if there is a possibility for such processes, which there does appear to be (Chapter 4).

Thinking about when people help, using either self-image or balance models in their responsibility formulations, is useful and very informative. For example, they assist in explaining why people would avoid information on suffering or why they might be motivated to interpret information so as to reduce their responsibility. The idea of resisting information because of the responsibility it might bring with it does not fit in well with other models, but does appear to occur. Statements like 'I don't want to know, because if I know then I'll have to...' are difficult to explain outside of these models. Opening up the possibility of self-manipulations, such as denial, takes most economics behavioural models into somewhat uncomfortable territory, but in return provide much more explanatory power, helping to explain a range of occurrences. If people simply wanted to help, why would they try so hard to avoid charity appeals or requests from beggars in the street? Some charity programmes even play to this characteristic: paying a monthly amount to a charity is 'doing your bit' absolving yourself of responsibility elsewhere or at least giving you the chance to argue that to yourself.

1.4 Cutting the number of alternative models

Introducing the responsibility formulation and the balance model expands the number of options for a theoretical framework for considering the issue of helping distant others. Problems with some of the older models have already been noted in the economics literature and are clearly apparent when applied to the problem at hand (Chapter 2). Incorporation models struggle to explain how helping behaviour is shaped by context. They appear to need a theory of when others' welfare will be incorporated and when it won't. Depending on which theory is used, incorporation models end up in one of the other groups. Private benefit of action models clearly have some strengths, but they struggle to explain attitudes towards information. A consideration of the economics literature alone does not, however, trim down the options sufficiently so that there are few enough alternatives to manageably proceed to their practical applications.

Fortunately, economics is not the only discipline to have examined such issues. Closest in character to economics is the psychology literature. In this literature a host of experiments have been conducted and theories proposed. Re-examining the evidence from experiments on helping behaviour and the theories proposed to explain their results and comparing them to the economics models assists in identifying the weaknesses of certain approaches (Chapter 3). Again the problems associated with incorporation models in terms of dealing with context show up, as well as the informational attitudes' issue for both incorporation models and private benefit of action models. The psychology literature, however, allows the distinctions between the balance model and the responsible self-image model to become clearer but the literature on helping behaviour alone takes this aspect of the examination only so far; to take it further requires a consideration of the literature on harming behaviour (Chapter 4). The balance model's assumptions that feelings are not perfectly substitutable, and that there are feelings which constrain harm, provide a useful

approach to explaining the apparent need for justification for violence and other crimes.

The comparison of the models with the psychology literature assists greatly in reducing the number of options. The self-image and balance models, particularly because of the possibility for motivated self-manipulation, do appear stronger. This strength appears to be further amplified by their consideration within a responsibility formulation. Arguing for the importance of responsibility, or the existence of altruism, in shaping people's actions opens one to being asked question about why people would have such feelings. Providing an evolutionary argument for the existence of characteristics would strengthen the argument and may indeed be possible (Chapter 3). It is, however, not the focus of this work and will not be considered in great detail. The topic is not covered in detail here, as the concern is with what the characteristics that shape behaviour are, rather than where they came from. It should also be noted that, while providing evolutionary arguments can be useful, they are often given too much weight. It is at times implied that, if it can be shown that purely self-interested actors would win out over altruists, altruism is impossible. This is not the case, because traits need not be shown to be fitter than any other, only fitter than others that existed. This is why we still have humans of the walking variety even though wings would have been an advantage. It may be that humans have always had such traits with no competitor. It could also be argued that traits may come in packages, not that this is suggested here, but maybe conscience came with reason and the advantage of the latter outweighs any possible disadvantage of the former.

Arguing for the importance of responsibility also opens up the discussion of what shapes responsibility or what makes an individual feel that they are responsible. This is an area that will be considered here. It provides one of the important links between the theoretical aspects of the thesis and the practical interests.

1.5 Application to practical questions

Developing the two responsibility models provides a framework within which to examine the practical questions of who might help whom and how the 'who' and the context might interact. When using these models the question becomes: when do people feel a responsibility to help? And do they? If so, how do they manipulate this feeling? It may seem strange to use two models, particularly when their predictive conclusions are so similar. At times both models are used in the analysis of the practical question as the balance model is more controversial. Using the responsible self-image model means that the predictions are not tied to the balance model. The balance model does, however, have different implications (Chapter 8) and so is used throughout. It also provides a different way of approaching the question, which may be more informative, even if the predictions of behaviour are the same.

The literature on the setting of health priorities and on charitable giving provides some guidance on how the above practical questions might be addressed. The health priorities literature has, however, typically been concerned with local contexts, which make it quite different; while the charitable giving literature is limited by its marketing approach. Both literatures can be used to assist in the design of an experimental economic approach better suited to the questions at hand by suggesting factors which may be important, and techniques for getting at these (Chapter 5).

In order to examine the applicability of the responsibility models and at the same time investigate the practical questions, an experiment was designed and conducted. Experimental approaches are popular in economics to test theoretical assertions, such as are being made here, with regard to behaviour. They are, however, not typically used to collect data on practical questions, although they can be. The experiment examined how small differences in

appeals for need influence individual responses, involving real financial allocations, to that need. What's more the experiment examined how individuals felt about different cases of need, so as to link these feelings with donated amounts.

The results from the experiment support the earlier conclusions relating to the weaknesses of alternative models and the strength of the maintenance of a responsible self-image model and the balance model (Chapter 6). Small changes in case characteristics were at times associated with significant differences in donations and feelings associated with cases. It appears as if individuals' feelings and donations are associated in very particular ways. For example, respondents appeared ready to assume fault and decrease donations. More importantly, the results provide evidence that a pressure to help was felt and was, independently of sadness felt, related to donations. The game theory component of the experiment provided interesting results in terms of attitudes towards information. This evidence further contributes to the earlier work reviewed which suggests that attitudes towards information are central to efforts to distinguish between alternative models of behaviour.

1.6 Considering the social context

The argument to this point has dealt with individual responses and has not considered the collective weight of many responses, or the social context in which they occur. The theory discussions and the experiments focus on the individual. This is particularly true for the economics literature. The more detailed discussions of the psychology literature, which are to come, show that this individualistic focus was softened by considering the evidence and theory presented in this research, although only to a limited extent.

Individuals do not, however, operate in isolation, particularly when considering responses to distant others. They typically do not have contact with those

distant to them. They hear about their need through the media and respond through charities or their influence on policy. To begin to consider the implications of the discussions it is necessary to expand the conceptual framework to consider the social context (Chapter 7). Information on the needs of distant others has to reach individuals before they can respond to it. In general, the conduit for such information is the media. Before the individual receives the information it has first to be selected for coverage and framed for reporting. Alternative frames and levels of attention given to events may well influence the individual's response. The information may reach not only one individual but many; the collective response in terms of public opinion may well impact on policy. Locating the balance model, or the responsible self-image model, in a social context considering these factors, allows for discussions relating to real-world events.

An argument can be made that the nature of the HIV and AIDS epidemic led, in the early stages, to low levels of media coverage (Chapter 7). The response to this coverage was minimal, because of the ease with which the suffering could be interpreted as the fault of the person suffering. Given a motive to avoid a pressure to help, this would make sense. Combine this with an elite, at the local and international level, unwilling to accept the socio-economic causes and the required responses, and the backdrop for inaction is clear. When the problem became about treatment, the media frame changed and the portrayal of responsibility shifted. This, combined with an elite divided on the issue of treatment, may have created the space for public opinion, shaped by this new frame, to have influenced the response leading to what we see today. If this is the case, then it is a worrying situation. While making it about treatment may have helped in generating a response in this regard, it may now be considered that the job is done once treatment rates are high enough. With responsibility being considered met, it may become increasingly difficult to gain support for prevention and mitigation interventions. This is an example of the dangers of the welfare of the 'have-nots' being determined by the 'haves'

1.7 Thesis structure

An enormous amount from a wide variety of sources has been said about why and when people help. Much of this work considers when people should help. This is not the focus here and so is largely ignored, but this still leaves a great deal to cover and a necessity to focus.

The theoretical component of this thesis is economic in character, so the thesis starts with the same focus. Chapter 2 reviews the economics literature in broad terms. The review aims to identify and discuss the key groups of behavioural models relating to helping behaviour. It is argued that, in so doing, a gap for the responsibility formulation and more specifically the balance model is identified; it is here that both are formally developed. This is primarily a development chapter. Alternative models of behaviour are identified and new ones introduced. Their strengths and weaknesses are discussed, but the chapter concludes with a range of alternative models still in contention to be used in framing research on the practical questions of interest.

Economics, as has already been noted, often aims to simplify, but at times complexities are needed. In terms of explaining helping behaviour these can be found in abundance in the psychology literature. Chapter 3 seeks not to review the psychology literature, but to re-examine some of the major results and debates that occur within in it, so as to assist in trimming down the number of economic models considered.

While comparisons with the helping literature from psychology are useful, it is also surprisingly informative to examine some of the literature on negative behaviour. Chapter 4 discusses the economics models in the context of the literature on negative behaviour, particularly violence, as this has a direct impact on others' well-being. The discussions in this chapter provide strong

support to the framing of the behaviour in the form of internal constraints which are subject to manipulation, considering how excuses or justification seem so important to perpetrators of violence.

The above process of development, followed by attrition, leads to the identification of two models as potentially useful: the responsible self-image model and the balance model. On the basis of these models, it is possible to develop an approach to examining the practical questions surrounding responses to distant others. This was done, but in such a way that it also allowed for the testing of some of the theoretical assertions and so for contributing to the literature in this regard. Chapter 5 outlines the experimental methods used that allowed for the simultaneous investigation of the two sets of questions. In addition to detailing the method of data collection the analysis plan is also laid out.

Chapter 6 presents the results of the method outlined in Chapter 5. It highlights the key findings and discusses the implications for the theoretical debate and draws out the practical findings.

A criticism of economics, and to some extent psychology, relates to what some see as their too great a focus on individuals. For the questions being considered, and the models being suggested, the importance of values and information flows are clear and these may well be influenced by social factors. Furthermore, the importance in responses, as was mentioned, is considered not only in terms of individual actions but also in the formation of public opinion. For these reasons, Chapter 7 attempts to place the individualistic models of behaviour discussed in the previous chapters into a social context. Based on this expanded model, the discussions in this chapter are able to return to the issue of HIV and AIDS and consider the possible reasons for the slow response.

The theory developed in the course of the discussions of helping distant others has potential applications beyond this subject area. Chapter 8 seeks to apply the theory to other relevant areas, such as violence and risky sexual behaviour. These discussions are intended to suggest areas of further research rather than to draw conclusions.

Throughout the thesis the differences between the balance model and the responsible self-image model are discussed. It is argued that, in terms of predictions, they are similar if not the same. It is, however, also argued that they differ in terms of their implications. This discussion is touched on in earlier chapters but is discussed in full in Chapter 8, so as to keep it separate from the predictive conclusions made previously.

Chapter 9 concludes the thesis by summarising the argument made and its implications.

Chapter 2: Economic theories of helping

2.1 Introduction

Explaining altruism within the neoclassical economics framework is like eating sushi with a fork. It can be done, it's not difficult; it might even be efficient but it still feels and, to most people, looks wrong. There have been numerous efforts to bring altruism, and moral action in general, within the standard utility maximisation framework, with the argument that at some level every action is self-interested and it is just a matter of finding that level. There have also been many efforts to depart from at least the mono-utility construction of the framework. It is interesting to note the language used by authors undertaking these latter efforts. They all struggle to express the importance of the difference in the type of motivation that underlies particular groups of actions and how uncomfortable it is to think of them as a form of want. Motivations associated with 'should', 'ought' and 'must' are argued to be of a fundamentally different character from 'wants' and 'likes'. The idea of thinking of heroes and martyrs as masochists does not sit well with many. It is interesting to note that the struggle with such differences did not take long to appear. While Jeremy Bentham wrote of pleasure and pain within a mono-utility framework, his student John Stuart Mill already started to identify higher and lower pleasures (Mill, 2002 [1863]).

The literature in this field of economics is vast. This chapter attempts to highlight the main trends and the limitations of different approaches. A current trend, which appears to be very promising, will be discussed and extended for use in the rest of this thesis. This trend revolves around the preference of individuals for maintaining a positive self-image. This conception opens the way for attitudes to information and self-deception that appear to align more closely to reality. That said, it is still like eating sushi with a fork.

This chapter will continue with the development, started in the introduction, of an alternative way of thinking about altruistic and moral behaviour. This framework essentially argues that individuals do have different types of motives, some physically and some cognitively generated. While these motives are not reducible, for the most part they are tradeable, the exception being when they are at close to their extremes; individuals seek balance across these motivations. While maintaining balance may be considered self-interested, not all of the aspects of the system can be. It is not suggested here that this approach necessarily predicts behaviour better than, or even differently from, models developed within a more traditional framework. What is suggested is that the use of the framework feels, if not right, a whole lot less wrong. Moreover, while its predictions may be similar, their implications may not be. This is an issue which will be touched on in this chapter and returned to in more detail in Chapter 8.

The outcome of this chapter will be the beginning of an argument for the predictive power in regard to altruism of two models of behaviour that will be continued throughout the thesis: one based on maintaining a positive self-image and the other on the alternative construction of the balance model. These models can both be placed in what is argued to be a useful formulation: the responsibility formulation, which is introduced in this chapter and discussed in detail in the following chapters. Once placed within this formulation, both models predict essentially the same behaviour as each other - though differently from other models - but the latter is argued to be more realistic in its assumptions. The former construction is included as it is less controversial and will be used to show that the predictive conclusions of this thesis do not require the acceptance of the latter.

2.2 The economics of altruism and moral behaviour

There are many examples of people acting in seemingly altruistic ways in everyday life. Someone falls in the street and a passer-by helps them to their feet with apparently no benefit to themselves, other than seeing someone else better off. The ways in which economists have attempted to explain such acts through behavioural models have differed in many respects, but arguably have some common strengths and weaknesses.

When examining the contribution of others, particularly if it is done critically, it is important to note what their purpose was for making their assumptions and developing their theories. For example, Sen points out that Edgeworth, in his 'Mathematical Psychics', while basing his results on the assumption that individual actors were motivated by self-interest, did so not because he believed that people were indeed so motivated but rather to show what implications such motivations would have¹ (1977). It is difficult, therefore, to criticise Edgeworth's assumptions on the basis of their ability to explain everyday behaviour; it would be more proper to discuss when the assumptions, and therefore the conclusions, are appropriate.

Similarly, Becker argues that critiques of behavioural models should be concerned with what implications the critique would have for theory (1962). He stresses that individual models are used as the basis for the development of theory. So, while there might be arguments about the appropriateness of alternative models of behaviour, what is more important is the implication for the theory. By way of illustration, Becker shows how models of seemingly irrational behaviour at the household level still lead to the same results at the market level. Specifically, he examines how impulsive households, or those characterised by inertia, will respond to price changes in such a way as to result

¹ Although Edgeworth did argue that in some cases they were, namely war and contract (Edgeworth, 1881).

in a negatively-sloped demand curve. In this example, Becker argues that it shows that an undue concentration on unit irrationality can lead to misplaced conclusions regarding market irrationality. In a similar vein, he notes that simply because markets may appear to act rationally is not grounds for assuming that individual units also do. In a related way, the following discussion is concerned not only with the realism of model assumptions, but also with how their variation alters theoretical predictions of behaviour and how supportable these predictions are.

Behavioural models, including altruistic or moral actions, have been developed in response to the rather limited explanatory power of the traditional economic rational actor in certain areas of study. These include charitable donations, volunteerism, intra-family transfers and moral behaviour. To understand how the models discussed here deviate from the traditional approach, it is important first to understand the point of departure in some detail.

The traditional economic actor is seen as a self-interested individual who seeks to maximise his or her utility, as defined by their preferences, through his or her choices. This definition of a 'rational' actor is typically treated as a whole. There are, however, a number of distinct aspects of the actor's character that are worth separating, as the models that will be discussed shortly vary in the aspects they adjust.

Sen (1985) identifies three aspects of the traditional economic actor:

Self-centred welfare: An actor's welfare is derived only from his or her own consumption.

Self-welfare goal: The only goal of an actor is to maximise his or her welfare.

Self-goal choice: All acts of choice made by an actor are made in furtherance of his or her goal.

Sen argues that these three aspects are independent of one another and that the violation of one does not necessitate the violation of the others. The majority of models discussed in this chapter seek to alter one or more of the above aspects to address their failure to deal with altruism or moral actions. They, therefore, aim to adjust the traditional approach rather than reformulate it.

Before moving on to an examination of variations it is worth noting some explanations provided within the existing framework. One explanation of altruism that requires no deviation from the traditional model is reciprocal altruism. Reciprocal altruism, where help is provided based on the belief that the recipient of the help will in some way repay the helper at a future date, is not really altruism and such models are inappropriate and typically provide no basis from which to explain single-spot exchanges – once-off events with no hope of reciprocity (Khalil, 2004a).

There are, however, more complex formulations of reciprocal altruism that may fare better, such as Hammond's (1975) suggestion that some seemingly charitable behaviour could be explained by cooperative egoism, a social contract of a kind linking benefits today to benefits in the future. But an argument such as this still fails to suggest how the contract operates within the individual and why an individual would not 'cheat' in once-off encounters.

A similar social contract argument, which similarly fails but it is interesting to note as it expands on how it might operate at the individual level, is contained in Buchanan (1975). He uses what he calls The Samaritan's Dilemma: a game situation where to influence the behaviour of a potential parasite the Samaritan must act strategically. The Samaritan must risk personal loss in order to shift the opponent's response. He uses a family example: the mother does not like spanking the child, but must do so to lead to the best outcome of no spanking and good behaviour, this as opposed to not spanking in the first instance, which for that instant would be the utility maximising action for the mother. He uses

this construction to argue for the benefits of rules for personal behaviour for maximising long-term utility. Buchanan, in the first instance, sets the dilemma up in the form of repeated games against the same opponent. He notes, however, that the conclusions hold in single encounters if the Samaritan's behaviour towards one potential parasite influences the behaviour of others they are yet to encounter. He further extends this to a public dilemma where the Samaritan is unlikely to encounter a parasite again but their actions may influence the behaviour of other parasites towards other Samaritans. In this case he uses a hijacking situation as an example: the Captain of a hijacked plane may do better to cooperate with hijackers, but does not, as this benefits all flight crews through its influence on potential hijackers. He does, however, argue that these are voluntary individual rules, which makes it difficult to explain great losses that have occurred as a result of heroic behaviour. He also argues that such behaviour is becoming less common as the population size increases and the ability to influence other parasites decreases. This leads, he argues, to the formalisation of rules. These models of self-interested help are interesting but the focus of this thesis on the helping of distant others makes explanation based on reciprocity or influence rather implausible.

Another argument in favour of self-interested moral behaviour has been provided powerfully by Frank (1987, 1988). He presents an interesting model that attempts to provide a framework capable of explaining seemingly irrational behaviour within a utility-maximising model. The paper concentrates on individuals who act honestly even when they would increase their pay-off by cheating with no fear of reprisal. To explain this, Frank takes an evolutionary approach, asking what kind of tastes might lead to an evolutionary advantage. He argues that if honesty can be signalled, thus identifying the honest actors, they will be advantaged as they would be disproportionately selected for profitable cooperative ventures. Harrington, when commenting on this, argues that, unless the signalling is perfect, the conclusions will be weakened substantially and that honest traits will only be stable in a population if a

sufficient proportion of the population began with them (1989). Frank defends his model by arguing that beliefs in signalling exist even if perfect signalling does not (1989). While such a model is interesting, it raises two problems. The case for evolutionary advantage for the honest can be made, but it is difficult to see how it could be done for the altruist (Khalil, 2004a). It is also not clear how the honesty or altruist motivation would operate for the individual – which is the main interest here.

Suggestions have been made as to how altruism might operate within an individual. These suggestions generally, although not always, involve some deviation from the traditional model. It is not possible here to review all proposed deviations from the traditional approach, but rather what follows focuses on some major contributions, which are grouped according to their common characteristics. Significant among these contributions are the various works on the topic by Becker, such as his Theory of Social Interaction (1974).

Becker's Theory of Social Interaction is based primarily on the concept of social income. He defines social income as the sum of the person's own income and the monetary value to them of the relevant characteristics of others, defined as the individual's social environment. Becker argues that standard economic theory is based on the assumption that the individual's social environment is beyond their control. Therefore, while utility may be in part determined by the environment, the individual must treat this as exogenous and maximise their utility function for a given environment. As an alternative, Becker presents a model where the individual can allocate income towards altering their environment. From this starting point he is able to develop his theory and apply it to a variety of situations. Becker's focus in his application of his theory is on the family, although he mentions charity, hatred and envy. The welfare of other family members, or others more generally, enters an individual's utility function through the social environment, implying essentially the inclusion of others' utility in the individual's utility function.

Becker is widely, and rightly, attributed with the popularisation of what is called here the incorporation approach. He was, however, by no means the first to suggest it. Adam Smith's comment on the pleasure of seeing others happy being the only reward for certain actions is often quoted. What is less often noted is Edgeworth's contribution (Collard, 1975). Collard presents an outline of Edgeworth's arguments on altruistic behaviour that involves the inclusion of others' welfare with some adjustment. He quotes Edgeworth as writing: 'We must modify the utilitarian integral by multiplying each pleasure, except the pleasures of the agent himself, by a fraction – a factor doubtless diminishing with what may be called the social distance between the individual agent and those of whose pleasures he accounts.' (in Collard, 1975: 355). This approach leads to the inclusion of the coefficient of effective sympathy in the analysis as well as the possibility for pure and impure models of behaviour. Quoting again, 'as the coefficients of sympathy increase, utilitarianism becomes more pure.' (in Collard, 1975: 355). Collard notes that the benevolence considered by Edgeworth is not of a paternalistic nature. That is, the weight is attached to the utility of others regardless from which goods it is derived. This is a useful extension of the incorporation approach and suggests a way to account for differences in the value of different others. What becomes difficult is that, to account for behaviour regarding distant others, it becomes clear that there is a need for the coefficient to vary by context. This necessity will be returned to numerous times, as failure to provide a theory for how the coefficient may vary dramatically weakens the approach.

Using Sen's division outlined above, incorporation theories differ from the traditional approach in their behavioural assumptions, in terms of the first aspect, that of self-centred welfare. By including the welfare of others, such as family, in an individual's utility function by way of the social environment, Becker shows how a deviation from the traditional behavioural model can be used.

Individuals, in Becker's theory, however, still conform to the second two aspects of self-welfare goal and self-goal choice.

Becker's work is a classic example of one of the main areas in which these approaches have been used, namely questions relating to intra-family distributions. As mentioned above, he also applied his model to another area of interest, one which has attracted considerable attention in the literature – that is, charitable giving.

Charitable giving has been dealt with again by varying the assumption of self-centred welfare. Individuals are seen to value the welfare of others: that is, the welfare of others makes its way into their welfare function. Charities that improve the welfare of others, therefore improve the welfare of individuals who value the welfare of those who benefit, even if they themselves do not benefit directly. The derivation of welfare from the improvement in the welfare of another does not exclude others from deriving similar welfare improvements. Indeed, they too can value the resultant welfare improvement of others; thus charities in this formulation can be seen as a public good. If it is the welfare of others that is being valued, then it is the total allocation to the charity that is valued not an individual's contribution. This has been seen in the literature as an example of pure altruism. Khalil comments on such models and argues that, if this is the case, the altruist would thank the poor for being so and would welcome disasters, as this gives the altruist the welcome opportunity to help (2004a). This is not strictly true, as the disaster and the poverty would bring the welfare of the potential altruist down, but this would require the altruist to have always had the welfare of others in their welfare function and this does not seem to be the case. There is also the issue of whether one individual can know the welfare implications of their actions for another (Folbre and Goodin, 2004).

The problem of continuous inclusion is an important issue for models and will be returned to, but it is the strong implications regarding crowding out implied by theories of pure altruism and the debates on these that have prompted the development of alternative models.

Warr shows, with a simple model, that if the utility of one individual enters the utility function of another, then redistributive taxes will have no net impact (1982). This results from the donor compensating 'dollar for dollar', as public fiscal redistribution is seen as a substitute for private charitable donations.

This argument has been repeated in a number of similar forms and is presented as the basis for assertions that public funding crowds out private charitable donations. Counters have been presented to this conclusion, notably by Andreoni (1989; 1990); such counters require an alternative model.

Andreoni expands on the behavioural models that include altruism to include what he refers to as 'warm glow' aspects (1989) in a model of impure altruism (1990). Essentially, he suggests that, while individuals may demand more of the public good altruism, they also derive some utility from the act of giving in and of itself, namely, a 'warm glow'. He shows that, if such a private benefit exists, then involuntary giving is not a perfect substitute for voluntary giving. Andreoni argues that the evidence supports this model, as opposed to pure altruism, as crowding out is not seen to be complete. The model of impure altruism allows for the special cases of pure altruism and the purely egotistic, the former's utility function including only the total provision and the latter only the individual's own donation. Impure altruism occurs between these two cases where the individual derives benefit/utility from both their own contribution and the total provision. The individual is then argued to maximise their utility function subject to wealth constraints and the link between individual giving and the total given.

Models such as Andreoni's impure altruism still vary from the traditional model only in terms of self-centred welfare. In actuality, one of the extreme cases of Andreoni's model, when the only gain is from the act of giving itself, could be argued to be no different from the traditional model.

Andreoni's model is one example of a group of models, termed here private benefit of action models, that in some way includes a value for the act of giving itself. For Andreoni, the value stems from a 'warm glow' the giver feels; others argue that value is derived from social recognition. These models, however, have been argued by some to miss the point, because they argue that the action is motivated by some benefit to self, but fail to identify what that benefit is. Andreoni's 'warm glow' is basically the by-product of pride, without suggesting why the pride should be there (Khalil, 2004a). Khalil argues that the same is true of arguments that the 'warm glow' comes from adherence to social norms through socialisation. The socialisation models are argued to beg the question. Suggesting that actors conduct themselves to gain approval, or in line with institutional arrangements, fails to consider how such actions were deemed approvable in the first place. This criticism may be a little harsh. Konow (2005), for example, suggests conditional altruism shaped by social norms. While it does indeed beg the question at a societal level, it does offer a suggestion as to what shapes the individual's behaviour. The criticism may also be a little misplaced. Andreoni set out to, and does, show that, if individuals gain some private benefit from giving, crowding out will be incomplete. It was not his primary aim to explain the source of such private benefit. It is, however, important to note the problems with using such a conception for reasons other than its original one. Other efforts have sought to expand on the explanations of where the glow comes from.

In a similar vein to models that argue that actors gain some benefit from acts of giving are those that claim that individuals value certain behaviours or maintaining beliefs about themselves. Essentially, they argue that individuals

value behaviours, such as honesty and altruism, or they value an image that they have of themselves and will take steps to preserve that image. How the valuing of a positive self-image affects behaviour has, however, been included in two distinct ways. It is either proposed that individuals derive utility directly from it, which is the same as valuing the behaviour itself (Johansson-Stenman and Svedsater, 2003), or that it acts as a constraint on their behaviour (Dana et al, 2006; Dana et al, 2007). The first of these approaches differs little from the 'warm glow' or other similar arguments, except in providing an argument as to the source of the glow; as such it is another example of a private benefit of action model; the second, however, is something of a departure. Deriving utility from acting in accordance with self-image is essentially modelling the value of self-image as a flow variable; that is, valuing each act of behaviour as a good in itself. Including self-image as a constraint is more like modelling it as a stock that the individual wants, or has, to maintain. This is an example of another class of models, which are referred to here as private benefit of maintenance models.

An example of a private benefit of maintenance model is Dana et al's (2007) presentation of a model of constrained behaviour, where the constraint is a 'want' (see also Minkler, 1999 for another example of 'want-based' constraints). The issue of constraints having to be adhered to will be returned to in due course, as when the constraint turns from being 'want-based' to 'have-based' (from soft to hard constraint) the models fall into a different group. Dana et al reported on the results of their experiments in which they eliminated the direct link between individual action and harm to others, although the link could easily be re-established by the participants if they wished. In doing so, they observed a significant increase in self-interested behaviour. They propose a model to explain this result, which suggests that individuals do not want to take actions that run contrary to their beliefs about themselves. They therefore will avoid information that may constrain their self-interested actions because it negatively affects their beliefs about themselves. Their actions, therefore, are not

influenced by a preference for the welfare of others or for a preference for certain kinds of acts, but rather by a constraint on their behaviour acting through a desire to maintain a certain stock of self-image. If they valued the behavioural act, they would want to know if they were harming others so that they could avoid such an act and gain utility from so doing.

Models such as the above, where the constraint is adhered to because maintaining the stock is valued, rather than moving further away from the traditional model actually move towards it. Altruistic behaviour is self-interested, it is undertaken to protect a self-image which is obviously only of direct benefit to the self. While this is so, the formulation does provide a framework for some interesting explanations and predictions. Valuing a stock of self-image, rather than a flow, provides a useful framework within which to understand the importance of context, self-deception and the avoidance of information. In such models, if an individual can maintain their stock at no cost, they prefer this to maintaining it at a cost. If, for example, they value a self-image of fairness, they do not want information on whether their actions are fair or not. If they are shown not to be fair, they may end up changing their actions to maintain their stock of self-image; but, if they never knew, they could tell themselves that their actions are fair and they have no need to change. Other examples of this type of formulation show more of the possibilities.

Within the same class of models, Khalil, for example, proposes what he calls a quantum model, which he argues addresses the existence of integrity, shame and self-rationalisation (2004b). The model essentially argues that individuals wish to maintain self-integrity and they do this by adhering to binding commitments that they have made to themselves. The breach of such a commitment results in shame, which can be damaging to the individual. The shame is, however, context dependent not simply action dependent. If the individual finds themselves in situations where maintenance of a binding commitment would see them plunged into an emergency state, they can breach

the commitment without feeling shame. This he suggests shows how his model differentiates shame from integrity utility, which would be an example of a flow utility resulting from the consumption of a positive self-image. He argues that, in non-emergency situations, integrity operates as a capital stock linked to identity; as such it is not separable at the margins and can only be traded in bundles.

Another example is Brekke et al (2003) who present a model of moral motivation. Their model is based on the assumption that 'people want to think of themselves as socially responsible' (Brekke, 2003: 1969). To determine the degree to which alternatives fulfil this preference, comparison is made against an endogenous ideal of moral behaviour. Thereafter they trade off acting responsibly against other demands. This, they argue, provides a framework that explains how information and the environment shape moral behaviour, by changing what is considered to be moral.

While private benefit of maintenance models are basically a regression to the traditional model, they do seem to have considerable explanatory power. Their ability to consider the possibility of self-rationalisation, information avoidance and context makes them a useful tool when explaining and predicting behaviour. Arguably, this is the fork with which neoclassical economics can eat sushi. To appreciate where the discomfort comes in, it is useful to examine somewhat different formulations of this type of model.

Take for example Konow (2000) who suggests that individuals have a desire to pursue self-interest and to act fairly and when they conflict they result in cognitive dissonance placing a tension on the individual. To reduce this tension, they either forgo self-interest or engage in self-deception. Similarly, Akerlof and Dickens (1982) developed a behavioural model based on cognitive dissonance. Their argument is based on three propositions: firstly, individuals have preferences not only over situations but over beliefs about those

situations; secondly, individuals to some extent have control over their beliefs; finally, once a belief is selected it persists. They cite a variety of psychological studies as evidence of these propositions, generally arguing that people prefer to believe themselves to be smart people and will interpret information to reinforce this view, but at the cost of cognitive dissonance. While these models are very similar to those just discussed, they are interesting as they suggest that there exists a tension between different types of motivation. It is a small step from here to arguing that individuals have more than one type of utility and associated preferences.

A number of authors have been arguing for some time that individuals are capable of having more than one set of preferences. Buchanan, for example, while discussing the differences between market and voting choice, noted the possibility that the individual will act according to a different preference scale when voting (1954). This possibility may come about as the individual responds to a sense of participation. Buchanan does note, however, that it is difficult to determine if this occurs or if it is alignment with self-interest or failure to consider or realise the individual costs.

One of the early major contributions to this debate was made by Harsanyi in his discussion of social welfare and individual utility functions where he argues that, as a result of increased awareness of the important role played by external economies and diseconomies of consumption, the concept of individual utility has changed (1955). He suggests that, given this importance, individual utility is taken to depend no longer only on the individual's conditions but also on the conditions of all individuals in the community. He stresses, however, that this should not be confused with an individual's social welfare function. Harsanyi proposes two sets of preferences, one being subjective preferences on which an individual's utility function is based and the second being ethical preferences on which their social welfare function is based. Ethical preferences are not considered as preferences in the full sense, but rather what social state an

individual would prefer if they forced a 'special impartial and impersonal attitude upon' (315: 1955) themselves. Ethical preferences in this discussion, however, are not offered by Harsanyi as motivations for individual action. As the interest here is in explaining individual altruistic behaviour, the focus should possibly be on subjective preferences, which, as mentioned, are included in the individual's utility function as preferences valuing all other members of the community, much as they are in incorporation models.

Harsanyi makes an important note that his formulation differs from Arrow's (1951) distinction between tastes and values. Arrow, he explains, views tastes as the basis for individuals' ordering of social situations in terms of their impact on personal consumption, whereas values are considered in regard to economies and diseconomies of consumption, as well as ethical considerations. This division, Harsanyi argues, does not explain how an individual might accept a social welfare function that conflicts with their values. According to his distinction, values form part of subjective preferences, and what is required is a different type of preferences to evaluate social welfare, for which he proposes his ethical preferences. Individuals will, however, be affected by the social context and so their ethical preferences may well affect their own welfare. If this line is taken, then the individual essentially has two utility functions affecting choice. This idea is taken up and argued for in the seminal work of Etzioni (1986).

Etzioni provides a very powerful piece of work gathering together insights from a range of sources. He makes an argument for at least two sources of utility, suggesting that the two, morality and pleasure, are distinct and irreducible. The argument is very strong as far as the existence of the distinct utilities is concerned but weak on how they operate collectively to shape decisions and this has led to criticisms of the approach (such as Khalil, 2004a).

The argument made by Etzioni begins by taking aim at the mono-utility concept. He argues that three interpretations of utility have developed and that defenders of the concept rotate their use. These three comprise, the pleasure utility, the interdependent utility and the grand X utility; his discussion of each and the differences between them provides some useful insights.

The pleasure utility concept is argued to be what most closely represents the original formulation of the concept. It is essentially a hedonistic concept where utility, happiness, pleasure and satisfaction can be used interchangeably. A model based on individual drive to maximise such utility, Etzioni argues, clearly struggles to explain a range of behaviours, including altruism, as there is no consideration for others.

The interdependent utility concept is invoked when the pleasure utility concept is challenged. This version includes such arguments as individuals valuing the welfare of other individuals and so holds to the pleasure motivation, but includes love for others. Such a conception does not allow for self-sacrifice and suggests that saints are masochists. Etzioni argues that this ignores the reality that people do things because they feel they ought to, not just because it pleases them, and that it even goes counter to Smith's differentiation between love for self and love for others.

The final conception of mono-utility that Etzioni discusses is what he calls the grand X. He suggests that this conception is totally divorced from the origins of the utility concept. Utility is not consumption or pleasure; it is just the value attached to different outcomes. Individuals choose and this reveals what they prefer; it is not seen as necessary to know why they prefer it, just that they do. This leaves utility as an empty concept. Any effort to better the situation requires the identification of how rankings occur and the problems of the above two conceptions are reintroduced.

Having addressed what he views as failings in the mono-utility concept, Etzioni takes a rather odd detour to discuss what constitutes a moral act. While he presents an interesting discussion, arguably it is unnecessary for the point he is trying to make. What he argues is that pleasure is measured with one type of utility and morality with another. The important argument relates to the existence of a moral utility not its correctness. That an individual feels a motivation 'should' is important, not whether the 'should' can indeed be considered as appropriate moral action.

Etzioni cites a variety of evidence in support of his argument that moral utility is distinct from pleasure utility and that the two cannot be reduced to one and why this difference is important. In defence of the division, he essentially makes an appeal to the difference in the feel of motivations driving actions: how feeling one ought to, or has a responsibility to, is different from one wanting to. Into this discussion he brings the feelings of shame and guilt associated with moral actions and not with others, as well as the possibility of conflicting motivations as characterised by dissonance. He also introduces an interesting dynamic element to his discussion, suggestive of the stock previously discussed, by citing how past transgressions influence future corrections, suggesting a debt to a particular source, which cannot be paid by a gain to an alternative source.

While Etzioni provides a compelling argument for a separate moral utility, there is a difficulty in that he does not propose a way for the two utilities to combine in decision making. If they combine to some higher level of utility, the difference in the construction is seemingly lost, but if they do not it is difficult to explain choice.

In response to the difficulties noted in deriving a decision theory from the above, Etzioni later argues that most decisions are made on the basis of emotions and values (1988). He argues that the bulk of decision making is based on normative-affective considerations as opposed to logical-empirical

considerations that characterise the traditional economic model. While most would accept a role for normative-affective considerations in the setting of goals, his thesis goes much further presenting them as the central factor in deciding means.

Etzioni argues that normative-affective considerations result in decisions falling into three zones. There is the exclusion zone, where as a result of these considerations whole series of options are seen as non-options; they cannot even be considered. Such a zone is characterised by what feels right. Then there exists what Etzioni refers to as the 'infusing zone'. This is when normative-affective considerations are combined with logical-empirical considerations. Such a situation results, he argues, in the disruption of the logical empirical approach, as individuals jump to conclusions or resist information. Finally, there is the legitimated indifference zone where logical empirical considerations are the primary concern; nevertheless, he stresses that the determination of what falls into the zone is a result of normative-affective considerations. For Etzioni, the relationship between emotions and reason and between values and reason are much the same. While again a useful and insightful contribution, it still avoids the question of how choices are made in terms of the evaluation of alternatives involving the two utilities. He does consider the process of choice in terms of identifying the importance of emotions, but how to weigh up between one emotion and another is not discussed, so the problem remains.

Contributions such as Etzioni's have been characterised as models of multiple-selves. These models have taken various forms, some taking the multiple-self approach literally. Lynne provides a model explicitly based on a multiple-self (1999). In what he calls 'meta-economics', he presents a model with three characters: the child interested in 'goods and wants' pursuing the 'I-interest', a parent pursuing the 'we-interest' and a rational, reasoning, mediating adult who finds the balance. He argues that what is needed is not a decision theory but a

command theory of how the mediating adult controls the, at times, competing demands.

Other conceptions of the multiple-self line have taken the form of introducing a 'have' or a hard constraint on behaviour. This differs from the 'want' or soft constraint mentioned previously, as individuals are seen to have to abide by the constraint. The issue of who constrains them if not themselves and, if it is themselves, how can it not be a 'want', is why they remain part of the multiple-self group of models.

While constraint models appear to have become more popular in recent times, they have been identified in the economics literature as possible formulations for some time. Arrow, for example, when reviewing a study by Richard Titmuss of blood donation, identified three motives for giving as part of a reformulation of Titmuss' argument into utility language (1972). These three are: firstly, the positive inclusion of others' welfare in individual welfare functions; secondly, the positive inclusion of others' utility and the individual's contribution to it; and finally that individuals are egotistic, but there is an implicit social contract that creates a duty to enhance the welfare of all, essentially along the lines of Kant's categorical imperative. The first of the motives matches closely to the incorporation models, the second to the private benefit of action models and the last is an example of a constraint which link to the multiple-self group.

The Kantian formulation of constraints has remained popular. White, for example, expands on his earlier work (2004) to present a Kantian model (2006). He argues that efforts to account for moral behaviour through a taste for morality miss the distinctive nature of the choice. While conceptions of multiple utilities remedy this, they do so by adding complexity, which he deems unnecessary, although by adding a constraint he is essentially joining the multiple-utility camp. The model he proposes differentiates between two sets of duties, perfect and imperfect. Perfect duties provide clear rules and allow no

flexibility; imperfect duties prescribe ends but there is some flexibility in their pursuit. He presents his model first under assumptions of perfect rationality/morality. In this model, perfect duties enter as constraints, whereas imperfect duties are considered as preferences. To allow for this conception, he argues the need to use the empty definition of utility with no psychological associations, simply defining preferences as determining ordering – the grand X as Etzioni refers to it. His use of the grand X is an effort to avoid the multiple-utility label. He argues that actions called moral under conceptions such as Etzioni's, where the pursuit of moral utility is motivated by satisfaction, are in the Kantian sense not moral. In White's model, perfect duties rule out certain options and the agent chooses between the remaining alternatives according to inclination or imperfect duty.

The assumptions of the model are then weakened to the contingent rationality/morality model. Here the constraints of perfect duties are not always applied, as a result of weakness of character or human failing. White argues that the constraint is applied on a probabilistic basis. The probability that the constraint will be applied is determined by the strength of character of the actor: the more moral the actor the higher the probability of application.

White uses a number of examples to illustrate his point; one interesting argument relates to crime. He argues that crime, in so far as it is immoral, is more likely to be committed by the weak of character. He also notes that, in the Kantian model, a rational agent's true goal is to be moral; immorality is a result of weakness of will.

White's model provides a useful example of the introduction of a constraint into the decision-making process. He does not, however, expand on the implications for behaviour of having such a constraint.

Rabin has also presented a constraint-based model and does expand on the implications: he argues that, if utility were derived from the welfare of others, that is, if individuals have a preference for the welfare of others, then they would have a motivation to obtain information relating to changes in others' welfare resulting from their actions (1995). He argues that this is not the case and asserts instead, in relation to moral behaviour in general, that individuals have internal moral constraints. This means that, unlike if they had a preference for moral behaviour (such as altruism), they self-servingly avoid and select information, thereby allowing them to weaken the constraints they have on their behaviour. These are the same predictions as 'want-based' soft commitments mentioned above, but with some important differences in terms of implications.

It has already been noted that deriving benefit from a positive self-image and maintaining a self-image as a constraint on behaviour are two very different assumptions. The former, as mentioned, results in a similar model to those that include some benefit associated with acts of giving. That is, self-image enters the utility function of an actor as with other goods and the predictions are the same as other private benefit of action models. The latter, as mentioned, has quite different predictions, but importantly also implications depending on whether the constraint is considered hard or soft.

Being within a constraint model framework does not, in and of itself, require that the individual may not be self-centred or still have self-welfare as their goal. It is, however, difficult to comprehend the impact on self-goal choice. Typically a constraint would not be seen as conflicting with self-goal choice; but typically constraints are not internal. What needs to be considered is whether the individual adheres to the constraint because they value the feeling they get when they do, or because the guilt resulting from not adhering is being avoided, or because they have no option but to adhere. If one or both of the first two is the basis for adherence, then it could simply be argued to form another input into the utility function, while still conforming to self-goal choice; this is where

the 'want-based' self-image maintenance models mentioned previously fall. When adherence to the constraint is incorporated into the utility function, the individual could still be argued to be welfare maximising in their choice if they adhere to the constraint, even if this results in a lower consumption level than was available. The loss from adherence would be outweighed by the benefit of maintaining a positive self-image, of feeling that they are complying with their moral constraints, or the value of avoiding the guilt of non-compliance.

If, however, the constraint, while internal, is not voluntarily adhered to because of its value, self-goal choice is constrained. This would mean that the rejection of a seemingly available option does not mean that it would generate lower utility for the chooser than the alternative selected, but maybe that the chooser is constrained from that option. This presents problems for the revealed preference approach. The possibility of counter-preferential choice was similarly raised, although from a slightly different angle, by Sen (1977).

Sen outlined the concepts of commitment and sympathy, the latter being considered as an externality, as actions motivated by such can still be seen as furthering a self-welfare goal along the same lines as 'warm glow'. Commitment, on the other hand, was defined as the situation where a choice that yields a lower level of personal welfare was possible. Arguably, individuals are at times motivated by commitment and their choices do not simply reflect the pursuance of a self-interested goal; this is the basis of the multiple-self models. The traditional economic behavioural model is clearly not able to account for such choices, although the counter argument is that they need not, as they do not exist.

Incorporation models and private benefit of action or maintenance can similarly not account for commitment. It could, however, be argued that, within such models, choices based on commitment need not be counter preferential: that the welfare benefits for others or the good feelings associated with the action or

the avoidance of guilt and remorse are such that they outweigh the apparent loss in individual welfare and that such choices are still self-goal, self-welfare maximising, but not always self-centred. Such an argument is, it seems, uncomfortable for many. It would suggest that the motivations for commitment simply generate utility like any other consumption. This would mean that, if enough compensation were offered, an individual could be persuaded to alter choices based on commitment. While it may be necessary to think of some rather extreme examples to argue that this is not the case, the possibility that the model may fail at the extremes is suggestive of problems at the centre.

Commitment can be addressed with multiple-self models, such as those based on hard constraints. Individuals may be subject to an internal moral constraint and are therefore 'left with no option' but the choice they make. While the constraint approach does appear, at least in regard to commitment, to have some advantages over other adaptations of the traditional approach, it is rather inflexible. Further, it does not address the possible interactions between self-interest and moral behaviour well. It fails to consider how moral values may still play a role in decision making, even when deciding between options that fall within the constraint, without them simply becoming conventional preferences. The inflexibility of the model will be returned to shortly. It is important first to consider the importance of commitment, because if it does exist in the way Sen described and is important, then the traditional model has no place, unless commitment can be incorporated in some way.

The commitment concept has been the focus of much debate. An examination of recent contributions gives an indication of the difficulty in agreeing on the issue. Take Khalil, for example, who takes issue with the notion of commitment (1999). He argues that Sen's conception is simply another version of a conflicting-self model. Khalil agrees that commitments are an issue and that dealing with them as normal preferences is problematic, but he argues that they can be dealt with within a unitary-self model. The model that he proposes

incorporates two kinds of preferences: non-binding and binding. Non-binding commitments are associated with outcomes that cannot be achieved with certainty, or rather that may be beyond the control of the individual, such as to be the best at something. These, he argues, stem from a desire for self-respect. Binding commitments, such as a commitment to be honest, on the other hand, are within the control of the individual and, in his framework, stem from self-integrity. Both of these types of commitment are considered in the model as voluntary. It is further argued that, while non-binding commitments are essentially substantive preferences, binding commitments are symbolic. Non-binding commitments are therefore part of substantive preferences and, as such, cannot be counter to them. He argues that the consideration of symbolic preferences in single spot encounters can lead to choices that are not in line with substantive preferences, giving the appearance of commitment. This is another example of a private benefit of maintenance of the 'want' constraint variety similar to other self-image models with soft constraints, which have already been noted to be powerful formulations of the traditional model. Even with such constructions, there are still those who argue that they somehow miss the nature of acts motivated by commitment by casting them as a form of self-interest.

Recently Sen has again argued that commitment is important (2005). In making his case, he suggests that there are two reasons for examining rationality. The first seeks to define what is rational and the second seeks to predict, based on a concept of rationality, observed behaviour. While the first can be done for its own sake, without reference to the second, the second is based on the first; although in modern applications this point has often been overlooked.

Dealing first with the importance of commitment in the definition of rationality, Sen starts by noting rational choice theorists' focus on rationality as the pursuit of self-interest. He goes on to note the variety of ways in which the self can be incorporated into choice and cites his earlier work distinguishing between self-

centred welfare, self-welfare goal, and self-goal choice, which has been discussed above. In Sen's own consideration of the concepts of sympathy and commitment, he argues that they affect different aspects of the rational choice model, with sympathy affecting self-centred welfare and commitment the self-welfare goal, since Sen argues that commitment can involve taking decisions to further others' goals.

Violations of self-centred welfare can be fairly easily dealt with within rational choice theory and have been, particularly in the work of Becker and other advocates of incorporation models, as mentioned. Sen, however, argues that the other two aspects of self in choice have been poorly addressed and that there is a need to address them, considering that people do take actions that are not intended to improve their own welfare.

Sen argues that, in both experimental conditions and real-world situations, there are behaviours that cannot be accounted for without a consideration of commitment, noting work motivation and environmental sensitivity as examples. He goes on to note the importance of social traditions in shaping these behaviours, as well as the related importance of established norms. There is also a role for ethical behaviour, as is seen, for example, in our sense of responsibility for future generations. Commitment is important, Sen argues, because it brings a fuller understanding of rationality and increases the range of behaviours that can be explained.

Petit, however, takes issue with Sen (2005). He does so in a rather interesting manner: he first takes it upon himself to summarise Sen's argument, then he hypothesises about what would lead Sen to these conclusions and then proceeds to dismantle his hypothetical explanations.

Sen's commitment, as Petit sees it, is characterised by the following of goals other than your own. Petit's paper is then essentially concerned with asking

how it is possible not to follow one's own goals. In his summary of Sen's arguments, Petit identifies and names two forms of commitment: goal-modifying and goal-displacing commitment. Goal-modifying commitment is not seen as a radical departure from rational choice theory, provided you take it at its minimum, as the pursuit of one's goals rather than as the pursuit of one's self-interested goals.

It is goal-displacing commitment with which Petit has a problem, struggling to understand how reasoned choice can be made in pursuit of a goal that is not your own. Sen, he argues, addresses this problem in an unsatisfactory way by suggesting that decision makers can act as if a goal is their own even when it is not, particularly in the group context. Petit sees no problem with group goals being followed, but argues that they become the individual's goals, as it is nonsensical to have an agent without a goal of their own.

Petit seeks then to ask how it is that Sen can come to have advocated such a position. He provides three theses, which if taken as the starting point can explain how Sen came to his position: the no-deliberation, the selfish-deliberation and the integrated-deliberation theses.

The no-deliberation thesis would assert that actors have desires and beliefs and seek satisfaction based on them. This view, however, does not require deliberation and is mechanical in nature, whereas deliberation involves the use of higher-order beliefs as a check on the mechanical response. If rational choice is argued not to allow for deliberation to occur, as in this thesis, and it is seen to occur, then an argument can be made for something that transcends it, such as commitment. Petit, however, acknowledges that, while this would form a basis for Sen's views, he does not believe it is and quotes Sen as seeing no problem with allowing for deliberation within the context of rational choice.

The selfish-deliberation thesis would state that, where there is deliberation, it is self-serving. Again, this conceptualisation would, according to Petit, give Sen grounds to argue that sometimes there is more to decision making, but again he concedes that he does not think this is the argument that Sen has made.

The integrated deliberation thesis allows for non-selfish deliberation, but only in so far as it is towards the furtherance of standing integrated goals. Thinking of rational choice along these lines would provide motivation to argue that certain behaviours transcend goals and the approach is therefore flawed. Petit, however, argues that, while the application of rational choice has typically assumed stable goals, it is not a necessity; therefore, if Sen's arguments are based on a problem with this assumption his criticisms are flawed.

Hausman places himself in Sen's camp, although he believes that he offers a critique that is conciliatory towards orthodox approaches (2005). He supports Sen in the rejection of a simple construction of preferences. Sen, he points out, distinguishes between different ways of using the word 'preference' and has a problem with arguments that require a move from one to another. Hausman describes the two uses of the word 'preference' that Sen refers to as: the one that refers to choice ranking and the other that refers to expected advantage ranking. He goes on to argue that there are other ways too that the word is used in economics.

Hausman argues that, while Sen is in favour of simply pointing out the different definitions and how these affect the arguments, he is not prescriptive about which use of the word he favours. Hausman, on the other hand, believes that there should be prescription in the usage and favours a construction of preference as all things considered preference, with a central role for beliefs. He argues that this is close to the everyday conception of preferences but not quite the same, as there is not the same contrast with duty. If this definition of preferences is used, then the concepts of sympathy and commitment can both

be included as rational, rather than resorting to reference to commitment as counter-preferential choice simply because it is not in line with expected advantage. All things considered, preferences already include influences of commitment.

Hausman suggests that Sen's argument is for a more considered look at the complexity of human decision making and that his is a similar argument. He then argues that this complexity can be considered by looking at the formation of all things considered preferences. Hausman then suggests that, while some economists may argue that this task be better left to the psychologists, he believes it is an important area for consideration. If it is not considered in economics, then there is no focus on when the simplifications that are employed are appropriate and when they are not. Furthermore, game theory would be extremely narrow if it made no effort to explain preferences and their formation.

The idea of having a goal that is not your own is difficult to comprehend, particularly if you accept individuals as having changing goals. The commitment argument is, however, not as clear as this. If you are committed, you follow a goal that is not your own; in some sense, you have it as a goal though not because you are better off pursuing it but for some other reason. The question is: do individuals select goals other than for their own benefit?

The re-conceptualisation of preferences as all things considered preferences again misses the subtlety of the problem. If individuals use all things considered preferences to rank alternatives, is that ranking reflective of relative advantage to themselves? If all things considered preferences do not reflect relative advantage, then choice is irrational so long as rational is considered as self-interested. If all things considered preferences are in line with relative advantage, then there is no such thing as commitment in Sen's terms, as all actions will be driven by self-interest.

Explaining commitment within the traditional framework is clearly still highly debateable. Khalil suggests that the problem with self-interested explanations is clearest when individuals engage in self-sacrifice or spontaneously (2004a). He notes that arguments have been put forward that individuals place a value on the afterlife or post-death fame. Such explanations, however, fail to consider why individuals would choose a particular religion or reading of it, or how they would weigh up post-death benefits against expected pay-offs from other actions.

The argument that, at some level, every choice and associated action is motivated by self-interest would seem impossible to beat, yet still many oppose it. Perhaps it is because such a dim view of human nature is hard to accept. It may, however, be possible to change the conceptualisation a little to make self-interest more palatable.

Imagine for a moment an individual in a control room. The control room contains a series of gauges each reporting a particular reading. The individual's job is to maintain those readings at the desired level; to do so requires her to spend time and resources attending to the various sources of the readings. At the end of the working day, two outputs are generated: one is a payment to the operator, the other is distributed to others. The controller attends to the gauges out of self-interest but as a result generates benefits for others. The benefits to others are not simply an externality, though some may be; they are one of the two direct outputs of production. During the course of operation, the operator may well have to attend to readings on gauges that do not directly affect output to her, but must be maintained for the health of the system. The operator is self-interested but the system benefits her and others.

In this conception, different pressures are not substitutable, but they are to an extent tradeable. The operator can decide to allow the pressure to increase to

a high level on one gauge because she prefers the output associated with keeping the pressure on another down. She can, however, only do this to a point and will eventually need to attend to the neglected gauge.

Think now in human terms, humans feel thirsty, hungry, tired, pained etc. These are all guides that pressure our consciousness to act in response to them. We can also anticipate them and act to avoid the pressure.

These feelings are physically derived but the body has a process of assessing and converting physical conditions, such as dehydration, into mental stimuli. These stimuli then act to prompt you in the direction of addressing the physical conditions. We are also equipped with the means to assess and convert information on other needs into emotional prompts. For example, we feel 'wants' for numerous different things. These are again felt prompts towards obtaining them, but they are not conversions of physical conditions, but of other needs such as prestige or self-worth. Some needs are combinations of psychologically-based needs and physical concerns, such as security, which again lead to emotions that place a pressure to act: an example in this case would be fear.

How these different emotions interact with one another is an interesting question. In this representation of emotions as pressures or prompts, how does the individual decide which pressures to heed or which prompt to follow, bearing in mind the possibility of predicting future pressures and planning to address or avoid them? In the mid-range, where no pressures are extreme, it would seem reasonable to suggest that the individual chooses the options that they feel will relieve the greatest pressure, resulting in the greatest comfort. This would seem little different from utility maximisation, but it is. Balancing pressures that are not all self-interested, may be considered self-interested if considered only in terms of responses, but in terms of outcomes the situation may not be so easily characterised. The implications of this distinction between

responses and motivations are discussed in more detail in Chapter 8. Some light can, however, be shed here by questioning what happens at the extremes.

The conception of having multiple motives and balancing between them is all that is needed for most of the arguments presented in this thesis. Essentially, the balance model that results is a multiple-utility model with a decision theory relating to balance. The discussions in Chapter 8 make the case that this decision theory does not reduce the model to utility maximisation. The following discussions focus in detail on decision making when pressures are approaching their extremes. This addition is included as it brings in the possibility of 'haves' and 'cannots', which prove useful in distinguishing this type of approach from those based on maintaining a positive self-image. They do this by highlighting the importance of the non-substitutability aspect of the argument. Much of the debate in this area is intuitively guided and the examination of extremes helps in this regard.

When you are not very hungry it would seem reasonable to consider trading going to eat for working overtime. This is a trade-off that people make all the time and arguably they are doing what they feel is best for them. When you are starving, it is difficult, if not virtually impossible, to even consider an offer of anything other than food, if it means delaying eating. This is perhaps a somewhat too general a conclusion and requires a limitation.

Assertion 1: As a pressure approaches its extreme so attention to it approaches fullness. The closer to its extreme the more difficult it is to consider options other than those which reduce this pressure, unless another pressure is similarly approaching its extreme.

The above formulation is somewhat imprecise. Any rise in a pressure could be argued to be an approach to its extreme. It would perhaps be more useful to cast the above in the language of constraints and to suggest that a critical point

being passed is the key. This would, however, imply a discontinuity which would fail to capture the concept of increasing need to respond to a pressure. What the above does is suggest that the need to respond is exponentially related to the pressure felt as a result of the simultaneously increasing pressure and attention to it. It may, therefore, be helpful to think of an inflection point in the relationship between pressure and the need to respond; once the pressure passes this point the need is akin to a constraint. The point here is not the development of a precise formulation, but rather to provide a framework within which to locate discussions so that the nature of the model can be understood more clearly.

As we are capable, at least to some extent, of predicting pressures, the above assertion can be extended to cover future choice.

Assertion 2: No option that is expected to result in a pressure approaching its extreme will be taken if another option is available that is not expected to lead to any pressures in such a range.

This essentially suggests that, close to the extremes, pressures are not tradeable as they are in what might be considered the normal range. This is fairly easy to accept when considering pressures that have a physical source. The situation is somewhat different with those having psychological origins. The pressures in these cases are subject to manipulation. Individuals can convince themselves, or at least try to, that a particular choice gives them more prestige or says more about their self-worth. Therefore, discomfort from very high pressure can, to some extent, be avoided by self-manipulation. This raises the question of the ability of individuals to deceive themselves, a topic well covered in psychology, which will be covered in detail in subsequent chapters.

So what happens when two pressures are approaching their extremes? Just as when they are not close to the extremes the individual is argued to decide between them, given their strengths; the same is argued here.

Assertion 3: When more than one pressure approaches its extreme, or will reach close to an extreme, and an individual has to choose which to relieve, they can decide on the basis of the relative strengths. Pressures close to extremes are tradeable with other pressures close to the extremes but not easily with lower level pressures.

The pressures are built in; some of what prompts them is similarly so, but other prompts are socialised and rationalised. This is not a model of instinctual behaviour. Some of the needs and resultant pressures we share with animals, such as hunger, thirst and fear, some with only a few others, such as the need for status – the prompts in that direction are shared with a number of other mammals.

Another set of pressures we as humans arguably have is a sense of right and wrong, with the pressure being for the right rather than the wrong. These pressures, resulting from a need to be sufficiently right, it is suggested, work in the same manner as those described above. They can be traded at normal levels, but as they approach their extremes they cannot, unless something else also approaches its extreme or is expected to. But this is a psychological need and so, like other psychological needs, is to some extent manipulable.

This is not to suggest that we all have some common conception of right and wrong, although there are those who argue that we as humans have certain boundaries within which to construct these conceptions (Hauser, 2007). Rather it is to assert that we have a common mechanism for indicating to ourselves the need for right over wrong. What prompts these pressures may or may not differ across societies and individuals. Some may indeed also be

common, such as a feeling that restraining or killing another is wrong (Eco, 2001). This obviously does not mean that people cannot kill or restrain others, but it is merely to suggest that to kill without justification is wrong and people feel that they should not do so. The very need for justification suggests that there is some pressure against the action, which has to be relieved in order to carry out the act. Murderers, particularly those involved in large-scale massacres, typically have some grounds for 'justifying' their actions; they do not simply participate because they want to. This issue will be discussed in detail in Chapter 4.

It is an obvious next step to ask where such a mechanism would come from. Is it part of our nature as social animals? Is it from a need to belong, to be part of something bigger? At this point, this is not such an important discussion, what is important is the assertion that the mechanism is.

Assertion 4: Humans have a sense of right and wrong preferring right to wrong. These senses produce a pressure to act in accordance with them, which in the extremes approaches a hard constraint, provided no other pressures are also close to their extreme.

This conception helps explain a number of observed behaviours discussed in detail in the following two chapters, such as the need for justification mentioned above, which is also seen with many other criminals. It helps to explain why torturers have to be desensitised, numbed down, during their training. They want the job, but cannot do it until a way around their constraining feelings has been found. This is an example of how a non-physical constraint can be manipulated.

An important area where this conception helps is explaining why people avoid knowing about atrocities, such as people did during Nazi rule in Germany. Knowing what is happening, places a pressure on you to do the right thing and

not to stand by. But it can be dangerous to do the right thing and people may not want to take that risk. What people appear to do is deny the information and so avoid the pressure. This fits well with the argument that we have a pressure to do right and that, when appeals are made to our humanity, this is the process people hope to invoke. Avoid the begging child's gaze or you will have to respond.

This is the essence of this extension to a consideration of the extremes, the presence of 'haves' and 'cannots' in human decision making as opposed to only 'wants' and 'don't wants'. It makes sense to distinguish between "I want to eat" and 'I have to eat'. 'I want to do this because it is right' and 'I have to do this because I have no choice but to do the right thing'. If we have 'haves' then we have constraints on self-interest and we have to ask how they operate, both when feelings are at their extremes and when they are not.

In the end, everything that shapes voluntary choice has to be internalised. So it is a self-interested model in so far as it suggests that good actions, or what people perceive to be good, come from a person following their sense of right and wrong and not manipulating or not being able to manipulate themselves out of the situation. But it suggests a not wholly self-interested nature: sometimes we have to do the right thing no matter what, which can hardly be described as self-interested.

The balance model is presented as a framework for understanding choice. The framework allows for the examination of a range of issues. For example, it deals well with issues of attention and seemingly inconsistent choice. Different contexts or framings lead to different feelings being invoked, leading potentially to different choices. This is also a way in which the different preference orderings, which are often suggested, could be understood. When an individual is in a situation that prompts attention to be paid to others, this may well prompt

considerations and associated feelings and pressures that are not related to self-satisfaction in the direct sense, but which prompt other-regarding motives.

To some degree this is a presentation of the way in which Etzioni's and others' multiple-utility models might operate. Although Etzioni attaches great importance to emotions (as does this model) it should not be seen as a purely emotional decision-making model - far from it. Emotions play a mediating role, cognitive processes are central in terms of generating emotions as a result of the reading of the environment, their manipulation in favour of the individual's well-being and the weighing up of pressures.

The argument that the individual seeks balance may well be challenged because, in some ways, it goes against the maximisation approach. Such a suggestion would, however, imply that people are capable of satisfying their wants. Nettle argues that people are designed only to be temporarily satisfied (2005). He suggests that feelings, such as joy, are temporary by design, as is a growing dissatisfaction with what we have. This design, Nettle argues, could well be considered a survival advantage and has helped people to prosper by always seeking more. This would suggest that, in the presented framework, pressures associated with 'wants' can generally only be kept down temporarily. Alternatively, it could be argued that the model is only in some ways counter to maximisation, as it could be suggested that the greatest benefit results from the greatest balance. This issue is discussed in Chapter 8.

The balance model presented here is suggested to be useful, as it draws attention to specific motivations, understanding that they interact, but also that they are distinct and cannot be easily collapsed. It is possible, with some imaginative formulations, to structure this same framework into a traditional approach. It would require some strong assumptions about the nature of different types of marginal utility, perhaps arguing that, when certain pressures approach extremes, the marginal utility from addressing them approaches

infinity, so they will always be addressed. Even in its current form it could be considered as utility maximisation, if utility relates only to choice and not personal benefit; that is, if it is of the grand X variety and provides no meaning beyond ordering. While this is the case, this model is not suggested because it predicts better, but because it feels less contrived and, as a result, is more natural to use. Moreover, the different conceptualisation leads to a different approach to practical questions and has different implications.

2.3 Summary and responsibility

This chapter aimed to track some of the main trends in the development of models to explain altruistic behaviour. The models that could potentially be used to consider the questions covered in this thesis can be grouped into four sets: incorporation, private benefit of action, private benefit of maintenance, and multiple-self models.

Incorporation models are those which involve the incorporation of others' welfare into the welfare function of the decision maker. Private benefit of action models are those which argue that it is the act of helping itself which generates a private benefit. Some formulations of this type argue that the value is derived from a taste for particular types of actions, such as honest or altruistic actions. Private benefit of maintenance models, notably the self-image as stock models, are based on the argument that individuals wish to maintain a certain stock of a positive self-image; as a result, the self-image acts as a voluntary adhered to constraint on behaviour. It has also been argued that self-image could lead to a hard constraint, but such formulations are grouped with other hard constraints. The hard constraint models are multiple-self models where self-interest is constrained by internal constraints which have to be adhered to.

In the following chapters, it will be argued that predictions based on the first two sets of models do not conform to observations, particularly with regard to

attitudes towards information, context and self-deception. Moreover, it will be argued that the evidence suggests that the assumptions of these models are largely flawed. The predictions of the second two sets are very similar, as has already been mentioned. The hard constraint models are, however, more rigid and difficult to apply and the use of the softer 'want constraint' formulation of self-image as stock models will be argued for.

As an alternative to 'the fork' of self-image maintenance, this chapter has presented an alternative model of behaviour. It will be argued in subsequent chapters that predictions based on this formulation will be largely similar to the maintenance of self-image models, but that the basis for the predictions may be more comfortable to accept and may lead to different implications.

More detail is, however, needed on both the maintenance of self-image models and the balance model. When comparing the models it is not useful to simply have a general self-image model. It is necessary to specify what that self-image is that is being protected, otherwise you could simply swap between images. Similarly, you could simply swap between pressures and explain everything and nothing with the balance model.

A possible useful formulation for addressing this problem is to consider the models in regard to responsibility. It will be argued in subsequent chapters, particularly the next, that using a responsibility formulation is useful when examining the helping of distant others. A responsibility formulation can be applied to a number of the models mentioned in this chapter, but it is suggested that it is particularly useful when applied to the self-image maintenance model and the balance model. That is, the self-image maintained is of a responsible person and the pressure felt is to act responsibly. Once it has been argued that responsibility plays a major role in helping behaviour, the issue of helping distant others will be examined through this lens.

This chapter has focused on the economics literature. It has considered the contributions made and some of the problems that persist. Economics is by no means the only discipline that has considered these issues: anthropology, sociology, psychology, among others, have also examined the field. Anthropology and sociology differ fundamentally from economics in their approach, making comparisons difficult. Psychology, however, does not. The following chapter compares the economics models to some of the wealth of psychology literature on helping behaviour. Helping and harming are, however, closely related issues and the chapter after next brings in some of the literature on harming. The aim is not to review this literature, but to re-examine the evidence and theory present in this literature with a view to critically evaluating the economic models identified in this chapter.

Chapter 3: Comparison with psychology literature on helping

3.1 Introduction

Within the economics literature there are a number of ways of conceptualising apparently altruistic behaviour. Economics is, of course, not the only discipline that has sought to explain such behaviour. There has also been considerable interest within psychology, especially social-psychology, as well as biology and sociology, not forgetting philosophy.

Given the wealth of research and consideration the topic has received, it is useful to examine the assumptions and predictions of the economic approaches against those that have been considered in other disciplines. It is, however, important to note that the economic approach in this field, as is generally the case in modern economics, has been to focus on the individual. Economic models of behaviour have sought to explain the decision process of the individual, with only limited consideration of the broader environment. Furthermore, economic discussions have focused on explaining individual behaviour rather than prescribing it or considering its moral worth. These aspects of the economics approach can complicate comparisons with work from other disciplines. Sociological research places a high level of importance, as would be expected, on the social structures and institutions and their role in shaping behaviour, making direct comparisons with economics arguments difficult. Philosophical debates have often leant towards discussing what should be done and how what is done compares – again a different focus that complicates comparisons.

The psychology literature on helping, moral behaviour in general, or what is often referred to in the literature as pro-social behaviour, is the most similar to the economics approach. While far more consideration is given to the broader environment, the individual is still the focus and the examination of the

individual's motivations or cognitive processes is a key consideration. This similarity suggests that the benefits of comparisons are potentially high. This chapter seeks to examine the degree to which the various economic models discussed in the previous chapter agree with the findings in the psychology, in particular the social psychology, literature that seek to explain the same behaviour. This is by no means to suggest that less individualistic approaches, such as those presented in the sociology literature, are not important and that their lack of consideration in much economics work is justified. These broader contextual factors will be introduced in Chapter 7. This chapter and the next are intended to strengthen understanding of individual motivations and actions in order that the individual can be meaningfully placed in context. Indeed, the consideration of social and contextual factors in the psychology literature acts as a bridge between some of the economics arguments and these broader issues.

The discussions here will focus on the four groups of economics models as presented at the close of the preceding chapter: the incorporation, private benefit of action, private benefit of maintenance (with a focus of maintenance of self-image models) and multiple-self (with a focus on hard constraint models). Alongside these models, the alternative balance model will be discussed. In addition to comparisons, this chapter seeks to continue the argument that the most useful formulation of these models, with the exception of incorporation models, is in terms of responsibility. For private benefit of action models, the formulation is that individuals value acting responsibly; for a private benefit of maintenance model it could be that individuals value the maintenance of a responsible self-image; for multiple-self models that the responsible self-image has to be maintained as a result of it at times being a hard constraint and, for the balance model, that individuals feel a responsibility to help. These collectively will be referred to as the *responsibility formulations*.

This chapter compares these models and makes these arguments, supported by reference to the literature on helping, moral, and pro-social behaviour. This involves the re-examination of the evidence and arguments, with the aim of selecting between the economics models or improving on them. As mentioned in the previous chapter, it is also telling to consider the literature on harming and avoidance of helping. This literature is considered in Chapter 4.

3.2 Comparing economic models with the psychology literature on helping

The psychology and social-psychology literature on altruistic, moral and pro-social behaviour generally, has grown over the years in terms of both depth of consideration and breadth of coverage, making it difficult to summarise. Penner et al recently argued that this literature is best dealt with by organising it from a multi-level perspective (2005). Their organisation does indeed facilitate understanding and provides a useful framework and will be used to provide a basic structure for the discussions presented here.

Penner et al divide the literature into micro-, meso- and macro- perspectives. The micro-level encompasses research relating to the origins and sources of variability in human pro-social tendencies. Meso-level studies comprise those that focus on individual helping behaviour, typically towards other individuals and how it may vary by context. Finally, the macro-level considers helping within a group context.

While it may seem obvious to start from a micro-level and work upwards, the discussion will first consider the insights from the meso-level. This was seen as the appropriate starting point, as it has been the typical focus of research within the literature and it is also the starting point used in the Penner review. Examining the early research at this meso-level is very informative in that it shows how this work shaped later contributions at the micro and macro levels.

Furthermore, in this early period of research, summarised by Penner, many experimental studies were carried out and this provides a wealth of data that can be re-examined with a view to selecting between the economics models.

Penner et al suggest that the literature at the meso-level has shifted from an initial focus on *when* people will help to increasingly considering *why* people help. Much of the literature in social psychology has been shaped by this early focus on when people help. To be more accurate, the initial focus was actually on when people don't help. Much of the work in this area was prompted by the highly-publicised and violent murder of Kitty Genovese by Winston Moseley in 1964 in the United States. Kitty arrived home early one morning and, while walking the 30 metres from where she had parked to her door, was attacked. In the initial encounter she was stabbed and screamed for help. In response, one neighbour shouted from a window and Moseley fled but returned a short while later when he raped and murdered the victim. Later investigations revealed that 38 people were aware of the attack at some point during the event but there was a substantial delay - over half an hour - before anyone called the police. This lack of help not only prompted much of the subsequent research; importantly, it also shaped the nature of that research, a point that will be returned to during the course of the discussion.

The concerning thing for many about the above story was that, for the attack to continue, so many people had to refrain from helping. This concern increased as the consequent research concentrated on what was presented as a perplexing outcome: the more people who are available to help the less likely it is that help will be provided. Before examining some of the research that was prompted, it is useful to consider for a moment how the economics models fare in explaining this outcome, which also provides an opportunity to examine the predictions of the models in regard to attitude towards information. It was noted in the previous chapter that these differ. These differences become more important in later discussions, so the opportunity to highlight them in this

example was taken. The full application and discussion of each of the models is provided in an annex to this chapter; the following summarises the results.

The predictive powers of the alternative models can be examined through the application of a game theory approach to the following stylised and much less disturbing example. Imagine the following: an individual is walking in the rain; in front of them someone slips and falls. The individual has two options: they can help the fallen person to their feet, but this will mean putting down their umbrella, in which case they will get wet; alternatively they can continue walking past. This can be formalised as follows:

Let $B = 1$ be the increase in welfare, to the individual being helped, resulting from them being helped to their feet. Let $c = 0.5$ be the cost, in terms of getting wet, of helping the fallen individual.

A self-centred individual whose welfare was determined only by their own situation would have no incentive to help the fallen individual to their feet. The expected utility of helping would be -0.5 , while the expected utility of ignoring would be 0 .

If, however, the potential helper's utility includes the utility of the fallen individual, then the predicted outcome will be different; this would be an example of applying an incorporation model. Assuming that B is included in the potential helper's utility function, with a coefficient of 1 , changes the pay-off structure substantially. Let H stand for help and L for leave.

H: $1 - 0.5 = 0.5$

L: 0

In such circumstances the individual would always help. In more general terms: the individual would always help, provided that the benefit was greater than the

cost; would be indifferent between helping or not if the benefit and cost were equal; and would not help if the cost exceeded the benefit.

The above, however, assumes that no one else was around. The presence of another potential helper increases the number of possible outcomes. The following pay-off matrix summarises the situation. Down the side is potential helper 1's (PH1) choice and along the top potential helper 2's choice (PH2). The first number in each cell reflects the pay-off to PH1 of that particular combination of choices and the second number the pay-off to PH2.

		PH2	
		H	L
PH1	H	0.5 , 0.5	0.5 , 1
	L	1 , 0.5	0 , 0

If both potential helpers helped (H,H) then both would experience an outcome of 0.5. If one helps and the other leaves (H,L or L,H) then the one who left would experience an outcome of 1, while the one who helped would experience an outcome of 0.5. If they both left the fallen individual (L,L) they would both experience a 0 pay-off. The preferred outcome for each individual would be for the other to help and for them to leave, but, as they only control their own actions and do not know in advance what the other will do, their decision has to be based on predicting the other's actions.

It is shown in the annex that, faced with such a situation, individuals will help half the time and not help the other half of the time. This result is, however,

based on the assumed values. More generally the above could be presented as follows:

		PH2	
		H	L
PH1	H	B-c , B-c	B-c , B
	L	B , B-c	0 , 0

As in the above example, no one strategy is dominant and potential helpers will help or leave based on a probability distribution over outcomes. The relationship between the costs and benefits will shape this distribution.

This result suggests that the greater the cost relative to the benefit, the lower the probability that an individual potential helper will help. If the cost is zero everyone would try to help; if the cost is greater than the benefit then no one would help. These are obvious conclusions. What is more interesting is the probability, at equilibrium, that the person who fell is helped by anyone. It is shown in the annex that, if the costs are less than the benefits, the chances that the individual will be helped by one of the potential helpers is positive but, so long as there is any cost involved, this chance is not a certainty. Recall that, in the case above with only one potential helper, help was provided with certainty under the same conditions. This result suggests that when there are two potential helpers the probability that the fallen individual is helped is smaller than when there is one, as long as there is any cost to helping.

It can also be shown that, as the number of potential helpers increases, so the probability that the individual will be helped falls. The incorporation model essentially predicts the outcome observed in the psychology literature.

It is also worth noting that, within the context of the above model and conditions, individuals would not refuse free information on what others were planning to do, as this would remove the possibility of both not helping, which generates the lowest pay-off, as well as avoiding the possibility of both helping, which is also a sub-optimal outcome for the individual.

It is possible to examine the results expected in the same circumstances if the actors were motivated by a warm glow from giving; this would be an example of a private benefit of action model. The results in the annex show that the inclusion of 'warm glow' does not make a great deal of difference to the predictions. The incorporation and 'warm glow' models presented above both make largely similar predictions in the current situation. Both suggest help will more often be provided the greater the benefits relative to the cost, with the only difference being the latter model's inclusion of the private benefit of 'warm glow'. Furthermore, both models predict that, as the number of potential helpers increases, so the probability that any individual helper will help falls, as does the probability that anyone at all will help. Finally, both predict the acceptance of information on how others will act, as it removes the possibility of sub-optimal outcomes for the individual receiving the information.

For the private benefit of maintenance models it is useful to examine the example of the maintenance of self-image models. With these models, the consideration is whether the action is in line with the individual's self-image. The individual holds a stock of self-image 'I' that will be affected by their choice. It is simplest in this example to suggest that they hold a self-image of a responsible person. As more people are available to help, the impact on 'I' (the image of the self as responsible) of not helping decreases. So long as there is

a cost of helping, in this case getting wet, the decreasing cost of not helping associated with an increasing population will lead to a reduced probability of help being provided by the individual. This fall in individual help will also be associated with a fall in the total probability of help being provided, if it is sufficiently high to offset the increased probability - associated with greater numbers - that the individual will be helped.

The hard constraint model, an example of the multiple-self models, would suggest that individuals may have to heed internal, context-dependent constraints. If they are the only potential helper, maybe the option of not helping falls outside of the constraint, and they help as it is the only option available to them. It would seem sensible to argue that this constraint is more likely to apply when there are fewer potential helpers. Obviously, the probability that an individual will help will be reduced as the population size increases to the point that the constraint no longer applies, as helping will move from a certainty to a probability. What happens with population increases beyond this point when the constraint is not applied would be shaped by what further assumptions were made regarding choices between options within the hard constraint. Any of the above three could be included and it has been shown that all could lead to a decline in helping.

The balance model would, similarly to the maintenance of a responsible self-image model, predict a decline in individual helping as population increases, given reasonable assumptions linking context to pressure. The individual feels a responsibility to help, which they respond to. This pressure to act is context-dependent and so can reasonably be assumed to decrease when there are more potential helpers reducing the probability that they will provide help. It is not that they derive pleasure from the helping but rather that they are motivated to avoid the pressure. They would have motivation to interpret more potential helpers as reducing their responsibility as much as possible. Given this motivation, it would seem reasonable to argue that the reduction in individual

helping would be great enough, with each increase in population size, to offset the positive impact on the possibility that anyone of the greater number of potential helpers would help and lead to a reduction in help at the population level. As a similar motivation can be argued to exist for the responsible self-image model, it would seem reasonable to suggest that all the economics models predict the outcome that individual and total probability of helping reduce with population size.

The maintenance of self-image models (assuming a hard or soft constraint) and the balance model also share a prediction regarding attitude towards information. Knowing that others have decided not to help would increase the cost to self-image of not helping, the likelihood of the hard constraint coming into affect, and the pressure to help, but would provide no advantage to the receiver. As a result, all of these models predict an avoidance of information.

From an economics point of view, therefore, with regard to the disturbing case, far from it being shocking that there were so many witnesses, all the models referred to above would suggest that the number of witnesses actually decreased the likelihood of help being provided. That said, the literature that resulted from this interest is useful in that it considered why this happened. Given that pretty much any reasonable explanation of helping would give the same predictions, the predictions alone in this example cannot be used to separate out which model is best describing reality.

The work conducted by Latane and Darley (1968, 1970) pioneered what has been called bystander intervention research, the authors themselves having been influenced by reports of the rape and murder. In their efforts to explain the events, they outlined five stages in the intervention process: firstly, the potential helper needs to notice the event in question; then interpret what they observe as an emergency; assume personal responsibility; feel that they are able to help; and, finally, provide help. As discussed, this early work and a

great many subsequent studies have found that, as the number of bystanders increases, so the probability of being helped falls. This has typically been attributed to the diffusion of responsibility that results from the presence of others, which occurs at stage 3 of the intervention process.

Essentially, Latane and Darley argue that, while an individual may perceive the need for help, they may opt not to provide it if they feel that they are not able to, or that the victim does not deserve help, or that there are others present who can provide the assistance (1970). They argue that, if only one potential helper is available, all the responsibility falls on them, as would all the blame and guilt associated with not helping. The diffusion of responsibility that occurs as the number of potential helpers increases, they argue, is a cause of reduced probability of helping. This type of explanation closely resembles the responsibility formulations. It stresses the importance of context in shaping the responsibility felt by any one individual. The incorporation model predicts the same outcome but by a different route and battles with including any arguments linked to differential helping across contexts. If helping changes because the individual values others differently, depending on the context, the incorporation model needs theory on how this operates.

Latane and Darley also suggest that the probability of helping reduces with group size because the cost of helping increases. Audience inhibition, they argued, results from the potential embarrassment associated with helping. Cacioppo et al present an interesting extension of the cost of helping argument associated with audience inhibition (1986). They argue that potential helpers believe that, if they provide help, as more onlookers arrive, the helpers will be held increasingly responsible by the new arrivals, for the victim's situation. If they are alone with the victim, with a low chance of new onlookers arriving, this cost is lower. They further assert that this belief has some objective basis and that new onlookers do indeed attribute greater responsibility to the helping agent. They do, however, note that this is only one cost and, while it may well

increase with group size, there may be instances where the opposite is true: for example, if onlookers see an accident and cheer on the helpers.

The reduced probability of helping as group size grows is well established in the literature, although there are some exceptions. Bihm et al, for example, conducted an experiment differing from most in that it did not involve face-to-face interaction but rather used a lost-letter approach (1979). They placed letters on cars, having recorded who had got out of the car. The letters had a note from the sender saying they had recognised the car and so decided not to post the letters, which were also addressed. They noted that the number of people in the car did not make a difference as to whether or not the letter was posted. The people in the car were likely, however, to be family or friends and the group effect in such circumstances may be very different.

A recently published study by Fischer et al examined the interaction of severity of situation and the bystander effect (2005). Their experiment involved participants watching couples interact with each other on what they believed to be live feed from a nearby room but was in actuality a video. They were also led to believe that they were there as part of a study on flirting, and the couples they watched would be doing just that. The first two interactions they viewed were uneventful but the third involved a couple, played by actors, where the man became increasingly aggressive towards the woman to the point of being abusive. The study used a 2 X 2 design, with participants either watching alone or with a member of the research team, who they believed was another participant. The men in the video also varied, one was skinny and the other was large and well built. When a skinny man was on the video the bystander effect was seen. That is, participants sought to help the woman far less when they were in the company of another. When it was a large man, however, the bystander effect practically disappeared, with hardly any difference in frequency of attempts to help when another participant was present.

It could be argued that, in the Fischer study, both the costs and potential benefits of helping changed. Tackling a bigger man could be seen as more costly, while helping a woman who was in danger from a bigger man could be seen as more beneficial or more of a responsibility. In the study, the probability of helping when a potential helper was alone was lower for the big man situation when compared to those alone observing the skinny guy scenario. This would suggest that individuals may have seen the risk to the woman as greater but this was counteracted by the perceived risk to themselves. Again, this result could be considered in line with any of the economics models, as could the reduced helping associated with an additional witness in the skinny guy scenario. What is more difficult to account for is the absence of the bystander effect in the big guy scenario.

Potential helpers may have thought the cost of helping against a large man too great when they were alone, but thought that if someone else was around they could help with a lower risk to themselves. This would reduce the costs of helping, which would be expected to increase helping when there were more people around. If this is considered alongside the prediction of a reduction in help associated with the bystander, the insignificant change in helping is not that surprising, with two factors working in opposite directions. These very specific circumstances, in which the increasing group size did not have an impact on helping, highlight its general applicability elsewhere.

There are, as noted above, alternative explanations for why helping reduces with the size of the group, but there is general agreement that it does. As already noted, all the economics models discussed in the previous chapter predict this outcome. The finding here that the result occurs does not assist in choosing between them. The argument that it is the rising cost of helping associated with either embarrassment or the potential helper's fear of having responsibility attributed to them, again does little to differentiate between the economics models, as they all consider the costs of helping as a deterrent to

help. The increased cost of helping is not a necessary aspect for any of the models to predict falling probability of help with rising numbers of potential helpers, just a cost.

Latane and Darley's (1970) argument relating to personal responsibility and how responsibility is diffused to other potential helpers as the number of helpers increases does fit very closely with the maintenance of a responsible self-image model and the balance model and, to some extent, with the private benefit of action models, where the benefit is gained from acting responsibly. The consideration of responsibility has been further examined in other studies.

Konecni and Ebbesen, for example, conducted an interesting experiment examining what impact the presence of a child would have on helping behaviour (1975). They examined the response in terms of the frequency with which help was provided and the time spent helping. They examined six categories, resulting from a 3 X 2 design, of passers-by comprising lone women, two women, and a woman with a man, considering each group with or without a child. They found that women alone were the most likely to help, although this likelihood dropped in the presence of a child, as it did when there were two women. A man and a woman together but without a child, responded least often, although helping behaviour increased substantially in the presence of a child, with the man being the more frequent helper. The authors argue that this is evidence of role-associated responsibility. Men, they suggest, are seen as the appropriate actors to demonstrate and teach altruistic behaviour to children in such settings, while women's responsibility is to the child, so the child's presence reduces helping. While this argument may appear to support the responsibility formulations, it must be considered in the context of the wider debate in the literature on to what extent social roles play a part in determining helping behaviour. This debate is dealt with shortly in more detail. At this stage, however, it is useful to note that the incorporation model would struggle to predict this behaviour, again as a result of the problem with considering context.

The studies mentioned thus far are along similar lines and focus on the initial arguments and model suggested by Latane and Darley. Others, however, have argued that the conceptualisation of the problem and the resultant model miss some important considerations relating to social constructions. For example, Shotland and Straw found that, when bystanders observed an attack on a woman by a man they believed to be her husband, they were less likely to help than if they believed they were strangers (1976). This result, they argue, occurred because bystanders did not believe it was their place to get involved in domestic disputes. Similarly, Levine argues that the non-intervention in the build-up to the murder of 2-year-old James Bulger by two 10-year-old boys occurred in part because bystanders they met believed the boys to be brothers (1999). Results such as these could be taken in two ways. Firstly, they could be taken as showing that social constructions can, at times, increase the costs of helping and therefore reduce its frequency. There would seem to be some sense to this argument. In addition to this, however, they could also be taken to suggest that individuals do not help when they do not feel or see a responsibility to do so. Given this continued discussion of increased costs of helping, it is interesting to ask the question as to what extent individuals conduct cost benefit analysis in such situations and how useful an approach it is to examining behaviour. Both these areas have been considered in the literature.

A powerful example of considering the costs and benefits in helping behaviour is provided by a number of studies that interrogate helping behaviour towards people with disabilities. Many of these are based on the argument that helping a person with a disability has greater rewards but also higher costs, typically attributed to the stigma. If this is the case, then in a situation with constant rewards a decrease in the cost of helping would prompt a greater increase in help for a recipient with a disability as compared to other recipients, whereas an increase in cost would decrease helping more for recipients with a disability. This would occur because, given the higher starting cost of recipients with a

disability, the individual considering helping is more likely to have negative or low positive net benefits from helping. A number of experiments have appeared to provide empirical support for this argument (Doob and Ecker, 1970; Ungar, 1979). Other examples of considering changing costs of helping can also be found, such as Edelman et al, who investigated the possibility that one of the considerations in a decision to help is the potential for embarrassment (1984). In their experiment, a confederate dropped a box of either tea or tampons. Assistance was less forthcoming when the tampons were dropped.

The above, and many similar studies, suggest that some cost benefit calculations do occur but they consider mainly the costs to the potential helper when the likely benefit to the individual in need of help is relatively small. What is more interesting is to examine the case when the costs to the potential helper and the potential benefit to the individual in need of help are both great. Such contexts also require the consideration of the nature of what are often psychic costs. Embarrassment, social stigma, and guilt could all be considered costs of not helping, or even of helping in some circumstances, but these are not external but internal costs. If they are internal it must be questioned to what extent they can be controlled by the individual experiencing them, particularly if they have a motive to do so such as when helping conflicts with narrow self-interest.

A scenario that shows the possible major conflicts of self-interest and interest for others occurs during drug overdoses. Bystanders have to decide if they will call the emergency services. As bystanders are typically other drug users they often fear the possibility of arrest. The possibility for self-manipulation in these circumstances can be seen in the results of Tobin et al (2005). As in other studies, larger group size diminished the probability of helping. They also found that the presence of a woman increased the probability of help being provided. The role of sex in determining helping behaviour has been seen as important and will be discussed shortly. What, however, is most interesting at this point is

that they found that bystanders who themselves had overdosed in the past were less likely to call emergency services, whereas those who had been present at a fatal overdose were far more likely. This suggests that, because of the high risk of helping, individuals were trying to play down the importance of the help, even if only to themselves. Those who had survived an overdose justified inaction by suggesting to themselves that it was not that bad, seeing that they did make it through, while those who had seen a fatal overdose themselves were unable to deny the severity of the situation. These actions would seem to suggest a manipulation of an internal pressure, as opposed to the evaluation of the costs to oneself and the welfare improvement to another, which would be appropriate in a utility-incorporating model. A counter is, however, again available in that those who had themselves experienced an overdose might genuinely believe it not to be as serious as others did, particularly those who had been present at a fatal one. It is also possible to account for such observations within the maintenance of a responsible self-image model, as individuals may find maintaining a positive self-image easier if they can downplay the significance of the situation.

While these possible explanations may clearly be thought about within a cost benefit framework, it has often been argued that that it is inappropriate. Dietz and Stern, for example, argue that the standard rational choice model is based on unrealistic assumptions about how individuals approach problems (1995). They suggest that individuals do not have the complex algebra and arithmetic skills required by standard models. Rather they are well developed in terms of pattern recognition and classification. Individuals, they argue, go through a process of problem simplification, at times to the point of habit and routine. Decisions are made, they contend, not through the weighing up of all options but through the truncation of outcome sets, often through the matching of the situation to pre-existing rules. Moral imperatives could well be included in this truncation process, providing points at which deliberation ceases. The decision process allows for the consideration of social influences on the truncation of

lists and also raises the importance of framing. While a potentially useful argument, and one which is repeated in various forms in the literature, it does not assist at this stage with the comparisons of models, as they could all be adapted in some way to include the above points.

The consideration of just a few studies on when people help has provided some support for the responsibility formulations in general, but not much to distinguish between them, as there has always been an alternative explanation that would fit one or all of the other proposed models. What may be more beneficial is to consider not simply when but, in more detail, why. Fortunately, Penner et al note that research has shifted in just such a direction and this shift links the meso- with the micro-level. From their review of the literature, they identify three mechanisms that have been discussed in relation to why people help. These consist of learning, social and personal standards, and arousal and effect. Learning mechanisms, they note, have been suggested, based on what they call operant conditioning and social learning. While the literature on social and personal standards stresses the role of the individual's wish to support a positive self-image, it is this approach, they argue, that has led to the consideration of long-term helping behaviour and volunteering - behaviours they consider to be part of the macro-level.

Related to the questions of learning and social and personal roles is the consideration of the differences in helping behaviour between men and women. Eagly and Crowley (1986), conducted a meta-analysis of the social-psychology literature in this area. They based their analysis on the consideration of gender roles. Gender roles associate certain types of behaviour with either sex and these associations create different norms that in turn play a role in governing behaviour. The female gender role, they argue, fosters certain types of help, including that associated with long-term kindness and caring, particularly towards family members and other persons close to them. The male gender role, however, contrasts sharply with this, associating helping with heroic acts

and chivalrous behaviours. This may result in men being more inclined than women to help strangers, particularly if there is an element of danger. The authors also note that women are often taught to avoid certain situations to decrease the risk of sexual assault.

As a result of different social roles, men and women may anticipate expectations of help in particular circumstances and as a result acquire the skills necessary to fulfil such expectations. The authors use the examples of men learning how to attend to mechanical problems and women learning to care for children. Because they do so, should the need for help arise, men and women are likely to view themselves as competent in different circumstances and as a result the nature of these circumstances is likely to shape who helps.

Eagly and Crowley suggest that taking such a social-role perspective highlights some deficiencies in the literature. They argue that it is not possible to label either men or women as more helpful, as the degree of help provided will vary by situation, depending on what social roles are called upon. Further, the authors stress that past work on helping has focused on brief encounters with strangers, with possibly some element of perceived risk. They argue that, given social roles, it would be expected that men are more helpful than women, not because they are generally more helpful but because they are in such situations. Their analysis of the literature supported this prediction, as men came across as helping more often. Interestingly, however, they also found that this tendency was more marked when there were onlookers. They argue that this suggests that 'chivalrous and heroic behaviour may be largely a product of social norms rather than ingrained motives or dispositions.' (1986: 301). They also found, as they expected to, that men were more likely to help women than other men.

Their analysis also showed some variation in helping associated with required skills. They did note, however, that if there was only one potential helper they tended to try to overcome any skills' deficiency.

It is interesting to ask how models that suggest that the welfare of others is incorporated into the utility function of potential helpers would account for the differences between men and women in helping behaviour. One possible way out is to suggest that, when they see themselves as occupying a helping role, men and women incorporate the utility of others with a higher coefficient. Another more plausible argument would be that they see the likely benefits that their actions could generate as different; also the costs they face may vary, as in the above-mentioned consideration of risk. The same argument could be applied to 'warm glow' approaches, if one assumes that warm glow is contingent on the success of the action. Models that consider the role of responsibility could still include the possible variation in perceived benefit associated with success and estimation of cost. More than that, they would consider how gendered roles attribute responsibility to people in certain circumstances and how this also shapes the acquisition of skills. If men and women incorporated the utility of motorists who have flat tyres, would they not have equal incentive to acquire tyre-changing skills?

The issue of context raises the question of when it is that the utility of another is included in the helper's welfare function. Connected to this issue, and identified in the literature in the Piliavin and Charng (1990) review, is the degree to which potential helpers perceived the potential recipient to be at fault for finding themselves in their predicament.

As discussed in the previous chapter, the utility-incorporating model battles to explain why the utility of a stranger would only enter the welfare function of the potential helper when they need help, and only in certain circumstances. Why not just walk down the street giving money to strangers who at the time might

need it more than you? There is an area of the psychology literature that discusses this in relation to arousal and action, which will be discussed shortly.

The literature on learning mechanisms has dealt not only with the long-term process of learning, but also with learning in the context of brief encounters. There is an area of research that examines the impact of modelled behaviour on subsequent actions. Much of the research in this area is along the lines provided by Harris and Samerotte (1975), who categorised the responses of shoppers to a survey and a subsequent request for a small amount of money after they had witnessed an aggressive response, an altruistic response, or without them having witnessed either. They found that an aggressive response was more likely after an aggressive model, although altruistic models appeared to make no difference. This result may well speak to the motivations in interpreting information. If an individual is looking for an excuse not to help, the aggressive model helps them. The models that suggest the use of internal excuses would predict such a finding; these include the private benefit of maintenance, multiple-self and the balance models.

A more general comment in the literature on the influence of a helpful model is provided by Solomon and Grotta (1976). They cite a number of studies that found that the presence of a helpful model increases helping behaviour. They note, however, that, while these studies were conducted for a variety of scenarios, they were all low-level emergencies, with little cost associated with helping. Solomon and Grotta attempted to examine how the example of an actor influences helping in varying levels of emergency. They found that, in their low-level emergency experiment, the presence of a helpful model increased helping, but in their high-level emergency situation the relationship was reversed. While the reduced helping associated with the presence of a model in the high-level experiment was not significant, the interaction between level of emergency and model presence was. This result, they argue, stems from the difference in level of cost associated with helping. In a low-level

emergency, a helpful model will serve to remind potential helpers of the appropriateness of helping and, as there is a low cost associated with the help, this increases the likelihood of help being provided. In the high-level emergency, potential helpers may perceive a high cost of helping and, if they can diffuse responsibility to the helpful model, they will do so in order to avoid the cost. This suggests resistance to information - which makes sense only if it comes as a cost. Information can be costly in the private benefit of maintenance, multiple-self and the balance models, but it is always of value in the incorporation and private benefit of action models.

Explanations considering social and personal norms are very similar to the self-image models; Piliavin and Charng cite a number of studies that suggest that helping behaviour is shaped by such norms (1990). They, however, similarly suffer from the difficulty of considering how these norms influence decisions: are they adhered to because benefit is gained from so adhering, as in the soft constraint version, or because they have to be, as in the hard constraint version?

Individuals who attribute responsibility to themselves, or who have high ethical values, may well be argued to gain more private benefit from acting in accordance with them. Such an argument in favour of private benefit of action models and maintenance of self-image models would suggest that people who risk their and, at times, their families', lives for others were simply maximising their own self-interest; that the warm-glow, avoidance of guilt, or maintenance of self-image was worth the risk. These types of situation show up the difficulties of using such approaches as advocated by either type of private benefit model. Similarly, they show problems with welfare-incorporating models. Can an individual, who risks their life to such an extent, really be described as maximising their utility simply because they benefit from the increased welfare of the helped? It would seem that the discomfort in the economics literature is similarly apparent here.

Linked to the literature on social norms are debates concerning the possibility that there is a pro-social personality. Penner et al summarise the literature in this area and cite a number of studies that suggest that there are differences in empathy and that these differences are relatively stable over an individual's lifetime (2005). Aspects of personality mentioned in the summary as being associated with pro-social behaviours, include agreeableness, a sense of responsibility, and a self-perception of being helpful. The second characteristic sits well with the balance model and the last with self-image models, although these efforts to identify an altruistic personality have met with some criticism.

Piliavin and Charng, on the basis of their review, argue that it is futile to seek to define an altruistic personality, given the wide variety of forms that altruistic behaviour can take (1990). They do, however, note that a number of studies suggest some regularities, notably: 'people high in self-esteem, high in competence, high in internal locus of control, low in need for approval and high in moral development appear to be more likely to engage in pro-social behaviour.' (1990: 31). Rushton was one researcher who argued that a 'positive sense of well-being' would be associated with altruistic behaviour (1980). Interestingly, Driver, on re-examining this hypothesis, argued that it held true only when the relationship of the other to whom the altruism would be directed was a friend, as opposed to a stranger or antagonist (1987). This type of argument fits in well with the balance model. If other aspects of the individual's life and personality are not placing a pressure on an individual to act, they have less motivation to avoid feelings of responsibility.

The final mechanism discussed by Penner et al for why people help is that concerning arousal and effect. The literature on this mechanism contains a lot of both agreement and disagreement. There appears to be considerable agreement on the existence of an emotional response to the distress of others, even among young children. While such responses are generally termed

empathy, it is further agreed that it takes a number of different forms and can involve sadness and distress or sympathy and compassion. The area of disagreement and debate centres on these differences. Those who focus on the sadness and distress argument consider empathy-associated helping as egotistic; that is, an effort to help motivated by a desire to remove the distress at its source. Others argue that, when motivated by sympathy and compassion, the helper may benefit themselves, but this is not their goal and motivations are in actuality altruistic. This is a very similar debate to that seen in economics surrounding issues such as commitment.

One of the major contributors in this field is Batson, who presents a discussion on the existence of altruism, examining his arguments against the results of experiments conducted by himself and others (1993). He begins his discussion with a definition of altruism as 'a motivational state with the ultimate goal of increasing another's welfare' (69: 1993) as opposed to egoism, which he defines as 'a motivational state with the ultimate goal on increasing one's welfare.' (69:1993). Batson argues that the use of such a definition has a number of implications. Importantly, it demands a focus on motivation and not outcome and requires a self-other distinction. Furthermore, he states that it implies that, while individuals can have both altruistic and egotistic motivations, a single motive cannot be both. The definition also does not require self-sacrifice and is not the same as morality. Batson's aim was to examine if experimental evidence suggests that altruism, so defined and with such implications, does exist. To this end he compares the explanatory power of three different egotistic-based motivations for helping behaviour with his empathy-altruism hypothesis.

The empathy-altruism hypothesis claims that 'empathy evokes motivation directed towards the ultimate goal of reducing the needy person's suffering' (72:1993). The more empathy felt, the more motivation there is to reduce that need. Batson admits that reducing need, the ultimate goal, may well have

benefits to the actor, but these benefits are unintended and, if the hypothesis is correct, their removal would not alter the selected course of action.

Batson takes as given² that the evidence suggests that empathic feelings increase helping behaviour. He notes, however, that the motivations in that helping could be argued to be either altruistic or egotistic. One possible argument suggesting egotistic motivation is the aversive-arousal reduction view. This explanation of helping behaviour essentially argues that empathic distress is a negative feeling, so individuals help in order that the cause of the negative feeling be addressed in order to reduce the unpleasantness. Batson argues that, if this explanation were true, then if individuals in situations where they were feeling high levels of empathy were offered an easy escape without helping, they would take it. Contrary to this, if the empathy-altruism hypothesis were correct, since the individuals' goal is not reduction in their own empathic distress, but relieving the suffering, the option of an easy escape would not alter their helping if they were highly motivated. Batson presents tables, similar to the following, of predicted behaviour.

Table 3.1: Predictions of helping behaviour 1 (73:1993)

Aversive-arousal reduction explanation		
	Empathy	
Escape	Low	High
Easy	Low	Low
Difficult	High	High/Very high
Empathy-altruism hypothesis		
	Empathy	
Escape	Low	High
Easy	Low	High
Difficult	High	High

² Based on an earlier review, Batson, 1991.

Batson states that, at low levels of empathy, both models would be based on egotistic motivations, resulting in identical predictions. He does not make it entirely clear why the empathy-altruism hypothesis only kicks in at sufficiently high levels of empathy. Batson argues that, at high levels of empathy, the option of an easy exit will be taken if the aversive-arousal reduction argument is correct, whereas it will not if the empathy-altruism hypothesis is. According to Batson's experiments, behaviour mirrors the predictions of the empathy-altruism hypothesis; more on counters to this argument later.

The second explanation based on egotistic motivations that seeks to explain the relationship between helping and empathy, which Batson compares to the empathy-altruism hypothesis, is the empathy-specific punishment explanation. This explanation, as he summarises it, is that, as a result of socialisation, people feel an obligation to help and an associated guilt if they don't. As a result, people learn that when they feel empathy they should help in order to avoid social- or self-censure. Batson argues that variations in information that make an individual feel more justified in not helping should reduce help if the empathy-specific punishment explanation is correct, whereas the empathy-altruism hypothesis would predict that it would not change the outcome. The following table summarises Batson's suggested predictions for the two alternatives.

Table 3.2: Predictions of helping behaviour 2 (1993:75)

Empathy-specific punishment explanation		
	Empathy	
Justification for not helping	Low	High
Low	Moderate	High
High	Low	Low
Empathy-altruism hypothesis		
	Empathy	
Justification for not helping	Low	High
Low	Moderate	High
High	Low	High

Again, Batson presents predictions that are identical at low levels of empathy but vary at high levels. He seeks to use experimental data to examine which argument better predicts the results. The data presented, however, are a little difficult. He reports on an experiment where individuals, who had been induced to feel either high or low empathy for a young woman, were given the opportunity to pledge time towards assisting her. Before their pledge, participants were told either that 5 of the previous 7 potential helpers had pledged to help, or that only 2 had. The plight of the woman was portrayed in such a manner that the number of previous pledges did not affect the need for help. These conditions, presented in an earlier paper, (Batson et al 1988), are argued by Batson to present conditions of low justification (when 5 had helped) and high justification (when only 2 had helped) for not helping. This is because if most people are helping this creates pressure to also help, but if few are, the potential helper will not feel such pressure. The results of the study reflect the pattern predicted by the empathy-altruism hypothesis.

Finally, the empathy-altruism hypothesis was compared to the empathy-specific rewards explanation. Empathy leads to a sadness and to improve their mood the individual provides help. If this is the case, Batson argues, then any mood-enhancing experience should do and, if potential helpers were presented with one, the empathy-helping link would be broken. The predictions for each explanation are summarised below.

Table 3.3: Predictions of helping behaviour 3 (1993: 76)

Aversive-arousal reduction explanation		
	Empathy	
Anticipated mood enhancement	Low	High
No	Low	High
Yes	Low	Low
Empathy-altruism hypothesis		
	Empathy	
Anticipated mood enhancement	Low	High
No	Low	High
Yes	Low	High

Again, according to the empathy-altruism hypothesis, the motivation to help is predicted to be unchanged by the anticipation of an alternative mood-enhancement experience. Referring again to a previous study, in this case Batson et al 1989, Batson argues that experiments of this nature yield results in line with the predictions based on the empathy-altruism hypothesis.

Hoffman takes issue with Batson's arguments, particularly with those regarding the exclusion of personal distress as a motivator for altruistic actions (1991). He suggests that the experiments designed to test the presence of this motive are problematic. Such experiments typically involve the offer of an easy or hard

escape and the argument goes that, if motivated by personal distress, altruistic behaviour will be reduced by the offer of an easy escape, such as presented in the experiment above. Hoffman, however, argues that, in Batson's experiments, an easy escape is not always easy and 'out of sight is not out of mind'. Further, he suggests that individuals can sustain the image of the suffering even after 'escaping'. Hoffman takes issue with the design but does himself argue that, once an individual feels sympathy and compassion for a suffering person, they are more concerned with helping than with relieving their own distress.

Batson is criticised for his approach of discounting one egotistic motive at a time (Cialdini et al, 1997). Such criticism is based on the argument that just because a motive is not responsible for action in one situation does not mean that it does not explain behaviour in another. Certainly, there would appear to be some evidence of egotistically-motivated help. For example, Harris et al examined the impact of confession on altruistic behaviour (1975). They found that subjects were significantly more likely to donate prior to confession than after, showing that once guilt was eased the self-interested motive was removed. Other studies, however, have argued that situations can result in both types of motivation.

Harris and Samerotte, for example, conducted a number of experiments to examine how acting either altruistically or transgressing affects attitudes towards future help (1976). Specifically, they were interested in how attitudes towards helping an individual a second time, after the potential helper had either helped or transgressed against the requester in the first interaction, compared to helping another stranger.

If following a transgression - in these experiments this involved allowing goods they had been asked to watch be stolen - results in sympathy toward the individual whose goods were stolen then the potential helper will be more likely

to help them than anyone else. If, however, the transgression resulted in general guilt, then the potential helper should be more willing to help anyone than if they had not transgressed. The results of the experiments presented suggest that both of these responses were evoked. Individuals who had allowed the theft were more likely to give to the individual whose items they believed to have been stolen and were more likely than other random individuals tested to give to anyone who asked. What is also interesting is that they were more willing to give when the request was for high-nutrition, high-need food than when it was for low-nutrition, low-need food.

Harris and Samerotte, on the basis of on another experiment, also argue that they can show that individuals who carried out a small favour for someone were more likely to help if that person made a subsequent request than if an identical request was made by another (1976). The individuals who were least responsive to the subsequent request by the same person were those who had already done them a large favour.

From their understanding of the literature, Harris and Samerotte state that it is a widely-accepted fact that individuals who have transgressed against, or harmed someone, are more likely to help in subsequent situations than those who have not. Their results do appear to show that individuals do feel obliged to people they have harmed or formed a connection with. This in itself is not proof of altruistic motivations, as helping the person you transgressed against may be a more efficient way of reducing guilt. Arguing that there is always some underlying egotistic motive is, as was discussed in the previous chapter, always possible, but in the psychology literature it has become less popular to do so.

Piliavin and Charng, as part of a detailed review of the period's literature on the topic of helping behaviour, argue that the literature suggests a shift away from believing that all altruistic behaviours are, on closer examination, actually motivated by self-interest (1990). They suggest that, during the period, there

was a shift towards the belief that indeed altruism does exist and is part of human nature. They state that:

“People do have ‘other-regarding sentiments,’ they do contribute to public goods from which they benefit little, they do sacrifice for their children and even for others to whom they are not related.” (Piliavin and Charng, 1990: 29)

The above examples suggest that there is evidence to at least argue for the empathy-altruism hypothesis and for an egotistic motivation to reduce guilt. The key point, however, is that people do appear to have motives beyond ‘warm glow’ or some other variant of individual benefit from helping others. That said, however, the model of impure altruism is based on exactly that: that altruism is based in part on warm glow and in part on the valuing of others’ welfare. It is difficult to conclude whether the empathy-altruism hypothesis is in agreement with models incorporating others’ welfare. The empathy-altruism hypothesis asserts that, when empathy is at a high enough level, the improvement in the welfare of the other, not the benefit to self, becomes the motivation, even if there still is a benefit to self. In a way, this is in agreement with incorporating models, as they suggest that the goal of altruism is the improvement of others’ welfare. The difficulty is that such models argue that this is the goal, because the improvement in welfare of the other improves the welfare of the person helping, so essentially the motivation is self-interest, which is the opposite of what the hypothesis is suggesting. This difficulty in reconciliation highlights a problem with the model, resulting from a simultaneous concern with others and oneself, which are at the same time the same and different. That said, another more recent experiment by Batson may highlight the distinction.

Batson and Ahmed (2001) examined what would in economics be considered a rather simple decision. They looked at ‘the prisoner’s dilemma’³, but in their

³ The prisoner’s dilemma is a simple and often quoted scenario. The scenario involves two prisoners who are under interrogation. They can each choose to defect or not. Whatever

experiment participants were confronted with the scenario that the other person had already decided to defect. In the base setting, almost all participants also selected to defect, which is what would be expected. In a second case, however, participants were induced to feel empathy toward the other person. When this was the case, almost half opted not to defect even though the other person had already done so.

This again is an example of where, for the incorporating models to generate the observed result, there would have to be a variation in the degree to which others' welfare was incorporated. It could possibly be argued that the coefficient on another's welfare varies according to the degree of empathy felt. This, however, would suggest that an individual benefits more from helping those whom they feel more empathy towards, which is exactly what the literature presented has been arguing against. Both the 'warm-glow' model and the welfare incorporation approaches do not appear to reconcile easily with the empathy research discussed. But do the other models fare any better?

The maintenance of a responsible self-image model as a soft constraint is, by definition, a self-interested model and so does not reconcile with arguments that non-self-interested motivation is possible. The hard constraint model fits neatly with Batson's arguments, if high empathy is argued to generate the constraint, but this runs into the argument of how an internal constraint can be anything other than voluntary. The balance model conceptualisation of responsibility as a pressure to act in a particular way, a way that may or may not yield other personal benefits beyond the reduction of the pressure, is very similar to aspects of the empathy-altruism hypothesis. If an individual feels a motivation to help another, and that motivation does not require there to be a benefit for the self and is a result of the degree to which they feel empathy, then it would seem reasonable to argue that empathy creates a pressure to act.

the other does, it is best for each individual to defect individually. But if both defect, both prisoners get a second best outcome. The best outcome is achieved by neither agreeing to defect.

Furthermore, that the action required to reduce the pressure may or may not necessarily align with self-interest, as in the above prisoner's dilemma example, is possible within the balance model. The one difference, which it is important to remember, is that the balance model does concede that heeding the motivation is self-interested, while the motivation itself may not be. Batson, however, only concedes this to some extent; he argues that there may be benefits to self but that these are not the primary reason for action.

In Batson's discussions of the models presented as egotistic alternatives to the empathy-altruism hypothesis outlined above, the empathy-specific punishment explanation, as he explained it, would seem, at least at first glance, to have some similarities with the balance model. Individuals learn to feel an obligation to help when they feel empathy, and to feel bad or face social censure if they don't. The important difference here is that the balance model suggests a pressure to act that, while it may generate a positive feeling if acted upon and a negative one if not, does not consider these feelings to be the sole basis for action. The model asserts that individuals will act even if the positive feelings generated are outweighed by the consequences, given, of course, sufficiently high levels of felt responsibility. The balance model is, therefore, distinct from the empathy-specific punishment explanation, as it is not based on benefits to oneself, beyond responding to the motive.

While the two explanations are distinct, the test suggested by Batson to examine the empathy-specific punishment explanation has implications for the balance model and the constraint models. The balance model suggests, as do both soft and hard constraint models, that if individuals can be led to believe, or even if they can manipulate themselves to believe, that they do not have a responsibility to act then helping would be reduced. Batson's experiment finds the opposite, but the experiment is debatable. His argument is that, if many others are helping, a potential helper should also feel obliged to help, while if few are helping they will also feel justified in not helping. He, however, ignores

the possibility that, given a motivation to avoid responsibility, individuals may well interpret both as justifying inaction. Alternatively, they may argue to themselves that, as other have helped, the individual is no longer unassisted and, while this may make no difference to the need, it could still be interpreted, with the motivation to do so, as reducing responsibility.

If Batson is right in terms of how he sees individuals reading the situation, then his results challenge the responsibility formulations. This raises a question, which is a major theme in this work, relating to where feelings of responsibility come from and what shapes them. This question is similar in many ways, at least in terms of the way responsibility is conceptualised as a pressure in the balance model, to the question of where empathy comes from and what shapes it.

Along this line of research, Russell and Mentzel investigated the relationship between sympathy and altruism (1990). Their experiment involved students being given descriptions of 20 disasters and having them rate the level of sympathy that they felt as a result. They were then asked to play the role of tax payers and to imagine there was a disaster relief fund from which they had to apportion funds between the various cases. While the study found a relationship between sympathy and allocations only for female participants, it did find, interestingly, that perceived culpability underlay the sympathy. The authors admit that the relationship between sympathy and altruism may have been complicated by the nature of the experiment, which is what is of primary interest here. By placing participants in the role of tax payers, they may have prompted them to consider other issues that they felt were appropriate for that role, thereby diluting the relationship between sympathy and altruism. A more direct impact on their own finances might have led to a different response. It is, however, interesting to think about what this – i.e. changing one's choices when assuming a particular role - means for the models. Such a situation does not fit easily within incorporation models, as it would require the inclusion or exclusion

of different people, depending on the individual's role. The responsibility formulations, on the other hand, easily explain how felt responsibility, or necessary actions to maintain a responsible self-image, could well vary by role.

Interestingly, efforts to examine variations in empathy have increased the debate on the empathy-altruism hypothesis. Cialdini et al, for example, argue that empathy is a result of a feeling of oneness with the other (1997): the greater the feeling of oneness the greater the empathy felt. Empathy can, therefore, be considered as a signal of oneness, but it is the oneness that is at the root of the helping. Oneness, in their argument, is derived from similarity. This argument has, however, met with a number of counters. Batson et al, for example, while asking the question of why there exists a variation in empathy felt towards strangers in need of help, point out that, while traditional explanations suggest that individuals prioritise strangers on the basis of similarity, their experiments did not support this view (2005). Rather they suggest that empathy variations were better accounted for by nurturing tendencies.

Furthermore, experiments have shown that feelings of oneness, empathy and helping are not so simply related. Examples involving helping friends and family highlight the difficulties. Kruger found that kinship did not increase empathy or feelings of oneness compared to close friends, although it still increased helping behaviour (2003). Although kinship did not appear to improve helping by way of empathy or feelings of oneness, it did have a unique positive impact. Kruger's results suggest that participants felt a greater sense of oneness with friends, but were more likely to help family. Again, the responsibility formulations deal well with such results; responsibility to family is obvious and difficult to deny.

Much of the research in social psychology, which was discussed earlier, examined the interaction in helping situations with strangers. The current

discussion suggests that connection may be important in the decision to help or not. Some of the earlier work did, however, seek to investigate a similar question relating to how familiarity affects helping behaviour. Macaulay, for example, presented results on this issue (1975). She, however, pointed out the need to separate the influence of attraction, which is itself ambiguous, from that of familiarity. The experiment on which her conclusions were based involved the staging of a conversation between two actors, which was held in front of a potential helper; after the conversation, one of the actors faked a need for help. The conversation was intended to introduce familiarity; however, to address the question of attraction, the conversation alternated between portraying the individual who needed help as either pleasant or unpleasant. Macaulay found that, having heard the conversation, individuals were more likely to help than those who had not had one staged for them. Further, it did not matter how the individual in need of help was portrayed.

The issue of connections brings into the debate the role that group perceptions and development of bonds plays in shaping helping. The link between empathy and helping has for some time been associated with inter-group influences, and these have been seen as a key factor in explaining variations in empathy felt. Penner et al, in their discussion of the literature in this field, note how it has been argued that perceiving a person in need of help as an in-group member increases the empathetic response of an individual who is considering providing help; as a result, it increases the probability of such help being forthcoming (2005). While the literature suggests that the group effect is clearly important, a number of studies considered by Penner et al have noted how such groups are not always rigid. Situations and information can alter who individuals perceive as in- or out-group members, which raises the question: if oneness is the, or even a, basis, on what basis is oneness determined? It also suggests a role for cognitive processes in defining the boundaries of groups, although typically the discussions focus on biological determinants.

The debate on the possibility of genuine altruism, and on how it might vary according to characteristics of the recipient and the helper's perception of them as in- or out- group members, leads into the discussion on the possibility of altruistic behaviour having a genetic origin. Such discussions, already touched on in Chapter 1, typically focus on how altruistic genes could survive selection, as surely selfish behaviour would result in greater success. Penner et al divide the arguments that altruism can actually provide an evolutionary advantage into three groups: kin selection, reciprocal altruism, and group selection.

Kin selection considers inclusive, rather than individual, fitness; what is important is the continuation of one's genes (Penner et al, 2005). Based on this, the argument for kin-based altruism suggests that kin, with genes predisposed to helping relatives, will fare better collectively. Penner et al cite a number of studies that claim to have found some empirical support for this argument. Such studies highlight the importance of relatedness in decisions to provide help, as evidence that this form of altruism exists, although emotional closeness has also been argued to play a role in mediating the effect. Piliavin and Charng, in their review, point to literature that has argued that recognition as genetically similar can be used as the basis for the targeting of altruistic actions (1990). They also comment, however, on literature that suggests that other similarities can also form a basis for allocations. These include similarities in nationality, political viewpoints, and even attitudes. Again, the possible use of responsibility is obvious.

The argument that reciprocal altruism can lead to advantage, attempts to address the possibility of that advantage existing even if the beneficiary and the benefactor are unrelated. The advantage, it is argued, would come from the provision of help in the future from the beneficiary to the benefactor. This, however, raises the difficulty of identifying beneficiaries who will reciprocate.

Kin selection only explains altruism to kin; it is even debateable whether reciprocal altruism should be called altruism. Therefore, only the group selection explanation provides an argument for the possibility that a genetic predisposition of altruism to non-kin would survive. Group selection arguments suggest that altruism provides an advantage to the group rather than the individual. As a result of this advantage groups of altruists will win out over groups of selfish individuals.

The evolutionary fitness approach does little to differentiate between the economic models. Private benefit and self-belief models are essentially egotistic, although it could be asked if such preferences would or would not be advantages. Incorporation models, and the balance model, could be considered as kin selection, reciprocal, or group selection, depending on what assumptions were made regarding whose welfare was incorporated, or on what basis responsibility was felt. All have their defendants in the literature arguing for their genetic fitness, even those who argue that being 'irrationally altruistic' creates benefits that would assist in the survival of the gene(s) (Frank, 1988). The point, however, should also be made that selfish people exist, as do seemingly altruistic people, so clearly, if these are genetic characteristics, one gene has not won out over the other; what's more, it can be risky to apply such evolutionary approaches to humans.

Ridley and Dawkins (1981) discuss altruism from a socio-biological perspective, since there is much interest in incorporating insights from natural selection to examine human behaviours. They first examine the arguments for the selection of altruism more generally, without a focus on humans. In evolutionary theory, they have a clear definition of altruistic actions as self-sacrificing behaviour for the benefit of others. The first possible explanation of how such a trait would survive that they examine is group selection, which maintains that, while an individual is at a disadvantage as a result of their altruistic behaviour, a group of altruists is at an advantage. Group selection, they argue, is problematic as an

explanation. While an altruistic group may be more successful than a group of selfish individuals, within the group the selfish would do better. It is the two-level selection that leads to the problem with the explanation.

Ridley and Dawkins also reviewed a number of explanations that they see as less problematic. Many of these arguments are concerned with relatedness, suggesting that genes of organisms that exhibit altruistic traits to relatives are more likely to survive. This is characterised in the theory of kin selection, as mentioned above. A non-related base for altruism is linked to reciprocal altruism, but this introduces the problem of cheating. They argue, however, that if organisms act as 'grudgers'- that is, they will not help another organism once cheated by them - an evolutionary, stable population, with reciprocal altruism, can result.

While convinced of the merits of the approach, Ridley and Dawkins warn of the need to approach the application of socio-biology to humans with some caution. They argue that 'the human brain has taken off on a non-genetic evolutionary trip' (32: 1981). The human environment changes too fast and evolution does not keep pace. We may well have traits that were selected many thousands of years ago, but which, given such different environments from when they evolved, now result in behaviour that may be very distant from the advantage they once gave.

3.3 Summary and conclusions

The psychology literature provides a range of interesting and insightful studies relating to helping, moral behaviour, and pro-social behaviour more generally. The comparison of this literature with the models presented in the previous chapter does not clearly point to one over the other. For example, a large segment of the literature, aiming to explain reduced helping when greater numbers of potential helpers are present, is trying to explain an outcome that all

the models could well predict. There are, however, aspects of the literature that do highlight some strengths and weaknesses of the alternatives.

The importance of context, perceived fault, and prior behaviour suggested in the literature is difficult to account for in the incorporation models. If they are to be useful, incorporation models appear to need a second, accompanying theory as to when, and to what degree, the utility of others is incorporated. Without such a theory, the approach is rather weak. Similarity, or oneness, could possibly be suggested as candidate theories, but the literature suggests that helping differs not only across individuals but across contexts more generally.

The other models similarly need some additional theory, which can be provided by the responsibility formulations. Individuals receive a warm glow if they act responsibly; they value a responsible self-image; they are constrained to act responsibly, or they feel a pressure to act responsibly. These formulations proved useful in explaining the behaviour observed in the studies reviewed and, indeed, the language of responsibility has also been used within the psychology literature from the very beginning of this body of work.

The responsibility formulation could arguably be used to try to solve the context problems of the incorporation model, arguing that individuals incorporate others' welfare when they feel a responsibility to do so. On the surface this seems helpful, but it raises another problem: why would an individual do this? Maybe because they gain benefit from the welfare of others when they feel a responsibility to help? In that case, the model is the same as the responsibility formulation of the warm-glow model. Alternatively, they do so because they value feeling responsible, in which case the model is the same as a responsible self-image model. In its standard form, it battles with contexts; efforts to address this problem turn the incorporation model into one of the alternatives.

Interestingly, although unsurprisingly, the psychology literature shares a debate with economics regarding the nature of self-, and other, interests. The whole debate around what it means to respond to empathy, asking if it is or is not self-interested, is much the same as the self-interest debate in economics. The debate here, as in the economics literature, seems difficult to resolve. There are those who argue that, as you respond to empathy or whatever, it is you who respond, so at some level it is self-interested. At the same time, there are those who seem determined to argue that non-self-interested action is possible and is distinct. The balance model can again fit into this debate as something of a compromise. The balance model would frame the debate in a somewhat different way: empathy may well be other-interested, but responding to it is not. That you feel empathy means your nature is not entirely self-interested, even if your responses to empathy and other pressures are.

From the psychology literature, there would seem to be a suggestion of the importance of context, support for the responsibility formulations in general, and a similar, possibly irresolvable, debate regarding self-interest. The responsibility formulations do, however, cover a range of models and it was hoped that some further help would be provided in choosing between them. There are two major differences across the models from which some insights can be drawn from the literature: attitudes towards information and having to act.

The warm-glow responsibility model predicts that individuals would want information on the behaviour of others. The maintenance of a responsible self-image model, in its soft or hard constraint form, and the balance model suggest that individuals might well want to avoid new information and have a motive to misinterpret the information they do have. A number of studies outlined in this chapter did suggest the possibility of information avoidance and, to a greater extent, of information distortion and self-deception. These behaviours can only be explained within the two constraint models and the balance model.

Then there is the issue of having to act. Do individuals avoid information because they do not want the damage to their self-image or is it because, if they receive the information, they may well have to act? Having to act makes no sense within the soft constraint models but it does within the hard constraint and balance models. This discussion, however, relates more to extreme situations, and is again central to the self- or non-self-interested question. The literature on helping touches on the extremes when there is consideration of heroes, which was only briefly mentioned here. 'Constraints', 'haves', 'can'ts' and 'excuses', however, are addressed in much more detail in the literature on failing to help and on harming. The following chapter examines this literature, raising issues of self-manipulation and denial, asking if we are constrained from harming if we can't find an excuse to avoid that constraint. If we are, then the human system is about far more than maintaining a positive self-image.

Chapter 3 Appendix: Game theory example

A3.1 Introduction

The predictive powers of the alternative models can be examined through the application of a game theory approach to the following stylised example:

An individual is walking in the rain; in front of them someone slips and falls. The individual has two options: they can help the fallen person to their feet, but this will mean putting down their umbrella, in which case they will get wet; alternatively they can continue walking past. If they decide to continue walking, the fallen individual will struggle to their feet on their own.

Let $B = 1$ be the increase in welfare to the individual being helped, resulting from them being helped to their feet.

Let $c = 0.5$ be the cost, in terms of getting wet, of helping the fallen individual.

Let H stand for help and L for leave.

Based on the above example, the probability of helping for each of the models can be examined:

A3.2 Purely self-interested

For a self-centred individual, whose welfare is determined only by their own situation, the pay-offs would be as follows:

H: -0.5

L: 0

The individual would have no motive to help and increasing numbers would make no difference; the fallen individual will never be helped.

A3.3 Incorporation models

In the incorporation model the potential helper's utility includes the utility of the fallen individual.

Assume that B is included in the potential helper's utility function, with a coefficient of 1.

The pay-offs are now as follows:

$$H: 1 - 0.5 = 0.5$$

$$L: 0$$

Given the above pay-offs, the potential helper would always help.

In more general terms, with only one potential helper:

$$H: B - c$$

$$L: 0$$

If:

$B > c$ The potential helper will help

$B = c$ The potential helper would be indifferent between helping or not

$B < c$ The potential helper would not help

The presence of another potential helper increases the number of possible outcomes. The following pay-off matrix summarises the situation. Down the side is potential helper 1's (PH1) choice and along the top potential helper 2's choice (PH2). The first number in each cell reflects the pay-off to PH1 of that particular combination of choices and the second number the pay-off to PH2.

		PH2	
		H	L
PH1	H	0.5 , 0.5	0.5 , 1
	L	1 , 0.5	0 , 0

In this example, each potential helper knows the pay-offs for themselves and the other potential helper.

Let p = the probability that PH2 will **not** help

PH1 will select H, if they assume a value for p such that $0.5 > (1-p) \cdot 1 + p \cdot 0 = 1-p$, and L if they assume a value such that the inequality is reversed; therefore, unless $0.5 = 1-p$ PH1 will favour one strategy over the other. If, for example PH1 assumed a $p=0.4$ then they would never help. As PH1 knows that PH2 would know this outcome, they would never assume that PH2 would have such a probability distribution. Unless p is such that both potential helpers are indifferent between helping and not helping, the outcome is unstable. At equilibrium, the probability distributions for both players are the same, as they face the same pay-offs.

There is, therefore, no strictly dominant strategy for either player, so Nash equilibrium will be based on a mixed strategy; that is a probability distribution over selecting each alternative. The Nash equilibrium, given the above values, is a mixed strategy, with each player selecting to help with a probability of 0.5 and not help with a probability of 0.5.

The above equilibrium is unique only because the decision is assumed to be once off, with no opportunity to alter the selected course of action having observed the other player's choice and it is assumed that both players make their choice at the same time. If following an outcome of (L, L) or (H, H) there was a chance to decide to help or stop helping then there would be other equilibria available in addition to the mixed-strategy equilibrium described above. These would be pure strategy equilibria (L, H) and (H, L), but in the current scenario these are not possible, given that if left the individual struggles to their feet alone not leaving time for minds to be changed. The situation is very similar to the Matching Pennies example described by Fudenberg and Tirole where the only stable equilibrium is a mixed strategy (1991: 16).

More generally the above could be presented as follows:

		PH2	
		H	L
PH1	H	B-c , B-c	B-c , B
	L	B , B-c	0 , 0

If $c > B$ then L is the strictly dominant strategy and no one will help.

If $c < B$ then, following the same logic as above, the equilibrium probability distribution over alternative strategies can be predicted by solving for what probability distribution over alternatives will result in each player being indifferent between alternatives. The prediction can be made as follows:

$$B-c=(1-p).B. + p.0 \quad \dots[1]$$

Which simplifies to:

$$B-c=(1-p).B \quad \dots[2]$$

The above can be used to predict the probability that each individual will help.

$$1-p = p_h = (B-c)/B \quad \dots[3]$$

Where p_h denotes the probability that a potential helper will select the help option.

This result implies the following:

- The greater the benefit to the fallen individual, the more likely a PH will help.
- The greater the cost, the lower the probability that an individual PH will help.
- If the cost is zero, everyone would try to help,
- If $c \geq B$ than no one would help.

The above are obvious conclusions, what is more interesting is the probability, at equilibrium, that the person who fell is helped by anyone, which can be predicted as follows:

$$PH = 1-p^2 \quad \dots[4]$$

Where PH denotes the probability help will be provided by anyone.

Using [3] [4] can be rewritten in terms of B and c as follows:

$$PH = 1-(1-(B-c)/B)^2 \quad \dots[5]$$

For the PH in [5] to be positive the following identify must hold:

$$(1-(B-c)/B)^2 < 1 \quad \dots[6]$$

If $B > 0$ and $0 < c < B$ then $PH > 0$, but also $PH < 1$.

If $c = 0$ then $PH = 1$

Remembering the case above of $n = 1$, where $PH = 1$ under the same conditions, this result suggests that, when $n = 2$, the probability that the fallen individual is helped is smaller than when $n = 1$, as long as there is any cost to helping.

The result can be examined at a population level. When there are N individuals present, equilibrium will again occur when each potential helper is indifferent between helping and not helping, based on the predictions of the probability that others will help. The equilibrium probability distribution can be predicted based on the following:

$$(B-c) = (1 - p^{N-1}) \cdot B + 0 \cdot p^{N-1} \quad \dots[7]$$

Which can be rewritten as follows:

$$(B-c)/B = 1 - p^{N-1} \quad \dots[8]$$

At the population level, the probability of help being provided is predicted by the following.

$$PH = 1 - p^N \quad \dots[9]$$

As N increases $(B-c)/B$ remains constant but as $(B-c)/B = 1 - p^{N-1}$ this means that p^{N-1} similarly remains constant. For this to be true, each individual potential helper's probability of helping must fall and their probability of not helping must increase as N increases. As $1 - p^{N-1}$ is constant as N increases and p is inversely related to N , the probability of help being provided [9] also decreases along with N .

It is also worth noting that, within the context of the above model and conditions, individuals would not refuse free information on what others were planning to do, as this would remove the possibility of both not helping, which generates the lowest pay-off, as well as avoiding the possibility of both helping, which is also a sub-optimal outcome for the individual.

A3.4 Private benefit of action

It is possible to examine the results expected in the same circumstances associated with actors motivated by a warm glow from giving, as an example of a private benefit of action model. 'Warm glow' is included alongside utility incorporation; that is, it is a model of impure altruism.

Assume that the warm glow is generated by the act itself and that act is excludable – i.e. only one person can help the fallen individual to their feet while everyone who comes to try and help can get wet. If both help, then each potential helper predicts that they have a 50% chance of being the one who gets the warm glow.

The pay-off structure for this scenario with two potential helpers is given below, where b is the private benefit – the warm glow.

		PH2	
		H	L
PH1	H	$B-c+b/2, B-c+b/2$	$B-c+b, B$
	L	$B, B-c+b$	$0, 0$

If $B+b < c$, then L is strictly dominant and no one will help.

If $B+(b/2) < c$, but $B+b > c$, then the same logic as above can again be applied.

The probability distribution over strategies that would lead to the mixed strategy Nash equilibrium can be predicted by identifying the probability of leaving that would result in each potential helper being indifferent between helping and leaving. This can be predicted as follows:

$$(1-p).(B - c + b/2) + p.(B - c + b) = (1-p).B + 0.p \quad \dots[10]$$

Which simplifies to:

$$p/(1-p) = (c - b/2) / (B - c + b) \quad \dots[11]$$

The above implies that:

- The larger the cost the less likely that a potential helper will help.
- The larger either the benefit or the warm glow, the higher the probability of helping.

For a population N, the mixed strategy Nash equilibrium is again shaped by the probability distribution that leads to indifference, which can be predicted as follows:

$$(1-p^{N-1}).(B - c + b/[(N-1).(1-p)]) + p^{N-1} .(B - c + b) = (1-p^{N-1}).B + 0.p^{N-1} \dots[12]$$

Which simplifies to:

$$p^{N-1}/(1-p^{N-1}) = c - b/[(N-1).(1-p)] / (B - c + b) \dots[13]$$

The above implies that:

- An increase in N decreases both the individual and collective probability of help being provided.
- The individual potential helper would want to accept information on how other potential helpers will act.

The incorporation models and the warm glow example of private benefit of action models, presented above, both make largely similar predictions in the current situation. Both suggest help will more often be provided the greater the benefits relative to the cost, with the only difference being the latter model's inclusion of the private benefit of warm glow. Furthermore, both models predict that, as N increases, so the probability that any individual helper will help falls, as does the probability that anyone at all will help. Finally, both predict the acceptance of information on how others will act, as it removes the possibility of sub-optimal outcomes for the individual receiving the information.

A3.5 Private benefit of maintenance

For the private benefit of maintenance models, it is useful to examine the example of the maintenance of self-image model. With these models the consideration is whether the action is in line with the individual's self-image. The individual holds a stock of self-image 'I' that will be affected by their choice.

Assume that each potential helper wishes to maintain a self-image of a responsible person.

Assume that as more people are available to help, the impact on 'I' (the image of the self as responsible) of not helping decreases.

So long as there is a cost of helping, in this case getting wet, the decreasing cost of not helping associated with an increasing population will lead to a reduced probability of help being provided by the individual. This fall in individual help will also be associated with a fall in the total probability of help being provided, if it is sufficiently large to off-set the increased probability of help being provided associated with the increase in numbers.

The pay-off matrix for this model is somewhat different. It does not involve direct benefits but rather the avoidance of costs. As with the previous examples, c is the cost of helping. In this construction, i_n is the cost of lost self-image associated with not helping when there are n individuals available to help. In a more general model, n would be replaced with context, as it is unlikely that the number of helpers will always be the only factor.

		PH2	
		H	L
PH1	H	$-c, -c$	$-c, -i_n$
	L	$-i_n, -c$	$-i_n, -i_n$

In this case, potential helpers will help if $i_n > c$.

The probability that an individual potential helper will help is given by:

$$ph = \text{Prob}(i_n > c \mid n) \quad \dots [14]$$

By assumption, i is decreasing in n , and c is constant therefore ph is also decreasing in n .

The probability of anyone helping is:

$$PH = 1 - (1 - ph)^n \quad \dots[15]$$

It is, however, more difficult to say, without adding more assumptions, if this is decreasing in n , as it is with the other models. It will depend on the relationship between the impact on the responsibility stock at n and at $n+1$. In this model, the individual does not try to predict the others' action directly; rather the responsibility cost includes this likelihood: the more people who can help the less responsibility it is of theirs because others could help, so there is less impact on their self-image if they decide not to help. While this obviously results in lower levels of individual helping, it will only result in less population helping if the decrease in individual helping is sufficient to offset the increase in helping associated with more potential helpers. There will be a decrease in helping if the following holds:

$$(1 - ph_n)^n > (1 - ph_{(n-1)})^{n-1} \quad \dots[16]$$

Where ph_n is the probability of helping when there is a population of n potential helpers and ph_{n-1} is the probability when there is a population of $n-1$ potential helpers.

If the reduction in the probability of helping is sufficient for this inequality to hold, as the population increases, so total helping will decrease.

A3.6 Multiple-self models

The hard constraint models, an example of the multiple-self models, would suggest that individuals may have to heed internal, context-dependent constraints. If they are the only potential helper, maybe the option of not helping falls outside of the constraint and they help, as it is the only option

available to them. It would seem sensible to argue that this constraint is more likely to apply when there are fewer potential helpers. Obviously, the probability that an individual will help will be reduced as the population size increases to the point that the constraint no longer applies, as helping will move from a certainty to a probability. What happens with population increases beyond this point when the constraint is not applied, would be shaped by what further assumptions were made regarding choices within the hard constraint. Any of the above three could be included and it has been shown that all could lead to a decline in helping.

A3.7 The balance model

The balance model would, similarly to the maintenance of a responsible self-image model, predict a decline in individual helping as population increases, given reasonable assumptions linking context, and population size as an element of that, to pressure. The individual feels a responsibility to help, which they respond to. This pressure to act is context-dependent and so can reasonably be assumed to decrease when there are more potential helpers, reducing the probability that they will provide help. As it is not that they derive pleasure from the helping but rather that they are motivated to avoid the pressure, they would have motivation to interpret more potential helpers as reducing their responsibility as much as possible. Given this motivation, it would seem reasonable to argue that the reduction in individual helping would be great enough, with each increase in population size, to offset the positive impact on the possibility that anyone of the greater number of potential helpers would help and lead to a reduction in help at the population level. As a similar motivation can be argued to exist for the responsible self-image model, it would seem reasonable to suggest that all the economics models predict the outcome that individual, and total, probability of helping reduce with population size.

The maintenance of self-image model (assuming a hard or soft constraint) and the balance model also share a prediction regarding attitude towards information. Knowing that others have decided not to help would increase the cost to self-image of not helping, the likelihood of the hard constraint coming into affect, and the pressure to help, but would provide no advantage to the receiver of the information. As a result, all of these models predict an avoidance of information.

Chapter 4: Comparison with psychology literature on harming

4.1 Introduction

Could you kill for money? For some people the answer, unfortunately, is yes, but we are not talking about them just yet; we're talking about you. If you were asked to kill a young child who stood in the way of a large inheritance for your prospective contractor, how much of that inheritance would be enough for you to agree to do the killing? You could plan it well, make sure you get away with it and take the money. The child is a stranger, what could be the problem? Yet, for most people, there *would* be a problem and the primary focus of this chapter is to discuss the nature of that problem.

Could you kill to protect your family? What about your country? Could you vote in favour of war? What is the difference between these and the above? Of the models of behaviour discussed in the previous chapters, there are two that propose clear explanations of this difference. The first is the maintenance of self-image: you value holding a positive self-image. Maybe in this case you value a self-image as a good person, whatever 'good' may mean to you; we will come back to this perception of good later. You do not want to kill an innocent child, as this would run contrary to your self-image, but killing to protect your family would not. The alternative explanation is that you feel a sense of right and wrong, again as you determine it. You cannot bring yourself to do something so wrong; you are constrained by your conscience, by some internal moral constraint, from pulling the trigger and killing the child.

The first explanation suggests that your aversion to killing the child is a preference, albeit a strong one. You simply do not want to do it. The second explanation suggests that you can't. In the mono-utility framework in which the self-image models are placed, the act of killing a child would generate disutility and, unless the pay-off for doing so is sufficient to outweigh this, you will not

undertake the act. Unless some fairly heavy assumptions regarding diminishing marginal returns to income are made, the argument would follow that there would be some amount of money that would offset the disutility and prompt you to undertake the killing. This is true because, as there is only a single utility, all things are measured on the same scale. Finding out that your milk has turned sour generates the same disutility as killing children: maybe less, but they are of the same type.

The balance model, on the other hand, suggests that the feeling of doing wrong is not substitutable, or at least not perfectly, with other feelings, as it is of a different type. While it cannot be substituted, it can be traded to an extent, provided it is not approaching its extreme. If you wish to find a way to undertake an action that will make you feel that you are doing wrong, you have to find a way to reduce the anticipated feeling of wrong that constrains you to a point sufficiently far from its extreme that it will be tolerable. The very idea that you can wish to take an action that you are constrained from doing is central to the argument. A less extreme example could be related in terms of embarrassment: you may want to do something but, until you can reduce your anticipated embarrassment, you cannot do it, as the embarrassment constrains you. Returning to the child killing example, you may want the money offered but, unless you can find a way to reduce the negative feeling, you cannot undertake the act. It is here that the role of mental manipulation comes in. 'If I kill this child I could take half the money and give it to charity to save many more children and still keep the other half for myself.' Or 'I owe this child nothing, I have no reason to feel bad about some rich kid, I need the money.' You could say the same under the maintenance of self-image model just to reduce the disutility, but that is the question. Are such excuses simply to make wrongdoers feel better about themselves or are they necessary for extreme wrongs to become options in the choice set?

The split between the two explanations is not as clear-cut as it may first appear. Maintaining a positive self-image may be so important that the disutility from damaging that image too much cannot be recouped from any other source. This would be an example of a heavy assumption regarding marginal utilities, mentioned previously as being necessary to bring the models' predictions in line. If this assumption were made, the two models would in many ways be very similar in their predictions, even at the extremes. With the above assumption, acting in accordance with some set of internal beliefs becomes necessary, at least when the damage to self-image would be extreme, even in a mono-utility construction. If it is not possible to interpret a situation so as not to damage one's self-image beyond a certain point, then the action that would cause this situation to arise is not considered an option, no matter what other benefits there are. To some extent, this chapter then could be framed as asking if such a strong assumption is appropriate.

Although, with such an assumption, the predictions of behaviour offered by the maintenance of self-image model become essentially the same as the balance model, they are still different in character and the implications of accepting one or the other model differ. The maintenance of self-image model would, with the above assumption, be rather clearer on the origins of the constraint being self-image, but the necessity of being able to interpret an action in a particular way before it can reasonably be considered as an option is apparent in both. Whether this necessity comes from the very strong need to maintain a positive self-image, or from some internal awareness of right and wrong, or a combination of both, is discussed in this chapter. This discussion is arguably secondary to the necessity discussion, but in some ways it is more important. Even with the strong assumption mentioned, the mono-utility model still suggests that all utility is of the same type. You may get lots more of it from protecting your self-image at the extremes than you would from eating ice cream, but utility from both is still conceptualised as being of the same nature. The difficulty of seeing utility in such narrow terms has been noted throughout

the preceding chapters but is perhaps highlighted best here when discussing such extremes.

Both the balance model and the maintenance of self-image model are based on interpretations of reality rather than reality itself. As discussed in the previous chapters, they both open the way for self-manipulation, information avoidance and rationalisation suggesting a motive for such exercises. All such exercises imply a strange process of interpreting situations in a favourable manner to the self. This may be by way of the exclusion of information or the misinterpretation of reality. These are commonly-encountered phenomena. When I am watching football, the ref is always against my team, he must be: how else could he miss such blatant fouls? Essentially this comes down to a denial of reality.

This chapter will first discuss some of the arguments relating to the psychology of denial, as it is useful to examine the possibility of self-manipulation before the examination of evidence relating to the how such manipulation plays a role in shaping behaviour, particularly with regard to behaviours that harm others. It is in the literature on harming others that the discussion of constraints and necessity can be dealt with in some detail. There is no proof offered in this chapter, none is available, but it is argued that people by their nature are constrained from certain negative behaviours, unless these can be sufficiently, not necessarily fully, justified to the self. It may be too kind to call an individual who has no constraints and for whom all wants are the same and are judged with a single utility function, a 'rational fool'. Such a person might better be described as a 'rational psychopath'.

4.2 Denial

The psychology literature on denial stretches back many years and has been through many phases. A thorough review is not necessary here; what is required is a summary highlighting the major disputes and explanations offered,

so that some sense of the possibilities and conceptual difficulties can be provided. Fortunately, Cohen (2005) provides just such a summary as part of his comprehensive work on denial of atrocities and suffering. The following section draws on Cohen's summary to identify the relevant aspects to support the discussions in this chapter. Before examining the summary of the psychology literature, it is worth taking a minor detour and examining some of Cohen's introductory discussions, as they too provide a useful background.

Cohen presents a very impressive and comprehensive work examining issues surrounding the denial of atrocities and general human suffering. His work stems from personal experience, first as a white South African during apartheid and, later, as an Israeli Jew. He felt from an early age that something was amiss with the South African system; as an adult he saw the evidence of state violence in Israel. In both contexts he felt uncomfortable and, as an adult, felt the need to do something about the situation. He appears to struggle to understand how others living in the same situations could behave as if there was no problem and felt no need to respond. While his work focuses on bystanders – be they individuals or states – denying their responsibilities and does not focus on their desires to help or harm, his motivation suggests why the work is so important for this thesis. He felt a need to respond because of what he saw, but others did not and do not. He does not seem to consider the possibility that others knew what was going on, accepted it was wrong and accepted their part in its continuation, but did not care because they benefited. Without ever making it explicit, he appears to base his arguments on the belief that, if only people knew what was going on, accepted the unjustifiability of these actions and saw their responsibility, they would feel the need, as he did, to respond. To avoid this situation, they deny what is happening in some way or they deny the implications of what they see. This is the origin of his focus on denial. He is clear that the process occurs and is widely acknowledged, but also that denial can take different forms.

The existence of common expressions of denial suggests the general acceptance that it occurs in some form or another. Cohen cites a number of common expressions to illustrate the point, among them:

Turning a blind eye

Burying your head in the sand

She saw what she wanted to see (Cohen, 2005: 1).

The possibility of denial seems to be widely considered, but there is not one simple form of denial and there is certainly not one neat explanation of how it might operate and what it actually involves.

Cohen identifies three forms of denial: literal, interpretive and implicative. All of Cohen's discussions and definitions relate to atrocities, but many can easily be transferred to other areas of interest. Literal denial, as Cohen defines it, is when the happening of the event is denied. It is said not to have happened at all. Interpretive denial is when the event is acknowledged but a different meaning is attached to it: it is not what it may seem. Implicative denial refers to denials of the implications of the event. This is again suggestive of an underlying assumption that there are implications for an individual associated with the acceptance of information.

In addition to the different forms of denial, Cohen discusses how there are differences depending on the individual's role in events, again with the focus on atrocities, but again with the possibility of wider application. There are the denials of victims: potential victims may deny the risk to prevent themselves from panicking or to allow themselves to continue with everyday life. There are the denials of perpetrators that, he argues, allow the atrocities to be committed and also allow the perpetrators to continue with normal life thereafter.

What interests Cohen most are the denials of bystanders, be they individuals who are close to the atrocities, or external to them, or if the denials are by bystander States. How is it that people lived next door to concentration camps and continued to live seemingly normal lives? People would not typically stand by and do nothing if their family was being persecuted. While Cohen regards it as obvious that the same level of response to strangers would not be expected, he does note that the 'boundaries of the moral universe vary from person to person.' (Cohen, 2005: 18). He also argues that they change, expanding and contracting across time and context. It is the differences in boundaries and how they alter that is central to the discussion. For Cohen, it seems simply knowing about atrocities and suffering was enough to prompt an urge to respond. He argues that many human rights activists, having had the same response, believe that if only the wider public could be provided with the information they too would have this urge to respond. He argues that the provision of information alone is not enough, as people appear to somehow deny it, in one of the ways described above. People who lived alongside concentration camps, where burning human hair rose out of the chimneys, certainly had access to information, but claim they never knew what was going on. These claims raise a number of important questions: firstly, are such claims for others only or are they in some way for the claimant? If the claim was for themselves, was it to avoid a compulsion to respond, or simply to feel better about not responding? And finally, how could they not know? The first two questions are essentially the topic of this chapter; the last speaks to the psychology of denial. This is a complex and highly-debated area. Fortunately, Cohen, having the same question, provides a review of the main arguments. This is briefly summarised in the following paragraphs.

Much of the work on denial has been shaped by Freud's early work on the issue. Cohen argues that the best interpretation of the term used by Freud in his work in this area is 'disavowal'. The original use related to disavowal as a defence mechanism: an individual refuses to recognise what, if recognised,

would be traumatic. This can take two forms: neurosis, where the reality is ignored; and psychosis, where the individual tries to replace reality. It is not that the individual is capable of totally blocking out the reality or wiping the perception from memory; rather the disavowal is a process of continuing suppression. Cohen notes that Freud, like later authors, gets a little inconsistent in his arguments. At times, he argues that it is not the reality that is denied but the unwanted implications, but at other times he seems to argue that it is the reality itself. For Freud, this is all an unconscious process; the defence mechanism would not work if the individual knew he was denying. This is where things become complicated. To know that the reality should be suppressed, or that the implications of reality should be suppressed, requires that the reality be perceived before the disavowal can occur. The person knows but does not know; explaining this is difficult. The Freudian response is built around ideas of the splitting of the ego, but such explanations are not, by any means, universally agreed.

Similar to splitting, but without the unconscious assumptions, is explaining knowing while not knowing with the concept of lying to oneself. Such arguments generate an image of some inner dialogue where one part of the self aims to convince another of an untruth. This is difficult to see as an unconscious process, which makes it more difficult, as the individual must be aware of the internal debate. Sartre, according to Cohen, found this idea of inner dialogue, or any idea based on the splitting ego, problematic. He proposed instead what he called 'bad faith'. This is where the individual does not want to know and refuses to face the facts. This links more generally to self-deception, rationalisation, wishful thinking and other forms of self-manipulation. Again, the arguments as to the process of how such activities might operate are confusing and often inconsistent.

The cognitive revolution saw a move away from talk of motivations to avoid information and a move towards concerns with attention, memory, perception

and other similar concepts. Despite this shift, within the cognitive approach there are a number of theories that speak to the possibility of awareness without knowing, such as negative hallucinations where people can imagine that something is not there but still act as if it were, blindsight and subliminal messaging. The argument runs that there can be an initial response to stimuli without conscious awareness, which may then prompt avoidance. This comes back to the same problem of knowing in order to avoid knowing.

Other cognitive arguments consider the possibility of filters that sort out what receives attention, prescinding from the problem of on what basis it is filtered and whether it does not need to be perceived in order to then be filtered. One solution offered is that it is not that information is missed, but that attention is concentrated on tasks selectively; information that is not relevant to the task is not attended to. But, again, this requires some initial perception.

An interesting argument within the cognitive framework centres on cognitive schema. The argument goes that people give priority to previously-held beliefs; they develop a picture and resist and try to discredit challenges to it. This provides a clear basis for the motivation for the process.

It is not the purpose of this work to argue for any particular position with regard to theories of denial. The purpose is rather is to note that, within psychology and indeed common understanding, the possibility of knowing and not knowing and of resisting information is considered as very real and many examples can be identified.

Cohen discusses examples of how denial plays out in everyday life. In this discussion, he suggests that macro-denial, at a wider social or group level, links in with micro-denial of perpetrators and victims. A powerful example of this is provided from domestic violence. Denial by perpetrators was/is easier when society as a whole was/is more tolerant of domestic violence.

Other examples of denial in everyday life are given, such as responses to serious illnesses, or accidents, or the death of someone close. In regard to some of these, it could be argued that the denial is good for the denier. By denying the reality, individuals prevent themselves from feeling the negative consequences. Cohen cites research that even suggests that denying illness, or the seriousness of it, can also lead to better survival rates.

An interesting argument that comes out of this line of thought is that, while mental health is normally characterised by clarity and mental ill health by illusion, this may not always be the case. Denying reality and replacing it with positive illusions can prevent the saddening and other negative emotions that may result from the acceptance of reality. Depressed people often see things very clearly and it is this that depresses them. While Cohen presents this as an interesting argument, he does note that it is a little simplistic.

In the process of shifting the debate from everyday life to atrocities, Cohen discusses criminals and delinquents and their excuses. He suggests that they do not challenge the validity of social norms so much as they challenge the applicability of them to their situation. It is not that they have different values but rather that they propose different interpretations. The question is whether their arguments and excuses are for the audience, or to ease guilt, or to justify their having undertaken the criminal or delinquent action in the first place. Cohen suggests that there are arguments that the excuses and justifications are needed before the violation; that they provide the space for it.

The same issue of what role denial plays is evident when Cohen examines issues regarding perpetrators of atrocities. Did those who were involved in some way in the Holocaust, but later claimed not to know what was going on, really not know? Are the reasons given for their lack of knowledge attempts to manipulate their audience, or evidence of previous manipulations of

themselves? Perpetrators do at times appeal to the morality of their acts. They acknowledge what they did but argue that they were right to do so. This would suggest that the excuses/reasoning predates the act. That said, Cohen notes how often perpetrators at lower levels appeal to lack of knowledge or lack of personal responsibility, rather than moral justifications. In this way, they use mechanisms of explanation that are very similar to those of everyday criminals.

One of the most interesting and relevant examples covered by Cohen was his discussion of the desensitisation of Greek trainee torturers. They were shown tapes of gradually more and more gruesome techniques and taught to focus on details rather than the entire event. It was as if they would be in some way held back from being 'good torturers' if they did not learn to manipulate their own natural responses to the violence. So, even though they wanted to do the job, they had first to be desensitised.

A less direct and more common form of desensitisation is the changing of language - its sanitisation. There are a number of possible explanations of why this would be done: for example, creating deniability at a later date for superiors who could claim to have been misunderstood. One explanation, suggested by Arendt and discussed by Cohen, is that the language is sanitised so that the previous associations with terms can be avoided. For example, avoiding terms such as 'kill' or 'murder' avoids the association previously attached to these words that would most likely be negative, while replacing them with 'sanitise' or 'treat appropriately' does not.

Cohen often discusses the similarities in the types of denial offered by individuals and states. He goes on to provide a clear explanation of why states do so. He suggests that with full acknowledgment comes responsibility; as long as states deny, they avoid much of the pressure to rectify. Could this explanation not similarly be applied to individuals?

While there is much hypothesising in the psychology literature, the only proof that can be offered is summarised in a quote by Christopher Bollas, cited by Cohen (2005: 24-25): 'Each of us is aware in ourselves of the workings of denial, of our need to be innocent of a troubling recognition.'

4.3 Harming others

The above discussions suggest denial is a reality, and a number of the examples suggest that it is used to get round your moral constraints when wanting to inflict extreme harm. The above examples only skim the surface, as a substantial literature exists on the topic of how perpetrators seek to do this. Bandura provides a useful, and relatively recent, review of the literature on moral disengagement as it relates to the perpetration of inhumanities (1999). Throughout this literature, many of the mechanisms discussed in the previous section are implicitly or explicitly applied. For example, Bandura argues that, in order to examine this area, it is necessary to have a theory of moral agency; that is, a theory not only of moral reason but one which links moral knowledge and reasoning to moral action. To this end, he postulates that individuals refrain from actions that run counter to their moral standards in order to avoid self-condemnation. This self-sanctioning system can, he suggests, be affected by what he refers to as self-influence, which links his discussions directly to the literature summarised in the previous section. It is the different processes of self-influence used to reduce or avoid the self-condemnation that are the focus of the review. Bandura examines the literature and identifies different approaches that individuals take to avoid the self-regulatory system when perpetrating inhumanities.

Bandura argues that people do not generally involve themselves in harmful behaviour until they have developed justifications to give to themselves. He cites the example of military involvement. Killing in the course of military duty is made morally acceptable. Individuals who would never do so outside of this role

become able to kill, not because they have become more aggressive or changed their moral standards, but because they are provided with justification. The provision of justifications is one way that creates the possibility of atrocities. Bandura (1999: 195) quotes Voltaire: "Those who can make you believe absurdities can make you commit atrocities."

Euphemistic labelling, as mentioned in the previous section, is another means of trying to avoid the evoking of self-regulatory systems. Again, the military is used as one example: bombing raids are described as clean, surgical strikes, sanitising the killing and even attempting to generate a sense of them as "a curative action".

Another 'justification' is 'comparative advantage', where it is suggested that the harmful actions are in some way justified when compared to the actions of others. Similarly, the comparison could be to future harmful actions that, it is argued, will be avoided by the actions currently being considered.

The above all seek to reduce the self-sanctioning; perhaps even to convert it to self-approval for carrying out harmful actions. A second group of disengagement practices, identified in the literature by Bandura, rather than seeking to lower or convert the sanction, avoid its application by denying agency. One approach along these lines is the displacement of responsibility. Individuals can argue, for example, that they were only following orders, as did guards in the Nazi concentration camps. If they are not responsible, then their self-regulation is not appropriate. Less extreme than the displacement of responsibility is its diffusion: the division of labour, or group decision making, appear to be useful ways of reducing moral responsibility by a process of diffusion.

The next mechanism discussed is the disregard or distortion of consequences. Bandura cites studies that have shown that the more removed from the harm,

the greater the harm that can be tolerated. It is far easier, it seems, to hurt from a distance than to do the harming oneself, as it is easier to convince oneself that the harm done is not so bad.

Dehumanisation is a powerful tool in avoiding self-regulation. If perpetrators can convince themselves that the victim does not warrant moral regard then they can avoid self-censure. This, Bandura argues, is essentially an effort to deny the common humanity that would activate an empathetic response. Potential victims can be labelled with animal terms or even associated with evil qualities. He does, however, make the positive point that, when potential victims are humanised, this can be a powerful factor in reducing harming behaviour, with individuals refusing to harm even with detrimental consequences to themselves. From the case of the Mai Lai massacre, he cites a powerful example of a helicopter pilot who identified with the victims and defended them against his own troops.

The above all relate to mechanisms of avoiding or denying self-regulation. It could, however, be argued that these are mechanisms used after the fact and are tools to reduce guilt for otherwise unjustifiable actions.

The participation in acts of genocide, or in military massacres, is difficult to comprehend. It is tempting to write off such events to mob hysteria and to argue that people lose control and act irrationally; only once the irrationality ends do they look for excuses. Closer examination of events in many instances shows up the inadequacy of such a simple explanation. In Rwanda, killers cut their victims' Achilles tendons so that there would be no escapes while the guards ate and slept; so the killing could continue in the morning. The Holocaust and the killing in Cambodia continued over much greater periods than in Rwanda and cannot be written off to hysteria. The Rape of Nanking, the killing of thousands of surrendered Chinese soldiers and the rape, mutilation and killing of many more thousands of civilians by Japanese soldiers, was

controlled in the sense that non-Chinese in the area were not harmed, thus challenging explanations centring on loss of control (Dutton et al, 2005). If then the individuals who commit such acts have not lost total control, are they simply considering their options and maximising their utility, believing this to be achieved by rape and murder, sometimes even of children? Are the techniques discussed above merely mental tools to minimise the damage to the perpetrators' self-image, while they enjoy the utility of the outcome? This is not the same as trying to rationalise the behaviour of individual rapists and murderers. Often, as in Rwanda, large parts of the population were involved in the killing; people who up until the genocide led normal lives.

Dutton et al (2005) discuss a number of genocides and military massacres and in doing so highlight the importance of a number of factors in shaping events. They note that some individuals are simply protecting themselves, participating, at least at first, to avoid being considered as part of the enemy. This in itself is an important observation. This would suggest that, rather than acting as a hard constraint, morals are considered along with a more narrow definition of self-interest. The balance model proposes that the only time a feeling of wrong that is approaching its extreme can be ignored, is when another feeling is also similarly close to its extreme. Thus, individuals who would be morally constrained from killing may do so if it protects their own life or the lives of those close to them. It is important, however, to consider not only those who are protecting themselves but also how those who believe in what they are doing, at least at the time, justify their actions. This brings the discussion again to the importance of justification or some form of excuse. Dutton points to the importance of the identification of the target group, how they are typically seen or portrayed as having unfairly benefited in the past or how they are a future threat (an important justification when killing children). The violence is then justified as revenge or protection. Indeed, the perpetrators may see themselves as doing the morally justified thing. This is also reflected in the language of the

events, with target groups being described as vermin or viral, and the need for cleaning or cleansing.

The issues of justification, sanitised language or denial of responsibility come up repeatedly in explanations of atrocities and even of lesser crimes. It is as if the authors assume self-regulation and the need for there to have been a process by which it was avoided. It is as if they assume a constraint on unjustifiable actions. The excuses then are not to make one feel better, but are necessary to be able to consider the act in question as an option. Higher pay-offs that generate other kinds of benefits may be motivation to try and justify the actions, but what is needed is a justification of the same type, one to remove or weaken the constraint. Saying that you participated in a massacre because you wanted more land does not seem to be enough; it may provide motivation but you need justification not just motivation. You need an excuse that influences the constraining feeling. They stole the land generations ago by killing your ancestors and you are taking back what is rightfully yours. The higher the pay-off, the greater the motivation to find justification, but they are still different.

While not dealing directly with atrocities, there does exist a body of work relating specifically to self-regulation, but even in this literature the self-regulation could be considered through a self-image frame or as an emotional pressure.

Wilson and O’Gorman, for example, examine norm-breaking behaviour and the emotions and actions associated with it (2003). Starting from the premise that norms can be internalised, they were interested in how these norms play out in emotional responses and actions in norm-governed interactions, an area that they argue is under-researched. They show, through experimental studies involving role playing, how norms do appear to affect emotions and that these emotions are linked to behaviour, even without third party presence. What is also interesting is that they find that participants distort what they believe others

will feel. For example, they tend to expect more aggression from other parties than the participants acting those other parties report. They argue that understanding the emotional response helps in understanding how internalised norms can be enforced, although they note major variations across individuals, which they attribute, in part, to the evoking of different emotions in the same situation. This type of argument and evidence has many similarities with the balance model. It suggests that emotions are linked to decisions as a manifestation of internalised norms, but that these emotions can also be manipulated according to the situation by the individual concerned. The importance of how acts are perceived by the individual and the possibility for them to alter that perception are often raised in the literature.

Anderson and Huesmann provide an interesting review of the social-cognitive approach to examining aggression (2003). They argue that the most notable developments, in terms of theory, in this area have been in regard to the development of social-cognitive models that suggest behaviour is shaped by internal self-regulatory controls. While these models take a number of forms, they argue that they all suggest that the manner in which individual characteristics and the situation combine influences the individual's internal state.

They suggest that most people will not commit major acts of violence even if they do not fear being caught. They do not do so because they cannot escape the self-censure that would follow. Justification and dehumanisation are used as ways to avoid the applicability, or the evoking, of moral standards. The authors go through a number of causes of aggression; included in their list as the strongest situational instigator is provocation. One reason they cite for this is evidence to suggest that individuals gain gratification from harming those who provoke them; if benefit can be gained from aggression there is a motive for perceiving there to be more provocation than there is. They also list other factors, such as alcohol and drugs, which they argue inhibit internal controls of

aggressive behaviour. They also noted the importance of beliefs about the appropriateness of aggressive behaviours. All of this again speaks to efforts to avoid an internal system of control so that other 'wanted' actions can be undertaken. The idea of getting drunk in order to do something you can't do sober is familiar to many.

Anderson and Huesmann go on to discuss the important theoretical move to considering biological factors, which is done best, in their opinion, by examining how they interact with other traits and the context. They note that numerous studies have found that individuals with low arousal tend to be more aggressive. Psychopaths, for example, show very little response to images of violence and this finding is repeated in less extreme cases.

All of the above suggest the importance of emotional responses in inhibiting harming behaviour. Harming action seems only to be possible when these emotional responses are avoided or reduced. Even psychopaths do not seem to endure the empathetic response and push through it; rather it is the response, which appears to be absent, that allows them to do harm as they will. Studies have shown that psychopaths are often quite capable of telling wrong from right, where they differ is in not having an emotional response to actions that are perceived as wrong (Hauser, 2007). Psychopaths fail to read cues of distress, which typically prompt empathy and limit aggression, and fail to distinguish between moral and social transgressions (Hauser, 2007). Psychopaths' system of internal constraints does not function in the same manner as other peoples; they simply base all decisions on what they want. In a mono-utility framework, they simply have different tastes and preferences, there is nothing irrational about their actions; in the balance model something is considered wrong. The balance model assumes the existence of a moral system, a system which leads to particular types of feelings related to particular types of situations. With psychopaths this system is not functioning fully, so they base decisions only on other motives.

None of this requires some objective morality. It all suggests a moral mechanism, but what prompts it could be learnt. Hauser, however, argues that, while much may be learnt, the system is universal (2007). What happens through socialisation and other forms of learning is calibration. He goes on to argue that this calibration is limited in its impact by the nature of the system. This implies a very specialised system distinct from other mechanisms, which again is difficult to conceive of in a mono-utility framework.

Nothing here would suggest that individuals are incapable of inflicting what might objectively be considered great unjustified harm, rather that they are not capable of such if they believe it to be so. These explanations and descriptions fit very neatly with the balance model; the question is whether they can also be explained by the self-image model. As mentioned, it would require some fairly strong assumptions, regarding the importance of self-image relative to other goods, to portray self-image as such a strong constraint on behaviour, but it is still possible. The major difference is in what the acceptance of one argument over the other implies about human nature. Returning for a moment to positive helping behaviour, but remaining with the theme of extreme situations, may make this difference in implication clearer.

4.4 Helping in extreme situations

Consideration of studies of behaviour in extreme situations is particularly relevant when discussing the role of preferences and constraints in shaping behaviour towards others. The examination of altruistic actions towards Jews in Nazi Europe is one such area of study that raises some interesting issues. The importance of denial for those who did not help was discussed above, but a number of authors have also attempted to explain why some people did help, often at great risk to themselves and their families. Elster argued that those who acted as rescuers were motivated to do so by moral principles (1989).

While it might seem that such an argument could be reconciled with a preference-based model of altruism, the way in which these motivations were described makes such reconciliation somewhat difficult. For example, a principle such as, 'Never turn away anyone who needs help' (Elster, 1989; 193) can hardly be described as a preference, unless it is accompanied by some very hefty assumptions regarding the strength of such a preference. Varese and Yaish note that a number of authors cite motivations relating to a 'sense of duty' (2000: 309). Oliner and Oliner stress the importance of feelings of responsibility, not the need to feel that one is responsible, for the welfare of others as shaping actions (1988). Monroe et al goes so far as to argue that the idea of cost-benefit calculus was meaningless for rescuers (1990). This, however, is not to suggest that the rescuers were not aware of the risks and potentially high costs of helping, as such awareness has been widely acknowledged in the literature (Varese and Yaish, 2000).

Varese and Yaish argue that helping behaviour towards Jews, particularly when the helped were strangers, was often the result of a request for help, which greatly increased the probability of helping (2000). In the balance model formulation, this could be argued to have prompted a feeling of responsibility that had to be responded to if the one asked could not think of an internal excuse to decrease the feeling of responsibility and allowed refusal. That said, however, and an issue that the authors raise, the selection of those asked may also play an important role. If they somehow signal their willingness to help and this increases their likelihood of being asked, then the request is rather part of a matching process.

Maintaining a positive self-image is not mentioned a great deal in the literature on helping in extreme circumstances. The feeling that this is just what they ought to do, that there was no choice but to help, appear as the central themes in accounts provided by helpers in such circumstances. The idea that people have a range of motivations that are not all self-interested and can sometimes

be activated by the needs of others provides a more comfortable explanation of heroes and martyrs: while reacting to these motives once they are activated is self-interested, their existence may not be. That Mandela went to prison for 27 years because maintaining a very positive self-image generated more utility than he lost from being there is not an easy explanation to accept. That he had a strong sense of responsibility, that this was what he 'had' to do and responded to it seems intuitively better. If he had not responded to it he may well have felt worse, so in this sense he was self-interested. That he felt it suggests he had concerns beyond his own welfare.

4.5 Conclusions

Maintaining a positive self-image is no doubt important and is often likely to prompt people to help or not to harm. Nevertheless, such an explanation alone, or indeed any explanation framed in terms of preferences and 'wants', appears to miss something. A sense of duty, of 'ought', of responsibility, or of conviction all exist and strong arguments can be provided to show that they are in some way different from other 'wants'.

It is this difference that theorists and researchers have struggled with in economics and psychology. In this chapter, the many examples of how individuals appear to need excuses or justifications before options can be considered do not sit easily within a self-interest model of behaviour. As mentioned in the introduction, with the help of the maintenance of self-image model and some strong assumptions regarding the utility from maintaining a positive self-image and the diminishing marginal utility of other goods, all the behaviours and mental acrobatics discussed in this chapter could be explained. But still the idea that it is all only about protecting a positive self-image does not seem right.

The balance model provides a framework within which the above discussions can be comfortably housed. The important roles of self-manipulation and information resistance are clear, although this is also the case with the maintenance of self-image models. The impact of emotions on decision making also fits well, although with some effort the self-image models could manage the same. What is very different, and separates the two, is the source of emotions. The self-image model suggests that they stem only from maintaining a positive self-image, while the balance model suggests that people have a variety of emotional responses, some of which involve others, but not all of these other-regarding responses are linked to self-image, although some of them may be. The idea of evoking a strong response – such as seeing a need, feeling a pressure to help and maybe even being unable to turn away, seeing a potential victim and being unable to harm them, feeling almost a physical barrier – gives meaning to ‘I can’t do this.’ It is here that the balance model stands apart. It should be remembered that the balance model does not suggest that responding to these evoked emotional pressures is not self-interested, but only that not all the emotional pressures are self-interested. The importance of this distinction is taken up again in Chapter 8.

This chapter and the two preceding it have focused on the individual. Some mention has been made of the role of socialisation and the internalisation of social norms, but links beyond this have not been discussed. Much has been made of the role of information and its avoidance. This thesis is primarily concerned with the helping or not of distant others and information regarding distant suffering is not, by definition, available first-hand. Given that information passes through others, and the importance of social norms, the discussions of individual responses have to be extended to place the individual behavioural models in a social context; this extension will be taken up in Chapter 7.

The discussions thus far have provided a range of arguments as to why the maintenance of self-image model and the balance model are best suited to

provide a theoretical framework from which to develop the discussion of helping distant others. The previous chapters, having drawn on a wide range of literature, have helped to narrow down the number of options to consider for use in the practical component of this thesis.

The theoretical questions are complex and no single study provides comprehensive answers. While the literature has supported a focused discussion of the theory, the studies used have not always been designed specifically with the proposed practical application in mind. The core of this thesis is the questions of when and why people would help distant others. Both the maintenance of responsible self-image model and the balance model provide arguments relating to the why; both also provide a theoretical framework within which to examine the question of when.

Examining real-world situations relating to responses to distant need is important. Such examinations would, as mentioned above, require a consideration of the social context. There is a wealth of literature which can be drawn on to expand these two models to consider the wider environment. Chapter 7 does indeed draw on this literature and expands the theoretical discussion; in so doing, it makes consideration of real-world events a possibility. It does, however, become difficult to isolate individual motivations within such discussions, as so many other factors are introduced.

Before moving on to the consideration of the social context, it is therefore useful to examine the practical question of when people help. A study was conducted with this in mind and is reported on in the next two chapters. The research conducted aims to consider individual motives to help and so focuses on hypothetical situations rather than real-world events. The conducting of research in this area also allows for the testing of some of the hypotheses raised in the discussion thus far. It is not, however, the primary purpose of the research to answer the unanswered theoretical questions, as that would not be

possible within one study. The aim is, primarily, to examine the question of how individual motives influence when those individuals will provide help and, secondarily, to make a contribution to the evidence regarding the theoretical questions.

The following two chapters outline the method used and the results of its application in an effort to examine these questions and to make this contribution. The empirical research presented aims to consider attitudes towards information, emotional responses and rationalisation, as all are central to the discussion, and will focus on these issues as they relate to distant suffering. Following on from the empirical research, the discussion will return to the development of the models and the inclusion of the social context in Chapter 7.

Chapter 5: Questions and experiment methodology

5.1 Introduction

When and why people help distant others are difficult questions to answer. The previous chapters have approached these questions from a theoretical standpoint. They have examined alternative behavioural models that have sought to provide a framework within which these questions can be addressed.

In an effort to make the discussions less abstract, a research project, which is explained and presented over this and the next chapter, was undertaken with two aims:

- To use the framework suggested by the balance model to examine questions of when and why individuals help distant others.
- To contribute to the evidence that is used to differentiate between the different models identified.

The balance model was used as the framework because, with the exception of the maintenance of a responsible self-image model, it seems to fare better in terms of predicting behaviour than the other models. The balance model was used, rather than the maintenance of a responsible self-image model, as it has been argued to better capture the decision-making process. It points clearly to the role of emotional responses and the possibilities to manipulate them. A maintenance of self-image model could be constructed to do the same and similarities will be noted, but using such an approach would involve a somewhat contrived set of assumptions.

The use of the balance model as a framework becomes an examination not only of when people help but also of the merits of using such an approach. This links closely with the secondary purpose of this research: to make a contribution

to the evidence used to differentiate between the models identified. The research project was also designed so as to be able to examine some of the key issues raised in the discussions thus far, such as attitudes towards information.

Specifically the experimental research addressed the following questions:

1. What characteristics of distant others and their need influence responses to them?
2. Are there emotional responses that place a pressure to help or not help?
3. Is there a pressure to act responsibly distinct from the pressure associated with relieving one's own feeling of sadness?
4. Do individuals engage in self-manipulation to avoid helping and how does this relate to attitudes towards information?

The first two questions relate to the application of the balance model to the question of when and why people help distant others. Questions 3 and 4 relate to efforts to distinguish between the models.

In aiming to design an approach that would facilitate collection of information on the way in which attitudes towards others' health are constructed, literature on past efforts to gather similar information was reviewed. These past efforts revealed that, in general, previous work has focused on attitudes towards local health priorities, often with the view to setting them. Little work has been done on identifying attitudes towards the health of distant others from the point of view of measurement.

A selection of studies considering health priorities, highlighting what lessons for the current research were found and where the differences and need for alternative methods arose, is presented below. Two other bodies of literature further helped in the design of the methods presented in this chapter: the

literature on charitable giving and the economics literature on giving in general. Both of these have already been discussed and such discussions will not be repeated in detail here, although relevant points that contributed to the design of the methods used will be mentioned. An outline of the research method is provided following discussion of lessons from the health priorities literature.

An experimental design, which allowed for both the theoretical tests and the practical data collection to be conducted simultaneously, was used. The focus of this chapter is a detailed discussion of the design and analysis plan. The results of the application of this design and analysis are the concern of the next chapter.

5.2 Health priorities literature

The question of what people value in the health of others, and why they value it, relates closely to the literature on public preferences for health priorities. It does, however, differ from simply determining public priorities, as it is necessary not only to consider what the priorities are, but also on what basis they are formed. The question also differs from existing work in that its focus necessitates the explicit detachment of participants from the groups being considered. In previous studies, the question of health prioritisation has generally been dealt with – at least implicitly – in the context of the participants' community or country and the bulk of such literature relates to developed countries. Nevertheless, a number of previous studies do provide useful insights and raise important issues for consideration in the design of the methods used in the examination of the questions at hand.

A number of studies have been conducted with the aim of determining what the general public consider to be health priorities. Kothari presents a review of the methods used in soliciting public preferences for health care priorities, noting that at the time of writing the methods were in their infancy (1999). Her review

examines a variety of methods used in a number of developed countries in terms of their internal validity, generalisability, reliability and objectivity. A number of key issues were raised through this process, notably the need to define who are considered as 'the public', the importance of considering the trade-off between cost and accuracy, and the instability of preferences. Another evaluation of methods used has been carried out by Ryan et al (2001). This review examined both quantitative and qualitative techniques, as well as investigating the importance of public opinion in decision making.

Both reviews raise a number of methodological considerations, particularly with regard to the instability of preferences. This was useful in the design of instruments in the present research. The focus of these reviews, however, is on studies that aim to elicit preferences rather than to understand their basis. It is in the understanding of the basis of preferences that difficulties similar to those experienced in the research occur. A number of different methods have been used to address this issue.

In trying to understand the basis of preferences, a number of reasons for variations in the value attached to health benefits have been raised, including age of the beneficiary, perceived liability of the patient for their condition, and the severity of the condition. Anand and Wailoo report on the results of a self-administered, anonymous questionnaire that sought to examine the basis for public priorities regarding the rationing of health services (2000). A number of factors that might affect rationing decisions were covered, including questions of fault, and socio-economic status. A commonly-used approach was undertaken, with respondents being asked to prioritise between groups on the basis of the information they were given. To address the question of fault, respondents were asked to rank groups with the same illness but with different routes to infection. This approach assists in identifying reasons, such as considerations of fault, but does not delve any further.

An approach that does disaggregate further is that of Tsuchiya et al who examined in detail one of the underlying causes of differential valuations of health benefits, namely, age (2003). They discuss the various arguments that can be used as justification for differential weights being attached to health outcomes for people of different ages. Their study consisted of three rounds of interviews. A first round, with a small group, was carried out to inform the design of the subsequent two rounds, which were used in the analysis. The first round aimed to categorise the different arguments for ageism, so that rounds two and three could be designed so as to examine the relative importance of each. Within each round, respondents were asked to rank groups, each group being characterised by differences in the age of beneficiaries, according to importance for treatment, with the outcome of the treatment being constant across groups. This allowed the authors to investigate the preferences of individuals but not their intensity. In later rounds, respondents were asked to give reasons for their rankings, which were grouped into the pre-defined categories from round one. This, combined with variations in the outcome, allowed for the examination of the importance of different justifications for ageism.

Neuberger et al looked at the importance of various factors within the context of the allocation of donor liver grafts and, among other things, investigated how responses differed across different groups (1998). Eight case studies were presented to participants who were asked to allocate four donor livers and to identify the least deserving case. The eight case studies differed in a number of respects, including age, misuse of drugs and alcohol, the likely outcome, and their time on the waiting list. Participants were also asked to rank which factors they considered to be the most important in allocation decisions. To examine how the responses differed depending on who the respondents were, the participants were drawn from the general public, as well as samples of doctors and gastroenterologists.

While all the above studies involve ranking, an alternative method is to pose a series of choices. While not directly health care related, Johansson-Stenman and Martinsson conducted an experimental study relating to the differential value placed on lives saved (2003). Their study considered measures related to road safety. Respondents to a postal survey were asked to make choices between seven pairs of different interventions. The interventions differed with regard to the ages of those saved and, in order to consider responsibility, they also differed in terms of whether the measures saved drivers or pedestrians. This experimental design allowed for an analysis of the relative value placed on age and responsibility.

Once-off rankings or choices, such as the above, however, can be criticised on the basis that preferences are not stable and may change as a result of additional information or discussion.

Ubel examined the stability of preferences with regard to allocations towards severe illness (1999). The study participants were issued with one of six questionnaires, which varied the scenarios given by providing reminders of illness severity, asking them to consider self-interest – that is the possibility that they were one of the patients – and explicitly offering an even split between groups. Ubel found that even small word changes in scenarios could alter preferences, suggesting instability.

Similar to the issue of wording is the instability associated with discussing conclusions and, as a result, hearing the views of others. Dolan et al investigated how people change their mind regarding health priorities following group discussions on the topic (1999). Participants took part in two group discussions: questionnaires were administered to them at the start of the first and at the end of the second group discussion, and the differences were examined. The sample was drawn from two lists of general practitioners and efforts were made to ensure representivity, particularly regarding age. The

group discussions were moderated by the researchers, but were so structured that their role was minimal in the first group and only facilitative in the second. The first round considered how priorities should be set and the second focused on ranking of hypothetical patients. This method of repeating the questionnaire is potentially useful in examining how framing and information affect preferences and rankings but, more importantly, it shows how discussions alter views.

Cookson and Dolan, drawing on the same study described in Dolan et al above, aimed to determine the ethical principles behind the rankings of the four patients conducted in the second round of discussions (2000). The authors note that there are various arguments about what ethical principles should govern rationing and they aimed to identify which ones the public support. During the second round of group discussions, participants, given a minimal description of the condition and treatment of each one, were asked to determine which of four patients they would treat. Minimal details were given so as to avoid a focus on the particular case, as opposed to the issues. The discussions were then directed to issues at a macro level concerning general principles. During the discussions, 'principles' (general rules) and 'factors' (specific aspects) were recorded and the number of times they were mentioned noted. The ranking and recording of principles and factors were then compared to a number of ethical standpoints raised in the literature and those which were echoed in the participants' views were identified. Efforts were made to avoid cases that might have been influenced by recent media coverage. This was done to follow the approach of giving minimal information so as to avoid undue focus on the specifics of a case, as opposed to general principles; others, however, have taken the opposite approach and sought to focus on providing much more detail.

Lenaghan presents the results and discusses the implications of the use of citizen juries as a means of eliciting public values on health priorities (1999).

These juries differed from the discussion groups described above in a number of respects. They concentrated on much more specific issues, were more in-depth, lasted longer, and participants were given more information. Such an approach may be more useful when considering published cases and could be combined with the repeated questionnaire to examine how viewpoints change. The differences between this approach and the group discussions previously mentioned raise the question of what views are important. If you wish to reflect public opinion, then opinions should be measured before any additional information is presented. If, however, these preferences change with information, then presumably it should be provided, as the result would be what opinion would be if information were made available; though once information is provided it is no longer currently-held public opinion.

The above literature suggests a number of options for investigating priorities on health care. The purpose of the bulk of these research projects is, however, fundamentally different from the questions being investigated here, which limits the transferability of methods. The purpose of the above has typically been to examine public preferences in order that policy formation may consider these. The research questions of this research simply centre on what they are. These may seem at first glance to be very similar questions, but they are not. Firstly, it has been suggested that people can force upon themselves a different set of values, depending on their role. Therefore, asking what they feel the state should prioritise is not the same as asking how they would like to see the money spent if it were to reflect what they wanted. This raises a further difference: while it may be appropriate to discuss opinions or conduct hypothetical rankings, if the purpose is to discuss priorities for public policy, it may not be so appropriate for measuring personal values. An individual may say that some cause is very important, but this does not mean that they will actively provide support.

A very useful point arising from the above literature, but which also highlights the differences, is the role of stability of preferences, which came across as key. What is important for policy makers – public opinion, considered public opinion, or informed public opinion? It would seem that they do differ, whereas the research interest of this investigation is on what they are and why they would differ with information.

Although there are a number of obvious differences, the literature on health priorities does provide some pointers. The use of scenarios and the implied ability of participants to engage with them suggests a route to investigating the practical questions of how people value distant others. The literature also suggests some candidate variables to examine as possible underlying causes of differential valuations, such as the age of the recipient and perceived fault. Many more can obviously be drawn from the psychology literature discussed in previous chapters. The major difference, however, remains that the literature on health priorities has also been exclusively concerned with local priorities.

The literature on charitable giving discussed in the opening chapter dealt more directly with people's attitudes towards distant others. The research and associated methods have, however, focused not on the underlying causes or processes of valuation, but rather on differences in the values attached to causes and how well alternative methods work in gaining a response. Similar approaches to those discussed above have often been employed in regard to varying information provision and, in the case of charitable research, the method of information provision, to identify how they affect behaviour. Studies on this topic often involve the individuals being asked to donate their own money. This makes the results far more real and personal, but typically these methods can only be used to examine one change at a time, due to the scale of the design, which usually involves a full-scale appeal for funds. Moreover, they cannot be easily followed up with questions on motivations, as they involve real

campaigns not surveys, whereas the surveys on these issues do not usually involve real money.

As has previously been discussed with regard to the economics literature, economists often set up experiments that use real allocations, in the sense that they involve real money, albeit in a laboratory environment, so that they can examine personal values rather than values associated with role play. It would seem that such approaches can be used to put the individual into a situation where they are making decisions that affect them directly, unlike in the health priorities research, while at the same time they do not need to be on the scale of the charitable-giving literature.

The economics literature has, as discussed previously, been primarily concerned with answering theoretical questions relating to behavioural models. The experimental economics approach has not been widely used as a means of collecting data on practical questions. By combining the experimental approach with the use of varied scenarios, data on how very specific changes in scenarios can influence behaviours involving real, personally-held money can be investigated. Furthermore, the experimental conditions allow for follow-up questioning that provides for deeper probing, linked to real allocations. As experimental designs are generally set up to test theoretical questions, their use for practical data gathering provides an ideal opportunity for the combining of the practical data gathering with the theoretical testing required.

5.3 Research design

The research was conducted using an experimental approach, based on a computer-based survey. The research combines the approach of varying scenarios with the experimental environment involving real allocations. The process essentially involved the use of the dictator game, in which the participant is asked to divide a sum of money between themselves and a passive

recipient. In this case, the passive recipient is a potential beneficiary of charitable assistance. Each participant was asked to make a number of allocations; the allocation decisions were then followed up with questions relating to how the participant felt about the scenarios, or rather cases, as they were called in the survey.

The following discussions outline the details of the experiment/survey and define the hypotheses relevant to each section, which link back to the research questions outlined in the introduction.

5.3.1 The sampling

Since the interest of this research was in the value of others' health when those others are distant, the survey was conducted only among citizens of European Union (EU) member states and the cases referred only to individuals outside of the EU. As the purpose was to investigate how individuals responded to the needs of different distant others, it was not necessary to sample a highly-varied participant group, but rather to ensure that participants were asked about a variety of potential recipients. Furthermore, there was no intention of generating generalisable results. It was not the purpose of the survey to say that this is what citizens of EU member states value; that would take a far larger-scale research project. The purpose was to examine the way in which the characteristics of others affect allocation decisions among a sample of people who are distant to them; in other words, to ask when and why this group of people would help those distant to them.

Since the research simply required a group of EU citizens, it was considered acceptable to take the easiest option, that of sampling students. This had the added benefit that students are on the whole poorer than working adults, therefore the money involved was likely to be more important to them than it would be to wealthier participants. This arguably makes the allocation decisions that they were asked to make similarly more important.

The survey was implemented at two universities. Participation was voluntary in both sites although recruitment methods differed. At the first university, participants were given the option of either participating during a class they were attending or of leaving the class early. At the second site, an email was circulated among students asking for volunteers to participate in a paid study on EU public opinion. The approach used to control for the implications of using a different method to recruit participants in the two sites is outlined in the analysis section of this chapter.

There was no obvious guide as to the appropriate sample size for the study. With very few similar studies having been conducted, it was difficult to predict what level of variability in allocation decisions to expect; without this information it was not possible to determine the sample size necessary to detect them. The final sample size was therefore determined on affordability grounds. Funding allowed for the survey to be run with a maximum of 100 participants. It was decided to allocate 10 of those places to a pilot of the survey and to aim to collect a sample of 90 for the final survey.

5.3.2 The instrument

The entire survey process was computer-based to ensure anonymity of responses. As discussed in previous chapters, a motivation for giving may be public appreciation. It would be particularly difficult to separate out the effect of benefits received from giving in public to different causes from the value of different causes to the individual, but maintaining anonymity removes this problem. During the course of completing the survey, each participant was provided with a unique identity code that they would later use when collecting payments.

Software was specially developed for the purposes of the study in Microsoft Visual Basic Version 6⁴. As far as possible, responses were obtained by selecting from drop-down menus to avoid out of range answers being given. The only point at which participants were required to type in a response was with regard to their age.

The software opened with a number of instruction pages explaining the survey. Participants were required to click through them before commencing. All of the screens seen by the participants are replicated in the appendix.

After the opening instructions, the survey consisted of three sections: basic socio-demographic questions, allocation decisions, and follow-up questions. Each of these sections is discussed in detail below.

5.3.2.1 Basic socio-demographic questions

The basic socio-demographic questions covered the participants' age, sex, religion, financial situation and asked if they had children or siblings. One of the major themes that has emerged in previous chapters has been the perceived importance in the literature of similarity of recipient to the potential donor. While the issue of similarity is addressed directly in the follow-up questions, the participants' age and sex were also requested to see if similarities in these respects play any role. The participants' religion was requested, as this may arguably play a role in determining participants' attitude towards others. It was noted in the psychology literature and the charitable-giving literature discussed previously that having children plays a role in the giving decision, so this question was included. Given that the sample was to consist of students, the incidence of children was expected to be low. The question was, therefore, supplemented with one on siblings, although there is little theoretical or previous practical motivation for this.

⁴ The program was designed by the researcher but the programming itself was contracted out and conducted by Lincoln King

In many studies on charitable giving, income has been shown to be related to donation levels. Originally a question asking for the participants' income was going to be included. After further consideration, however, this question was replaced with one asking for the respondents' self-assessed financial circumstances. In line with the general argument of the thesis, what is important in the allocation decision is how well off an individual believes themselves to be, rather than their actual income. In retrospect, it may have been appropriate to ask both questions.

The sample was drawn to try to ensure that the participants were responding to distant others. Knowing that a not inconsequential proportion of EU citizens, or their parents, may originally come from outside of the EU, it was necessary to control for the possible effect of this. Participants were therefore asked if they had close family ties with Eastern Europe, Africa, or Asia – the regions in which the cases were located.

Having your financial situation, your religion, and your close family ties on your mind may well change your responses. To avoid this effect, the basic socio-demographic questions were divided into two sections. The first questions participants were asked were basic demographics: age, sex and if they had siblings or children. These were placed at the opening of the survey to familiarise the participants with the interface while responding to simple questions. The questions on religion, financial status, and family ties were asked at the end of the survey. The final wording of socio-demographic questions is given in the appendix.

5.3.2.2 Allocation decisions

Drawing on the experience of previous researchers, the use of varying scenarios was used as the basis for examining allocation decisions. The approach used, however, differed substantially from that which is generally

used in the health priorities literature discussed above – as previously mentioned, the method draws on the experimental economics literature, presenting participants with allocation decisions involving real money.

The allocation decisions were made within the context of the dictator game, which was mentioned in Chapter 2 and earlier in this chapter. To recap, the dictator game involves a participant being asked to divide a sum of money between themselves and a passive recipient, however they see fit. At the end of the game they receive their allocation and the passive recipient receives theirs.

This game is typically constructed so that another anonymous member of the group participating is the passive recipient. In the version of the game conducted in this study, the recipient was a person described to the participants in a scenario presented to them. The scenarios were concerned with individuals in need of assistance as a result of poor health. All the scenarios referred to distant others in the sense that their nationality was given and they were never from an EU member state. While the scenarios, or cases, as they were referred to in the tool, were clearly hypothetical, the participants were informed verbally before the survey, and in the introductory text, that the cases had been linked to charities working with as similar as possible individuals to those described. Participants were asked to act as if the money would go to the person described.

Examining the impact of varying the recipient has been done before (see for example Eckel and Grossman, 1996). This experiment, however, differed in a number of ways, including incorporating approaches from elsewhere in the experimental literature. Firstly, each participant was asked to repeat the allocation for nine cases. For each case, they were asked to divide 10 pounds⁵

⁵ After discussions with fellow students, 10 pounds was deemed to be a fair amount of compensation for students' time, given the going rate of student pay for casual work of between 6 and 10 pounds an hour. As an added incentive they were also given a 5 pounds show up payment.

between themselves and the person described in the case. They were informed verbally prior to the survey, and in the introductory text, that, on completion, one case would be selected at random and that they would receive the payment due to them from that case, while the balance would be donated to the most appropriate charity that had been linked to the case. In this way, every case had the possibility of being real. The approach of repeated decisions and the final selection of one is not uncommon in the experimental literature (see for example Andreoni and Miller, 1998).

The second difference from previous varied-recipient work was the manner in which the recipients were varied. Previous work had considered, for example, the difference in allocations when the recipient is another member of the group or when it is a charity (Eckel and Grossman, 1996). In this survey, all the cases/scenarios referred to distant others in need of help because of their health.

The variation in the cases needed to be constructed in a very careful manner. If there were too many differences between cases, it would not be clear what any possible difference in response to the case resulted from. If the cases were too similar and varied only in one way, it would be obvious to the participant what was being tested and this might influence the results. For example, if a case varied only in terms of the recipient's nationality, it would be clear that this is what was being tested and, as the participant may not want to appear, even if only to themselves, to be influenced by this one factor they may alter their response from what it would have been. Previous work discussed earlier highlighted how opinions about oneself may affect allocation decisions (Dana et al, 2007).

To overcome the problem of isolating the cause without making the test obvious, participants were randomly divided into two groups. Each group faced a set of cases that were all fairly different from each other and so not making

any testing obvious. Each case faced by group one, however, only differed from the corresponding case for group two in one aspect, allowing for the testing of the impact of this aspect, similar to Ubel's (1999) small variations. That the two cases were presented to different groups of individuals presents a particular challenge for analysis, which will be discussed in the analysis section of this chapter.

Each case presented to the participants was therefore constructed to test a specific issue. The issues tested were those that arose from the literature and those which were of particular interest to the discussions in this work. The nine cases are presented in the following table. The first column is the version presented to individuals allocated to set one and the second column is the version presented to those faced with set two.

Table 5.1: Set and case differences

Set 1	Set 2
Case 1: Religion	
<p>An appeal has been made by a Mosque in Pakistan for a member of its community. Nazir is a 42-year-old man who is suffering from cancer. He requires treatment that he cannot afford. The treatment will increase his survival chances substantially.</p>	<p>An appeal has been made by a Church in Pakistan for a member of its community. Nazir is a 42-year-old man who is suffering from cancer. He requires treatment that he cannot afford. The treatment will increase his survival chances substantially.</p>

<i>Case 2: Context</i>	
<i>Simpiwe, aged 24, lives in Swaziland and has malaria. She lives in an area which is currently experiencing a food shortage and has high rates of HIV, malaria and cholera. A medical NGO is appealing for support to provide treatment that would improve her chances of survival greatly.</i>	<i>Simpiwe, aged 24, lives in Swaziland and has malaria. She lives in an area which has high rates of malaria, but is otherwise doing well. A medical NGO is appealing for support to provide treatment that would improve her chances of survival greatly.</i>
<i>Case 3: Possible fault</i>	
<i>A 50-year-old Kenyan man was injured in a car accident while driving home at night from a bar. A local charity is appealing for funds to assist in his physical rehabilitation, which would greatly improve his quality of life.</i>	<i>A 50-year-old Kenyan man was injured in a car accident while driving home at night from work. A local charity is appealing for funds to assist in his physical rehabilitation, which would greatly improve his quality of life.</i>

<i>Case 4: Age – adult to child</i>	
<i>Joseph, a 35-year-old Zambian, is suffering from an unpleasant skin condition that, while not fatal, is extremely uncomfortable. If left untreated it will likely last for up to five years, whereas with treatment he should recover in a few weeks. A charity is appealing for help to pay for the treatment.</i>	<i>Joseph, a 3-year-old Zambian, is suffering from an unpleasant skin condition that, while not fatal, is extremely uncomfortable. If left untreated it will likely last for up to five years, whereas with treatment he should recover in a few weeks. A charity is appealing for help to pay for the treatment.</i>
<i>Case 5: Sex</i>	
<i>A 55-year-old Indian woman working as a cleaner has contracted pneumonia. She is very poor and has asked a local NGO for support while she recovers. The NGO is raising funds so that they can provide the support.</i>	<i>A 55-year-old Indian man working as a cleaner has contracted pneumonia. He is very poor and has asked a local NGO for support while he recovers. The NGO is raising funds so that they can provide the support.</i>

<i>Case 6: Nationality / ethnic group</i>	
<i>A medical NGO is seeking support to provide treatment for an HIV+ Zimbabwean man, aged 34, who has recently progressed to AIDS. While treatment will not cure him, it will likely extend his life for a number of years.</i>	<i>A medical NGO is seeking support to provide treatment for a HIV+ Ukrainian man, aged 34, who has recently progressed to AIDS. While treatment will not cure him, it will likely extend his life for a number of years.</i>
<i>Case 7: Implied wealth</i>	
<i>Mary, a 47-year-old community worker, was hit by a car outside her office in Lusaka, Zambia. She has suffered internal injuries and needs surgery, which is not available in Zambia, and her church is raising funds for her to go to South Africa.</i>	<i>Mary, a 47-year-old corporate lawyer, was hit by a car outside her office in Lusaka, Zambia. She has suffered internal injuries and needs surgery, which is not available in Zambia, and her church is raising funds for her to go to South Africa.</i>
<i>Case 8: Age – young to old</i>	
<i>Afina, aged 22, from Romania, is suffering from tuberculosis. This illness can be cured, but funds are currently not available to provide treatment. Her local Church is appealing for help.</i>	<i>Afina, aged 72, from Romania, is suffering from tuberculosis. This illness can be cured, but funds are currently not available to provide treatment. Her local Church is appealing for help.</i>

Case 9: Emotive issue

*The following case occurs all the time; the money donated will go towards organisations providing such services. A Malawian woman has been **raped** and needs funds to be provided with a short course of drugs to reduce her chances of being infected with HIV.*

*The following case occurs all the time; the money donated will go towards organisations providing such services. A Malawian woman has been **injured in a medical accident** and needs funds to be provided with a short course of drugs to reduce her chances of being infected with HIV.*

The choice of variations was not only made to reflect issues raised in the literature. They were further selected so as to examine not only if the above variations prompted a difference in donations, but whether they prompted a difference in responses to follow-up questions. The basis for their selections will therefore be discussed more fully once the follow-up questions have been outlined.

It was originally intended that the participants would be asked to type in the amount that they wished to donate, so as to generate as close to a continuous variable as possible. This would, however, have slowed the distribution of monies after the experiment and have required large amounts of change to be carried around. It was decided that it would be easier and quicker to have a drop-down menu where participants could select how many whole pounds they wished to donate, from none to all. This again affects the analysis approach and will be discussed shortly.

To avoid any possible ordering effects, once a set was randomly assigned to a participant, the cases were asked in random order. It may be that cases asked

earlier would get systematically more or less, but random ordering was introduced so as to minimise the impact of this should it occur.

The process of collecting allocation decisions was done with two levels of analysis in mind: testing firstly between the two versions of each case for the impact of the single change; and testing across cases for the causes of differences more generally. So as to test not only for differences associated with specific cases but to try to explain variation, a set of follow-up questions were asked in regard to each case. Before these questions were asked, participants were prompted to make another allocation decision, but of a different sort. This involved a pay-off matrix and was designed specifically with the aim of examining some of the theoretical assertions. It will be discussed shortly but should be noted to occur as part of the allocation decisions.

5.3.2.3 Follow-up questions

After participants had made their allocation decisions they were asked a number of questions about each case. These follow-up questions were asked for each case in the same order as the cases were presented. On the screen showing questions, participants were provided with a reminder of the case they were considering.

In total, nine questions were asked for each case. These questions were designed to capture arguments in the literature and the theory put forward in this thesis as to why people respond to help others. A list of the questions is presented in the following table. Participants were asked to select a response from a drop-down menu of between 0 and 9, allowing for the use of a 10-point scale without the distraction of any double figures.

Table 5.2: Follow-up questions

1. How saddened were you by this story?
2. Does this story make you feel a pressure to help?
3. How much are they to blame for their current situation?
4. How much is it up to them to get themselves out of their current situation?
5. What do you think their level of suffering is?
6. Do you feel you have a responsibility to help people like this?
7. How similar would you say they are to you?
8. How much help do you think a donation of 5 pounds would be?
9. How urgent would you say this case was?

All of the questions are linked to arguments and issues raised in previous chapters. The sadness question was included so as to examine the argued link between an individual's own suffering from hearing of a situation and their response. The questions on felt pressure and responsibility were included as they are two attempts to capture the link, assumed in the balance model, to helping as distinct from sadness. Blame and means to get out of their situation, both of which have been discussed previously, link with responsibility and were therefore included. The similarity argument has played a major role in arguments around motivation for helping. The importance of urgency and perceived impact and suffering speak to the rationalisation of help, again an important issue in the discussions thus far.

5.3.2.4 Pay-off matrix

As mentioned previously, after having been through the first nine cases, participants were presented with an allocation decision of a different type. This final case had three versions. Participants in set one were presented with one of two versions, whereas participants in set two were all given the same version as each other, although it differed from both versions from set one. This meant that the third version of the case was weighted to be asked half the time, while each version in set one was asked a quarter of the time. This weighting was

deliberate, as the final version was more complicated and a larger sample was therefore considered appropriate.

The approach to this case was designed to examine the impact of involving others in the decision and the attitude towards information. The attitude towards the information component bears similarities to Dana et al (2007).

For the first version of case ten, presented with a probability of 0.25, the case was introduced and they were then asked to select one of the following two options:

Option A: You receive 8 pounds Oxfam receives 1

Option B: You receive 6 pounds Oxfam receives 5

In the second version, presented with a probability of 0.25, they were told that they were linked with another member of the group and that the final outcome depended on what they both selected. They were given a pay-off matrix and told that the first number would be their pay-off, the second would be their partner's, and the last would be the amount donated to Oxfam. The pay-off matrix appeared as follows:

		Partner Chooses	
		A	B
You choose	A	8, 8, 1	8, 6, 5
Or	B	6, 8, 5	6, 6, 5

Choosing A ensures that the participant gets 8 rather than 6 and leaves the determination of what Oxfam receives to the partner's choice. Choosing B in the

above matrix guarantees that Oxfam receives 5 rather than 1, but also guarantees that the participant drops from 8 to 6 as in the first version.

The above two versions were intended to replicate the impact of the increased number of potential helpers discussed previously. The final version of the case was presented with a probability of 0.5. The case was the same as the second version with only one addition. Participants were told that they were responding second and could find out what their partner had selected. To see what their partner had selected, they had to click a button, which they could simply ignore if they wished, thereby avoiding any need for an active refusal. This version was intended to examine attitudes towards information. The participant could choose to see their partner's response before or after responding themselves and could change their response after seeing.

The choices of participants were in actuality not linked and so not included in the random selection of the case that was played out. If participants opted in the final version to see what their partner had selected, they were always told that their partner had decided not to help. Participants therefore believed that the case could possibly be real, but it was not. The analysis of this aspect of the tool, as well as the other aspects, will be discussed in detail shortly.

5.3.3 The process

In order to summarise how the above aspects of the tool fitted together and how the implementation occurred, the following section outlines the process followed.

The survey was implemented in groups in university computer rooms. The programme was started and set up prior to the participants' arrival. A sheet of paper was placed over the key board asking participants not to begin until asked to do so.

Once participants arrived, a standard welcome and introduction was given, outlining the process and the payment methods. A summary of the introduction is provided in the appendix. The introduction included a clear message that there would be no linking of the answers with any individual and that payments would be made in sealed envelopes and claimed on the basis of a code generated by the programme. They were told that some of them would be presented with a pay-off matrix. They were asked to turn over the sheet covering the key board to find instructions explaining how to read a pay-off matrix. These were run through and the participants were then asked to begin.

The programme began with a series of information pages, all of which are replicated in the appendix. The programme then assigned each participant one of two sets of allocation cases. These sets were selected with a 0.5 probability each. For those who were allocated set one, a further allocation of either version 1 or 2 of case 10 was made, with a 0.5 probability. Those allocated set two were all allocated the same version of case 10.

The participants were asked to enter some basic demographic information. Thereafter they were asked to make their allocations for each case, asked in random order. Following their allocations, they were given a claim code and asked to record it on the claim forms provided. They were then asked the follow-up questions for the 9 cases; there were no follow-up questions for case 10. Finally, they were asked some socio-demographic questions and reminded of their claim code. The results were saved and could be immediately accessed from another computer by the facilitators who placed the amount of money due to the participant, based on the case randomly selected by the computer, in an envelope marked with the participant's claim code. Once all participants had finished, monies were distributed according to the claim codes.

5.3.4 The pilot

A small pilot study was conducted to test the use of the tool. The pilot was carried out at the London School of Economics with a small group of 7 students. After they had completed the survey they were asked to fill in a short questionnaire.

Following the pilot and a review of the responses to the short questionnaire, a number of changes were made. These related mainly to time issues. Originally there were 12 follow-up questions for each case, but participants in the pilot felt that it was difficult to concentrate on answering the same questions 9 times when each case had so many.

The questions included two that asked, firstly, if they would change their donation having been through the questions; if they responded that they would, they were given the option of doing so. This option was almost never taken, so it was dropped. The second question asked if the participant could imagine the situation that was described in the case. This question had the weakest link to the literature and, as the number of questions was an issue, it was also dropped.

In the final version of the tool, screens encouraging the participant to continue to pay attention and, towards the end, reminding them that it was almost done were introduced. These were inserted to break the repetition of the screens and to encourage concentration.

There was also some confusion regarding the pay-off matrix. For this reason the information sheet was extended and an explanation was included in the introduction.

5.3.5 The basis for content hypotheses tested

The research questions, as framed in the introduction, are too general to be tested with a few direct hypotheses. Rather what is necessary is the testing of a number, in the hope that collectively they will inform a discussion that will shed light on the answers to the broader questions.

Once combined, the follow-up questions, the socio-demographic questions, and the allocations allow for a number of hypotheses to be investigated. These hypotheses can be grouped into related batches.

The first group involves testing whether the changes to cases outlined in table 1 above led to significant changes in donations and feelings associated with each case. All case variations were designed with the intention of examining the difference in donation associated with the change, but certain variations were selected to also address specific questions relating to the link between feelings, socio-demographic characteristics of respondents, and actions.

Case one varied the religious association of the appealing institution, while remaining silent on the religion of the recipient. An underlying question being asked by these variations is whether participants take the bait to assume what is only implied, as this suggests motivation to do so. Why they might have such a motivation speaks to the fundamental questions of the thesis. Religion was included as, once linked with socio-demographic questions, it allows the question of whether mentioning a religious institution associated with the respondent's own religion leads to higher donations or a greater pressure to help.

Case two concerned context: while the recipient was the same, one version painted a poor picture of the background context. It was thought that this might lead to respondents feeling that there was no point in helping and that the impact would be minimal. The specific interest was then in whether this change

would lead to a significant difference in terms of responses given to the marginal impact question.

Case three, dealing with implied fault, is a particularly important case. The case does not make any reference to fault in either version. The mention of a bar allows the respondent the opportunity to jump to a conclusion. The man could well work in a bar, in which event the versions would be identical. If the mention of the bar does increase the perception of fault and lowers donations, it lends weight to arguments that it is not only the potential benefit of the recipient that is of concern to the helper.

Cases four, five, six and eight all deal with demographic characteristics. The inter-relationship between these changes and reported feelings is potentially very informative. Much has been made in the literature of these differences and how they lead to variations in donations. Examining which of the feelings are significantly different when these variations are made, may shed some light on why these characteristics are so important.

Case seven implies a difference in wealth. The individual who, it is implied, is wealthier is expected to receive less. The case was included to see whether this occurrence is associated with a perception that it is their responsibility to get out of the situation they find themselves in, or because the respondent thinks their money will have less impact.

Case nine introduces the emotive issue of rape. While difficult with a small sample, the interest here was not only in the difference between the versions of the case, but whether that difference meant more or less depending on the respondent's sex.

The case variations allowed for the examination of specific issues. The next batch of hypotheses relates to the association between the feelings questions,

the donations, and the socio-demographic characteristics of the respondents in general and between the feelings questions themselves.

The first batch of these examines the evidence for some of the theories on motivation for helping that have been discussed.

If the respondents help in order to reduce their own sadness, then the sadness question should be positively associated with the donated amount. This is tested as the first hypothesis of this section.

If the respondent helps more when they believe the help will be more beneficial, there will be a positive relationship between believed marginal impact and donations, and possibly also between urgency and donations, both of which were tested.

A central argument of this thesis is that individuals feel a differential responsibility to help in different circumstances, regardless of either of the above two arguments. A key hypothesis tested with these data is that there will be a positive association between pressure/responsibility reported and amount donated, even after controlling for the impact of the above two influences - that is, it has a distinct, independent role.

The second batch of these considers the associations between feelings. In the balance model, feelings are what link assessments of situations to decision making. Using such a framework suggests that it may be useful to examine the associations between assessments of situations and the feelings generated. Drawing on the discussions covered in the previous chapters, and some intuition regarding the likely relationships, a number of these relationships can be hypothesised.

Sadness, it is hypothesised, is positively related to perceived suffering and similarity and negatively related to perceived fault of the potential beneficiary in getting themselves into the situation outlined in the case.

Similarly, responsibility/pressure is positively linked to suffering, urgency of the situation, expected marginal impact, and similarity, but negatively to the perceived responsibility for the beneficiary to get themselves out of the situation described.

Related to this is the hypothesis that similarity and suffering act through sadness and responsibility and do not have an impact independent of them.

The final batch of this section deals with the socio-demographic characteristics: testing if the sex of respondent, having a sibling, religion, and financial situation lead to significantly different donations.

Case ten, being a little different from the other cases, obviously relates to its own hypotheses. The two versions of the case in set one - in one of which the respondent decides alone and in the other they decide in conjunction with another but with no option to know what they decided - were designed to test the traditional hypothesis relating to number of helpers. That is, to examine whether, as the number of potential helpers increases, so the probability that help will be provided decreases. This hypothesis has two parts: firstly, that the individual will help less when there is another potential helper; secondly, that the probability of help being provided at all will fall with the increase in the number of potential helpers.

The version of case ten, which was responded to by those allocated set two, was designed to examine attitudes towards information. If the respondents were interested in maximising the net benefit to themselves and Oxfam, they would all want to know the decision taken by their supposed partner. If,

however, they placed little value on benefit to Oxfam, they might not be interested in the information and would simply maximise their own benefit. If would, however, make little sense to select to help Oxfam but to refuse the information. The hypothesis that no one would take this option, was tested.

For those individuals who take the option of information, the case reduces to the version in set one, where the decision is entirely up to them. It is then tested whether those who take the information select differently from those who were faced with the simple version of deciding alone in set one. If they are found to decide differently, they are clearly considering more than their direct benefit and the benefit to Oxfam.

The above hypotheses are formally presented in the methods appendix. The following sections discuss the techniques used to analyse the data to evaluate the hypotheses outlined.

5.4 The analysis

The following chapter presents the results of the implementation of the process as previously described. This section outlines the approach used in examining each aspect and the motivation for the choice of such approaches. The analysis has three major components: the analysis of within-case variations, that is, between the version presented in set one and the slightly different version of the same case presented in set two; the analysis across cases; and the analysis of the somewhat different case ten. The first two components test different hypotheses but require very similar approaches to analysis and so for the purposes of method discussions are often discussed together. Before dealing with the methods for each component, a few general methodological issues in regard to the analysis are discussed.

The pilot study was conducted under the same conditions as the final survey. There were, however, differences in the number of follow-up questions and the manner of recruitment. As mentioned, more questions were asked as part of the follow-up section of the survey, so the experience was different; as a result, it was felt that even the responses to questions that were retained in the final survey were not comparable. For this reason, only follow-up questions from the main survey were included in the analysis. The allocation questions were, however, exactly the same, so these could be included in aspects of the analysis that did not require a link to the follow-up questions. If included efforts would have to be made to control for possible bias resulting from the different approaches used in the recruitment of participants. This problem was similarly encountered when pooling the results from the two different universities. As the manner of recruitment differed between the universities, it may be that one site would systematically donate more. Where controls of this nature were required they are outlined.

As has been described, the survey used an approach of presenting multiple cases to each participant and randomly selecting one to be played out. This was done to increase the number of recorded donations without having to increase the sample size and associated costs. The disadvantage of this approach is that, while the number of records is increased, they are not independent of each other and are essentially clustered in groups of nine for each individual. Such clustering can be dealt with and, where it is an issue, the means for controlling are discussed. Clustering is a disadvantage in some respects, but it can also be a great advantage. Comparing donations across individuals is a difficult way of assessing the importance they attach to a case. One individual may donate more than another, but this does not mean that they necessarily place more value on the case; it may just reflect that they value money less or they are generally more generous. Having clustered responses allows the relative donations to cases to be compared, which can at times be more informative. The cluster similarly allows for the calculation of average

donation that may provide a suggestion of generosity, even if still limited by the potential difference in the marginal value of money to the individual, which could anyway be seen as a determinant of generosity rather than a source of bias.

The above issues cut across the sections of the survey; the following discussions now address the specific analytical tools employed for each component of analysis. These components can be grouped into two areas, each area being characterised by the need for similar analytical tools: assessment of donations and associated feelings, and decisions relating to case ten. In the analysis itself the division is three-way as the assessment of donations and associated feelings are dealt with separately for within case and across case analysis. As the methods used are very similar for within and across case analysis they are treated here as sub-groups.

5.4.1 Assessment of donations and associated feelings

This area comprises a number of different components brought together by the need for a similar approach to analysis: differences within cases; determination of donations and feelings across cases. Differences within cases refer to the comparison of donations and feelings associated with cases in set one, with corresponding cases in set two. Determination of donations considers which factors determine the level of donations; determination of feelings similarly refers to investigations as to what prompts specific feelings. Both of these analyse consider determination across cases.

For the purposes of this area of analysis, the data consisted of the donation data, the coding of the cases, demographic characteristics of respondents, the site at which they participated, and responses to the feelings questions.

5.4.1.1 Differences within cases

This section of the analysis sought to examine whether the subtle changes in cases presented in set one, as opposed to set two, altered the level of

donations and reported feelings. At first glance, it may seem appropriate to simply compare the average donations for each case in set one with the average of the corresponding case in set two. This, however, would have ignored the fact that it was not simply the same set of respondents responding to different cases, but rather different respondents responding to different cases. The random allocation of respondents to one set or the other would be expected to reduce the impact of this situation, but it can be further addressed through the use of regression analysis to account for further potential differences across individuals.

To conduct regression analysis requires the selection of an appropriate model. In this analysis, the interpretation of what the data represents plays an important role in determining model choice.

If the data had been collected as originally intended, it would have provided a variable 'donated' for each case and for each individual, ranging from 0-10. This could arguably have been treated as a continuous variable and examined with ordinary least squares (OLS). The data on donations from both sets could then have been pooled and a regression conducted for each case, with 'donated' as the dependent variable and participant characteristics as explanatory variables, with the addition of a dummy variable identifying which set the response was from. If this dummy were significant, then the subtle change could be considered to have potentially led to a significant change in the amount donated.

There would, however, be a number of problems with the above approach, not least because only whole pound donations were made. Firstly, it was not the intention of the study to determine what portion of 10 pounds individuals would donate to particular cases, but rather the level of help, as indicated by donations, that cases prompt. Donations are considered as a proxy measure of some underlying wish to help. Considering the data in this way presents a

problem with using unadjusted OLS. The problem occurs at the extremes of zero and maximum donations. It is not possible, from the data collected, to distinguish between cases where a respondent donated 10 pounds and would have donated only 10 pounds even if they had been given more, and a case where the respondent donated 10 pounds but would have donated more if they had been able to. The data are essentially censored at 10. Less obvious, but arguably the case, the data are censored at 0 as well. This is because a case to which an individual is indifferent, and one which they would positively dislike helping, are both associated with a zero donation. The latent unobserved variable of level of help may well run into the negative range. There are approaches to addressing data of this type. Using a tobit model, this two-sided censoring could be adjusted for in the analysis.

The second problem with using OLS is again associated with the relationship between the observed 'donated' variable and the unobserved latent variable of motivation to help. It is possible that the relationship may not always be linear. In other words, the increase in willingness to help associated with an increase in the amount donated from 4 pounds to 5, may not be the same as the increase associated with a change from 8 to 9 pounds. To consider this possibility would be to consider the data as ordered data. The argument for considering it as ordered data is strengthened by the change from a continuous variable to allowing only whole pound donations. If the respondent had wanted to donate 1.20 pounds, they would have had to choose between 1 and 2 pounds. The collected variable, therefore, is essentially an ordered representation of donations between 0 and 10 pounds. This problem can be remedied by using an ordered regression model, which has the advantage of dealing with the extremes of zero and 10 in such a way as the censoring is not a problem. Censoring is not an issue, as the categories at the extremes are treated as open-ended - that is, 10 or higher and 0 or lower.

An ordinal regression model was, therefore, selected as the most appropriate model for the analysis of within-case variances. The model included the amount donated to a particular case as the dependent variable. The explanatory variables were made up of variables detailing the characteristics of the respondent, including where they responded, a dummy variable indicating the set from which the case was drawn, and the average amount donated by the individual to other cases. This last variable was included, as without it the analysis would not consider the impact of more or less generous individuals who are such independently of the specific case under consideration. The dummy variable was the main variable of interest. If the coefficient of this variable is significant, the hypothesis that the subtle change led to a change in donation cannot be rejected. The interest was, therefore, in the significance and direction of the impact. Technical details of the approach used, and the econometric issues that require consideration, are discussed fully in the technical section of the methods appendix to this chapter.

Differences in feelings reported for each case were similarly compared with feelings reported for the corresponding case in the other set. The feelings data is again ordinal in nature. Although it is a scale, it is only intended as a representation of an underlying unobservable feeling and should therefore be considered ordered. An ordered regression model was, therefore, used for examining each reported feeling for each case. The explanatory variables were again the characteristics of the respondents, where they responded, and their average level – excluding the current case – of each feeling being examined. This final variable was included for the same reason that average donation was included when examining donations. Some people may have a tendency to report a feeling using a higher number, but what is important is this number relative to their average.

5.4.1.2 Differences across cases

The determination of donations section is the core section of the analysis, which seeks to examine factors and feelings that drive the determination of the amount donated. As this section again uses donations as the dependent variable, the above discussion is again relevant and the use of ordered regression analysis was employed for the same reasons. There are, however, some special methodological considerations that become relevant in this section.

The most important of these considerations is the clustering of responses. The design of the study results in the generation of data on nine donations for each individual. This increases the total number of data points hugely, but it is important to recognise that they are not all independent. They are independent across individuals but not within groups of individual responses. There are means of addressing this problem and they were employed; details are provided in the appendix.

The analysis, therefore, involves an ordered regression model accounting for clustering, with 'donated' as the independent variable. There are three groups from which independent variables can be drawn: the characteristics of respondents; the characteristics of the cases; and the feelings reported. These could all be dropped into a single regression, but this would complicate the analysis. Respondent and case characteristics should interact to generate feelings, so including all in one would be inappropriate. Furthermore, it is difficult to include case characteristics across cases, as the study was not designed to do this.

The within-case analysis discussed above aims to examine whether different case characteristics shape giving, but to do so across cases would be complicated by differences in cases not associated with the change being examined. For example, comparing cases to see whether those with women

recipients received more than those with male recipients would be complicated by differences in cases that have nothing to do with sex. It is arguably possible to control for these differences in a regression by coding other aspects of the case. This was done, but was inconclusive and the results will not be presented. The characteristics of the respondents themselves were also included in this regression but given the lack of variation within the group none of these characteristics were significant.

A regression was conducted, with the third group of variables, examining the relationship between feelings and donations. This regression required some transformation of variables. As has been mentioned, comparing individual responses to feelings questions can be complicated. That one individual ranks their feeling of sadness higher than another individual does not mean that they were indeed sadder. Inter-personal comparisons of feelings are very difficult, but fortunately not necessary. What is of more interest is the relative feeling prompted by one case as compared to another. To examine this, variables, measuring the difference between the feeling response for a particular case and the individual's average response to that feeling question, were constructed and included in the regression. This is slightly problematic in that not all individuals faced the exact same set of cases and the average would be altered by this. As the case differences were very subtle and did not all vary in favour of more donations being made to one set, the impact of this is likely to be minor.

Related to this area of analysis was an investigation into the relationship between reported feelings. This section of the analysis considered whether certain feelings appeared to lead to others. Ordinal regression analysis was again used, this time with the feelings question under consideration as the dependent variable and other feeling variables as the explanatory variables, along with the individual's average response to the feelings question, excluding the case for which the dependent observation was recorded. This analysis, as above, was done using transformed feelings data.

5.4.2 Analysis of case ten responses

The analysis of responses to case ten was somewhat simpler, as it did not require regression analysis. Respondents could decide to ensure that Oxfam was helped or that they themselves received a higher pay-off. The analysis aimed to identify whether the probability of ensuring Oxfam was helped differed significantly in the four different situations: when they were alone in their decision; when there was another participant involved but they could not find out what their response was; when with another where they could find out the partner's response but did not; finally, where they could find out the response of the partner and did.

The response in each of the situations was compared with the responses in each of the other three situations, using the chi-squared test. This allowed the hypotheses previously outlined to be tested.

5.5 Summary

This chapter has outlined how an experimental approach was used to examine some practical research questions and, at the same time, to test theoretical arguments. The approach was developed by combining methods from the charitable-giving literature, health priorities, and experimental economics.

As far as possible, this chapter has remained non-technical. Obviously there are technical issues that did require careful consideration. These are outlined in detail in the methods appendix, together with the formal statement of hypotheses. This appendix also includes more detail on the experiment/survey in terms of design and content. The following chapter presents the results of the analysis as described above.

Chapter 5 Appendix: Technical issues and details

A5.1 Technical issues

A5.1.1 Model outline: Ordered regression models ⁶

Ordered regression models (ORM) are typically used when it is efficient to consider the natural order of certain multinomial variables. Common examples include analyses of opinion surveys and bond ratings. Using a multinomial logit or probit would ignore the ordinal character of such variables, while OLS would consider all differences to be equal. So, in an opinion survey analysis with the responses 'strongly agree', 'agree', 'disagree', and 'strongly disagree' a multinomial model would examine only the groups, while OLS would consider a shift from 'agree' to 'strongly agree' to be the same as one from 'disagree' to 'agree'. ORM seek to address both of these drawbacks.

Typically ORM are explained through the consideration of a latent variable. The latent variable in the problem at hand being the desire to help (h^*) or associated feelings. The latent variable is defined as follows:

$$h^* = \beta'x + \varepsilon$$

The desire to help is unobserved; what is observed is:

$$d = 0 \text{ if } h^* \leq 0$$

$$d = 1 \text{ if } 0 < h^* \leq \mu_1$$

$$d = 2 \text{ if } \mu_1 < h^* \leq \mu_2$$

.

$$d = 10 \text{ if } \mu_9 \leq h^*$$

⁶ The introduction to ordinal regression models is adapted from Greene(1997: 926 – 929) unless otherwise stated.

In the case being examined, the above suggests that individual respondents have a desire to help. They were not asked how much they wanted to help but rather how much they were willing to donate (d) as a proxy. It is assumed that the more they wish to help the more they will donate. Assuming the desire to help is defined as being zero for an indifferent individual, negative for those who dislike a case and would rather take from it than give to it, and positive for those who wish to help, the model can be understood as follows. The desire to help is not observed, what is observed is a zero donation ($d = 0$): for those who are indifferent or dislike a case, a one pound donation ($d = 1$); for respondents who are more than indifferent but feel less desire to help than μ_1 ; and so on. This structure continues until the desire to help equals or passes μ_9 , from which point on a donation of 10 pounds is observed. The desire to help is assumed to be dependent on 'x' measured factors and ε unmeasured and possible unobservable factors.

The μ 's and the β 's are unknown and so have to be estimated. Maximum likelihood estimation is used, but to do this the probabilities have to be defined. How they are defined is determined by what assumptions are made with regard to ε . Typically ε is assumed to be either normally or logistically distributed, the former resulting in the ordered probit and the latter in the ordered logit, the choice making little difference to the results.

Assuming a normal distribution, the probabilities can be defined as follows, which can be shown to sum to unitary:

$$\text{Prob}(d = 0) = \Phi(-\beta'x)$$

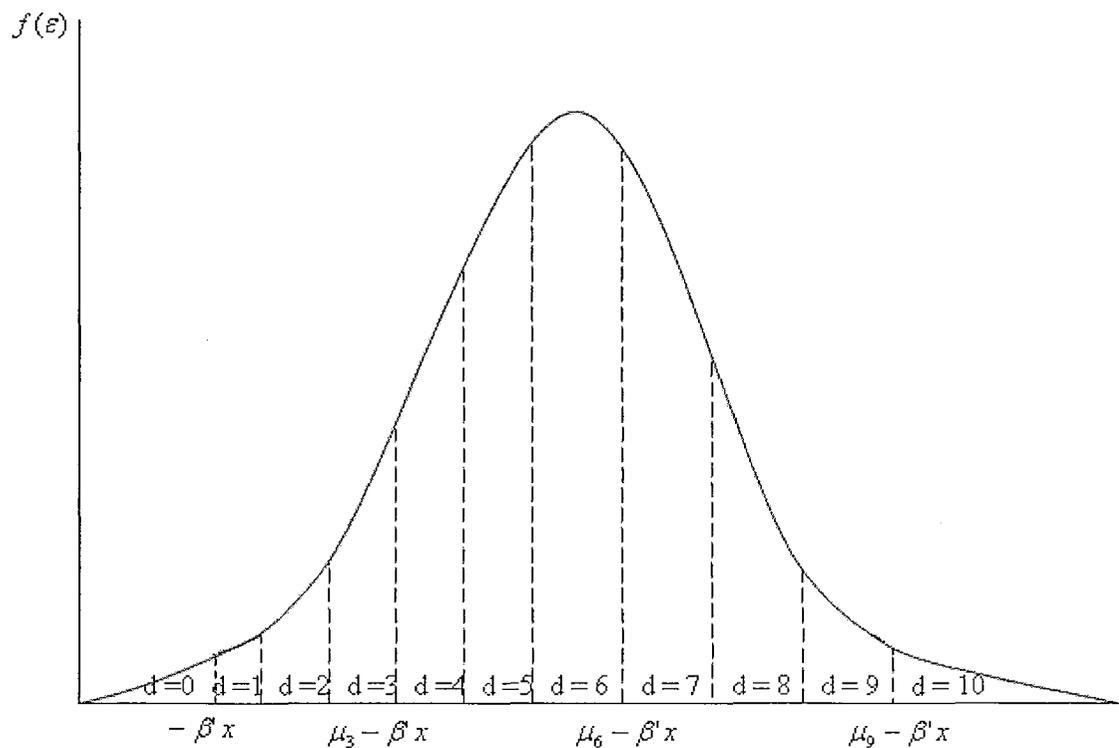
$$\text{Prob}(d = 1) = \Phi(\mu_1 - \beta'x) - \Phi(-\beta'x)$$

$$\text{Prob}(d = 2) = \Phi(\mu_2 - \beta'x) - \Phi(\mu_1 - \beta'x)$$

$$\text{Prob}(d = 10) = 1 - \Phi(\mu_9 - \beta'x)$$

The above can be represented graphically as follows:

Figure A5.1: The ordered probit model



What the graphic shows is how the donations of 0 – 10 each correspond to a not always identical space under the curve. The probabilities reflect the likelihood, given 'x', of an observation being in that space.

Once the model is estimated, it is important to note that the coefficients are not the marginal impacts of the explanatory variables. Increasing 'x', while holding β and μ constant, is equivalent to shifting the curve. If β is positive, the curve will shift to the right with an increase in 'x'. This shift will increase the probability of a donation of 10 and decrease the probability of a donation of zero, but the impact on the intermediary amounts is unclear.

The above model was used to examine the differences between single cases across sets, in terms of both donations and feelings, and the determination of donation amount and feelings reported across all cases and sets. In the latter cases, it was necessary to account for the relatedness of the observations. The assumption that the observations are independent could not be maintained, as repeated observations were taken from the same individuals. To address this, robust standard errors that considered the clustering of donations and feelings around individuals were calculated.

To examine the importance of model choice, regressions of differences across sets were conducted using ordered probit, as outlined above, censored regression models, accounting for upper and lower limit censoring (referred to in the results as interval regressions), and finally OLS. Only the results of the ordered probit are reported on in the results section, unless important differences, such as a change in a relevant coefficients sign or significance, were noted. The results of the other regressions are reported in the appendix to the results chapter.

The ordered probits conducted were tested for violation of the parallel regression assumption using an approximate likelihood ratio test as described by Long and Freese (2006). Ordered probits are essentially equivalent to running a series of binary probits with the assumption that the slopes are equal for each of the regressions, Long and Freese note that this assumption is

frequently violated. As three models for each regression were already been run it was decided not to add a fourth if the assumption was violated but rather to treat the results with some caution.

A5.1.2 Model specifications and specification of formal hypotheses

In keeping with the methods chapter, the models and associated hypotheses can be grouped into three areas: within case differences; across case differences; case ten analysis. The following deals with the first two of these as more detail is required on these than was provided in the body of the chapter.

A5.1.2.1 Differences within cases

For each case a series of ten ordered probits were conducted.

$$Y = \beta_1'x + \beta_2's + \beta_3'ave + \varepsilon$$

Where 'x' denotes a basic set of participant details, including a dummy variable denoting from which university the observation was collected. The 's' denotes which set was responded to and 'ave' the average, excluding the case under consideration, of whichever variable was being examined. 'Y' is a vector containing the amount donated and the responses to all feelings questions. If 'Y_i' is the donated amount then 'ave_i' is the average donations of respondents to other cases. The ten hypotheses tested (donations and nine feelings questions) for each case if $\beta_2' \neq [0]$. This involved a total of 90 hypotheses. Significance testing here, and throughout the analysis, was based on a $\alpha=0.05$.

A5.1.2.2 Differences across cases

Determination of donations

An ordered regression was estimated for the model, where the dependent variables consisted of the responses to the feelings questions. The regression was estimated with robust standard errors accounting for the clustering of responses around individuals as recommended by Long and Freese (2006).

$$d = \beta_0 sad + \beta_1 res + \beta_2 fault + \beta_3 get + \beta_4 sim + \beta_5 suffer + \beta_6 marginal + \beta_7 urgent + \beta_8 ave + \varepsilon$$

Sad = Difference from average response to sadness question

Res = Difference from average response to responsibility or pressure question

Fault = Difference from average response to question on fault of potential recipient

Get = Difference from average response to question of to what degree it is the recipient's responsibility to get out of their situation

Sim = Difference from average response to similarity question

Suffer = Difference from average response to perceived suffering question

Marginal = Difference from average response to marginal impact reported

Urgent = Difference from average perceived urgency response

Ave = Average donation by respondent to all other cases

The following hypotheses were tested for both regressions:

H0: $\beta_0 > 0$ (Reported sadness is positively related to higher donations)

H1: $\beta_1 > 0$ (Reported felt responsibility is positively related to higher donations)

H2: $\beta_2 < 0$ (Perceived fault is negatively related to higher donations)

H3: $\beta_3 < 0$ (Perceived responsibility to get themselves out is neg. related to donations)

H4: $\beta_4 = 0$ (Perceived similarity has no independent impact on donations)

H5: $\beta_5 = 0$ (Perceived suffering has no independent impact on donations)

H6: $\beta_6 > 0$ (Perceived marginal impact is positively related to higher donations)

H7: $\beta_7 > 0$ (Perceived urgency is positively related to higher donations)

Determination of feelings

The following ordered probits were estimated with robust standard errors to account for clustering.

$$sad = \lambda_0 ave + \lambda_1 suffer + \lambda_2 sim + \lambda_3 fault + \varepsilon$$

$$res = \alpha_0 ave + \alpha_1 sim + \alpha_2 suffer + \alpha_3 get + \alpha_4 marginal + \alpha_5 urgent + \varepsilon$$

Each regression was again conducted twice: once with the data as they were reported and once with the transformed data. The following hypotheses were tested for both versions.

H8: $\lambda_0 > 0$ (Average reported sadness to other cases is pos. related to reported sadness)

H9: $\lambda_1 > 0$ (Perceived suffering is positively related to reported sadness)

H10: $\lambda_2 > 0$ (Perceived similarity is positively related to reported sadness)

H11: $\lambda_3 < 0$ (Perceived fault is neg. related to reported sadness)

H12: $\alpha_0 > 0$ (Average felt responsibility to other cases is pos. related to felt responsibility)

H13: $\alpha_1 > 0$ (Perceived similarity is positively related to felt responsibility)

H14: $\alpha_2 > 0$ (Perceived suffering is positively related to felt responsibility)

H15: $\alpha_3 < 0$ (Perceived res. to get themselves out is neg. related to felt responsibility)

H16: $\alpha_4 > 0$ (Perceived marginal impact is positively related felt responsibility)

H17: $\alpha_5 > 0$ (Perceived urgency is positively related felt responsibility)

A5.2 Experiment design

A5.2.1 Summary of verbal introduction

Introduction

- Welcome and thanks for coming.
- This is a survey of Europeans' opinions regarding charitable causes.
- It should take about 30 minutes to complete, plus a few minutes to sort out your payments.
- You will all be paid 5 pounds for showing up here today.
- You will receive an extra 10 pounds to split between yourself and a charitable cause.
- You will be asked to do this 10 times for 10 different cases, but I am afraid that you will not get 100 pounds – what will happen is the programme will select one case at random and that one will become real. You will get the share you allocated to yourself; a charity working with people like the one described in the case will get the balance.
- We have done our best to match so you can think of the money going to someone like the person described.
- The money will be distributed once everyone is finished. If you finish early please could you wait quietly and allow others to finish without distraction.
- Twice in the programme you will be given an ID number. It is very important that you write it down – otherwise you cannot be paid.
- Some of you will be faced with a payoff matrix – please look at your sheets so that we can explain.
- Please read through the handout and begin – thank you.

A5.2.2 Handout: Explaining a pay-off matrix

A pay-off matrix outlines what you and other participants will receive, depending on the choices you make.

If you are faced with a choice between A or B and are paired with someone else in the room who has to choose between X and Y the pay-off may be summarised by the following.

		Your pair chooses	
		X	Y
You choose	A	8, 5	9, 6
Or	B	9, 4	2, 7

The combination of choices determines which pair of numbers will be selected. For example, if you choose A and your pair chooses X, the result is 8,5; on the other hand, if you choose B and your pair chooses X, the result is 9, 4. The resultant pair of numbers determines the pay-off. The first number is what you will get and the second number is what your pair gets.

The pay-off matrix can be extended to determine payments to non-participants, as in the following example:

		Your pair chooses	
		X	Y
You choose	A	8, 5, 1	9, 6, 5
Or	B	9, 4, 5	2, 7, 5

In this case the choices made by you and your pair still link to a set of numbers, but this time to three numbers. The first number is what you will get, the second will be what your pair gets and the last will be what the two of you determine a third party gets. So, if you choose A and your pair chooses Y, you will get 9, your pair will get 6 and the third party will get 5.

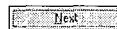
The idea is that you can try and predict what your pair will choose and then make your choice accordingly. For example, in the first matrix on this page you can see that it is always better for your pair to choose Y instead of X. So you could guess that this is what they will choose. Having assumed this, you can see that it is best for you to choose A in order to get the highest expected payoff.

A5.2.3 Introductory text and the interface

The following screens show the programme interface. They include all the introductory text the questions and examples of cases.



Welcome and thank you
for
agreeing to be part of this study



The responses you give here will be anonymous.

At the end of the study you will be paid an amount of up to 10 pounds, depending on your decisions.

[Back](#)[Next](#)

This study seeks to examine charitable behaviour. You will be presented with a number of cases. Each case is based on a real life situation and outlines the predicament of an individual. For each case you will be asked to divide 10 pounds between yourself and the individual described in the case

[Back](#)[Next](#)

After you have completed your allocations for each of the ten cases, the computer will select one case that will be played out. You will be paid the money you allocated to yourself for that particular case; the amount you allocated to the individual will be donated to a charity working with such people.

Back

Next

Every effort has been made to identify a charity working on cases as similar as possible to those outlined. Please try to think of your allocation going to support the individual presented in the case.

Before and after asking you to make allocations, we would like to ask you a number of questions. The whole exercise should take approximately 30minutes.

Back

Next

At the end of this session you will be allocated an identification number. You will be able to use this number to collect your share of the 10 pounds from the randomly selected case. Payments will be made in envelopes and the person distributing them will not know what you are receiving.

If you are not clear on anything please raise your hand and a researcher will gladly assist you.

Back

Next

Introductory questions

The following questions will be used as background information.

What will your age be at your next birthday?

What sex are you?

Do you have any siblings?

Do you have any children?

Back

Next

Case 1

A 55-year-old Indian man working as a cleaner has contracted pneumonia. He is very poor and has asked a local NGO for support while he recovers. The NGO is raising funds so that they can provide the support.

If this scenario is randomly selected, how much would you like to be donated to her cause, bearing in mind the balance would be paid to you for your participation today?

Amount paid to cause (0-10):

Case 10

This case is a little different and a little more complicated. Please choose one of the two following options after reading the explanation provided.

You have been paired with another member of the group. The outcome of this case will depend on what both of you decide to do. The following is a table of payoffs: if you choose A and your pair chooses A you get 8 pounds, your pair gets 8 pounds and 1 pound is donated to Oxfam. On the other hand if you choose A and your pair chooses B you get 8, they get 6 and 5 is donated to Oxfam, and so on for the other two combinations

		Your pair chooses	
		A	B
You choose Or	A	8, 8, 1	8, 6, 5
	B	6, 8, 5	6, 6, 5

Your pair has already responded to this case.
Would you like to know which option they selected

Select Option

Thank you for completing the allocations section of the study. One of the cases you have just considered will be selected at random and you can collect your portion of the money using the following ID number. Please note it down, you will not be able to collect your payment without it.

QHKQEL420293

The balance will be donated to the most appropriate charity. Before you go to collect your payment please could you answer the following questions in regard to each case. No need to spend a great deal of time on them, simple answer as you see fit, we would really like to know your first response.

Next

Case Reminder : 1

A 55-year-old Indian man working as a cleaner has contracted pneumonia. He is very poor and has asked a local NGO for support while he recovers. The NGO is raising funds so that they can provide the support.

Please answer the following in relation to the case, on a scale of 0-9: (0 - low level, 9 - high level)

- 1. How saddened were you by this story?
- 2. Does this story make you feel a pressure to help?
- 3. How much do you think people like this are to blame for their situation?
- 4. How much is it up to them to get themselves out of their current situation?
- 5. What do you think their level of suffering is?
- 6. Do you feel you have a responsibility to help people like this?
- 7. How similar would you say this person is to you?
- 8. How much help do you think a donation of 5 pounds would be?
- 9. How urgent would you say this case was?

Next

Just a final few more questions about yourself:

1.How would you rate your current financial position?
(1 = very bad, 5 = very good)

2.Do you have close family ties with:

Asia:

Africa:

Eastern Europe:

North America

3.Which of the following would you say best describes you:

a) Not religious b) Christian c) Hindu d) Buddhist
e) Muslim f) Jewish g) Religious but not one mentioned h) Decline

Chapter 6: Experiment results

6.1 Introduction

The experiment was conducted and the data collected. This chapter presents the analysis of those data as per the design in Chapter 5. Firstly, a description is provided of the data collected, raising any issues that were of concern. Secondly, the results of the analysis of within-case variations are presented. This section examines whether the subtle changes to cases were associated with significantly different donations and responses to follow-up questions.

The third section considers what association was observed between the donations made and the follow-up questions, across all cases. Thereafter, the question of the relationship between the follow-up questions is examined; finally, there is a section presenting the analysis of case ten, which was of a different type from the other cases.

This chapter seeks to present the results in as clear a manner as possible so as to highlight the major findings. More detailed results are provided in the appendix. The implications of these results will then be considered.

6.2 Description of data

In the final analysis, it was decided to use only the data from the main collection and to exclude the data collected during the pilot study. It was noted in the methods chapter that it was possible to use the donations data from the pilot study, as the process and presentation were identical to the main survey, whereas the follow-up questions were not identical and could not be used. Although this was the case, it is simpler to analyse and present one set of data rather than include some parts sometimes. The data described in this section and used in all the analysis therefore exclude the pilot results.

The experiment was implemented at two sites. At the first site, a total of 35 respondents took part; this was the site where a class was given the option of participating or taking the time off. Due to a programming error, six responses were lost, leaving 29 usable records. At the second site, a total of 52 responses were collected, meaning that a total of 87 volunteers participated, generating 81 usable records. The characteristics of the respondents are outlined in the following table.

Table 6.1: Respondent characteristics

Variable	Range (mean : SD ⁷) or Percentage of respondents
Age	19 – 37 (24.18 : 4.55)
Male	60%
No children	99%
Have siblings	80%
Financial status 1 – Very bad	7%
Financial status 2	32%
Financial status 3	48%
Financial status 4	9%
Financial status 5 – Very good	4%
Close ties to Asia	15%
Close ties to Africa	4%
Close ties to E. Europe	6%
Not religious	48%
Christian	40%
Hindu	0%
Buddhist	1%

⁷ Standard deviation

Muslim	2.5%
Jewish	1%
Religious but religion not specified	6.5%
Declined to select religion	1%

As the sample was comprised of students, the average age of 24 is in the range expected. As was also predicted, most of the sample – with one exception - did not report having children; eighty percent, however, reported having siblings. The distribution of participants in terms of their ranking of their financial status centred on the mid-point of three but was skewed towards 'bad'. Again, this generally mid- to bad range of responses is what would be expected of a student population.

Fifteen percent of respondents reported having close family ties with Asia. This was far more than reported family ties with Africa or Eastern Europe – a difference that is important in later analysis.

The largest single response to the religious question was 'not religious', although a large proportion responded that they were Christian. Because of the small numbers reporting that they were religious but not Christian, in the analysis the sample was often simply divided into those who responded that they were religious and those who did not.

The following table presents a summary of the responses of the participants. It details the range of responses given, the mean response, and its standard deviation for the amount donated, and all the follow-up questions. In each case, the average is of 729 responses, that is 81 individuals each providing 9 responses.

Table 6.2: Responses

Variable	Range (mean : SD)
Donation amount	0 – 10 (3.96 : 3.00)
Sadness felt	0 – 9 (4.67 : 2.53)
Pressure felt	0 – 9 (3.90 : 2.40)
Blame attributed	0 – 9 (1.70 : 2.17)
Perceived responsibility of potential beneficiary to get themselves out	0 – 9 (2.36 : 2.26)
Suffering perceived	0 – 9 (5.98 : 2.09)
Responsibility felt	0 – 9 (3.84 : 2.42)
Similarity perceived	0 – 9 (2.06 : 2.43)
Marginal impact predicted	0 – 9 (5.04 : 2.15)
Urgency perceived	0 – 9 (5.23 : 2.57)

The table shows that, on average, close to 4 of the 10 pounds allocated to respondents was donated to the case they were asked to consider. This is a fairly high percentage of the money they were given and suggests that they took the exercise seriously. Only 5 of the respondents gave nothing to every case they were asked to consider, thus guaranteeing themselves the full amount. It is also worth noting that 3 participants gave 10 pounds to every case, ensuring that they received nothing other than the attendance fee, which two people returned, asking that it also be donated to charity; there is, however, no way of knowing if these were among the three.

For all of the variables, responses were given across the entire range with responses at both extremes recorded. It is noteworthy that respondents generally reported low levels of feelings of similarity with the cases being considered, and generally attached low levels of blame. Suffering, marginal impact, and urgency stand out as generally promoting higher responses. There is no way of judging from this table why responses varied. It simply notes that

they did; the why is the question that dominates the bulk of the rest of this chapter.

6.3 Examining within-case variations

Two versions of each of the 9 cases considered were prepared and participants were randomly assigned one or the other set of these cases. This ensured that each participant was faced with each of the nine cases, but never with both versions. The two versions of the cases differed in only one way. The interest of this section is whether that one change was sufficient to prompt participants to differ in their donations and responses to follow-up questions. Each of the nine cases is discussed below, examining what significant differences were observed between the responses to the two versions. The left hand column in all of the following tables presents the version of the case provided in set one.

Case 1: Religion	
An appeal has been made by a Mosque in Pakistan for a member of its community. Nazir is a 42-year-old man who is suffering from cancer. He requires treatment that he cannot afford. The treatment will increase his survival chances substantially.	An appeal has been made by a Church in Pakistan for a member of its community. Nazir is a 42-year-old man who is suffering from cancer. He requires treatment that he cannot afford. The treatment will increase his survival chances substantially.

The change in the above case was not sufficient, according to the criteria for significance employed in the analysis, to prompt a significant difference in the amount donated. That said, it should be noted that, if the significance level

were taken at 10% rather than 5%, it would be recorded as a significant difference in favour of a higher donation to set two, which mentions the Church.

While the donation was not significantly different across versions, respondents to set one did report a significantly lower level of sadness and feelings of responsibility to help than those faced with set two. Respondents who reported that they were Christians and were faced with set one gave significantly less than those faced with set two, whereas the opposite relationship was recorded for respondents with close family ties with Asia, who gave more in set one. These differences were all observed, while there was no significant variation in feelings of suffering, impact, or urgency, which might be considered possible reasons for differences in donations in general.

Case 2: Context	
<p>Simpiwe, aged 24, lives in Swaziland and has malaria. She lives in an area which is currently experiencing a food shortage and has high rates of HIV, malaria and cholera. A medical NGO is appealing for support to provide treatment that would improve her chances of survival greatly.</p>	<p>Simpiwe, aged 24, lives in Swaziland and has malaria. She lives in an area which has high rates of malaria, but is otherwise doing well. A medical NGO is appealing for support to provide treatment that would improve her chances of survival greatly.</p>

Case two was initially included to examine whether participants write off helping when everything is going wrong. So, while the individual situation of the potential recipient is unchanged across the versions, the context alters. This change did prompt a number of significant differences in responses, but not in the manner expected.

Participants presented with the set one version of the case (on the left) gave more, reported greater sadness, felt more pressure to help, felt there was more suffering and that the situation was more urgent. This was also the only case where one version - the version in set one - prompted participants with close family ties with Africa to donate more. No significant difference in marginal impact was observed.

Case 3: Possible fault	
A 50-year-old Kenyan man was injured in a car accident while driving home at night from a bar . A local charity is appealing for funds to assist in his physical rehabilitation, which would greatly improve his quality of life.	A 50-year-old Kenyan man was injured in a car accident while driving home at night from work . A local charity is appealing for funds to assist in his physical rehabilitation, which would greatly improve his quality of life.

The differences in responses in the above case were in line with expectations. While the case does not say that the man was drunk, it was probably assumed that he had been drinking. The set one version received significantly lower donations and participants reported less sadness and pressure to help; felt that there was more fault and a greater responsibility for the potential recipient to get themselves out of the situation.

Case 4: Age – adult to child	
Joseph, a 35-year-old Zambian, is suffering from an unpleasant skin condition that, while not fatal, is extremely uncomfortable. If left untreated it will likely last for up to five years, whereas with treatment he should recover in a few weeks. A charity is appealing for help to pay for the treatment.	Joseph, a 3-year-old Zambian, is suffering from an unpleasant skin condition that, while not fatal, is extremely uncomfortable. If left untreated it will likely last for up to five years, whereas with treatment he should recover in a few weeks. A charity is appealing for help to pay for the treatment.

Participants who considered the set one version of the case reported lower levels of sadness and feelings of responsibility to help and a higher rating of the extent to which it is the recipient's responsibility to get themselves out. These differences, however, were not associated with a significantly different level of donation.

Case 5: Sex	
A 55-year-old Indian woman working as a cleaner has contracted pneumonia. She is very poor and has asked a local NGO for support while she recovers. The NGO is raising funds so that they can provide the support.	A 55-year-old Indian man working as a cleaner has contracted pneumonia. He is very poor and has asked a local NGO for support while he recovers. The NGO is raising funds so that they can provide the support.

This case went against expectation in that a significantly lower donation was recorded for those faced with the set one version, which mentions a woman. While the difference in donations was significant, nothing else was.

Case 6: Nationality / ethnic group	
A medical NGO is seeking support to provide treatment for an HIV+ Zimbabwean man, aged 34, who has recently progressed to AIDS. While treatment will not cure him, it will likely extend his life for a number of years.	A medical NGO is seeking support to provide treatment for a HIV+ Ukrainian man, aged 34, who has recently progressed to AIDS. While treatment will not cure him, it will likely extend his life for a number of years.

Very little difference was observed when comparing responses to the two versions. The only one was that, in set one, respondents reported a lower level of similarity with the recipient than those in set two. This is what was expected, but this was not carried through to a difference in sadness, pressure to help, or amount donated. What was recorded was that individuals reporting close family ties with Eastern Europe gave more when presented with the set two version.

Case 7: Implied wealth	
Mary, a 47-year-old community worker , was hit by a car outside her office in Lusaka, Zambia. She has suffered internal injuries and needs surgery, which is not available in Zambia, and her church is raising funds for her to go to South Africa.	Mary, a 47-year-old corporate lawyer , was hit by a car outside her office in Lusaka, Zambia. She has suffered internal injuries and needs surgery, which is not available in Zambia, and her church is raising funds for her to go to South Africa.

Of all the case variations, case 7 was associated with the most significant differences. To start with, set one received higher relative donations. The idea behind the case was that the lawyer would be seen as no different in terms of similarity and that the cases would be as urgent, but more importantly that it would be her responsibility to get out of the situation. All of these differences, or lack thereof, were observed.

In addition to these, the set one version was also associated with more sadness, pressure to help, perceived suffering, more responsibility to help, and less blame. It is as if the obvious need for help, and the lack of excuse not to provide support, led to everything falling in line, whereas when there is a way out other factors are also downplayed.

Case 8: Age – young to old	
Afina, aged 22 , from Romania, is suffering from tuberculosis. This illness can be cured, but funds are currently not available to provide treatment. Her local Church is appealing for help.	Afina, aged 72 , from Romania, is suffering from tuberculosis. This illness can be cured, but funds are currently not available to provide treatment. Her local Church is appealing for help.

Age is often raised as a factor in allocations, but in this case it did not lead to significantly different donations. Participants considering set one reported higher relative feelings of similarity, possibly a result of the similar age, but with seemingly no major impact on donations.

Case 9: Emotive issue	
The following case occurs all the time; the money donated will go towards organisations providing such services. A Malawian woman has been raped and needs funds to be provided with a short course of drugs to reduce her chances of being infected with HIV.	The following case occurs all the time; the money donated will go towards organisations providing such services. A Malawian woman has been injured in a medical accident and needs funds to be provided with a short course of drugs to reduce her chances of being infected with HIV.

The mention of rape was associated with feelings of greater pressure and greater responsibility to help, a higher relative level of suffering, and a lower attribution of responsibility to get out of the situation. All of these findings could

be expected. What was also expected was a higher donation to the rape victim; both versions of this case did lead to high donations. The version in set one received the highest average donation of 5.8 pounds and the set two version the second highest donation of 4.7 pounds. The set one version also led to more 10 pound donations than any other case. Despite this, the difference in donations was not significant, as it appears that both versions prompted a strong response.

6.3.1 Emerging across-case issues

Considering the results across the different cases, a number of interesting findings can be noted. Firstly, the importance of factors other than the expected marginal impact and the urgency of the situation appear to be related to differences in donations. Any possible attachment of fault appears to be latched onto by respondents, as does the suggestion that the potential recipient may have means of their own. Moreover, they then appear to bring other feelings and assessments into line with the decision. Helping when help is not needed is irrational under any of the models discussed, so it would seem appropriate to take the fact that an individual had means as a pointer. What is more interesting is the alignment of other responses: for example, why should the lawyer in case 7 be thought to be more to blame for her situation?

In neither of the cases that differed in terms of age was there a difference in donations; nor did the change in sex lead to the expected difference. The age differences were expected but possibly were not a strong enough factor to prompt the change. The difference in favour of the man is problematic, as it was also not associated with any other differences that may have helped to explain the observation.

Case 6, dealing with changed nationalities, also did not lead to significant differences, as it was expected to. In fact, while not significant, the case mentioning the Zambian actually got on average more than the case mentioning

the Ukrainian, although a difference in favour of the Ukrainian was significant for those who reported family ties with Eastern Europe. The case was intended to see whether a geographically-closer recipient was considered more similar and whether such a consideration might lead to higher donations, but this design may have underestimated current EU sensitivities towards Eastern Europe. While this case did not bring out such biases, the results in case one to some extent did, with the apparent religious biases of Christian respondents and those with close family ties to Asia.

When there are significant differences in the donation amount, the case variation is typically associated with other significant differences. On average, versions with significantly different donation levels are associated with 4 other significant differences, as opposed to 2.4 for cases with non-significant differences in donations. The relationship between donations and follow-up questions can be examined in much more detail by looking across cases, which is the topic of the next section.

6.4 Determination of donations and feelings across cases

The design of the experiment allowed for the examination of the association between follow-up questions and the amount donated, as well as between the follow-up questions themselves. To this end, a series of regressions, as described in the previous chapter, was conducted. The full regression results are provided in the appendix; this section concentrates on the major findings.

The first area of examination considered the relationship between the follow-up questions and the amount donated; a number of hypotheses were discussed in relation to this. Key among these was the hypothesis that the variable associated with pressure would have a positive impact on the amount donated, independently of the positive impact associated with sadness. A primary argument of this thesis is that there exists a motivation to help associated with

responsibility felt as a pressure, which is distinct from sadness. This distinct association was observed. The following table presents the results of the regression, listing the direction of the relationship with the amount donated and the p-value of the coefficient. A p-value of less than 0.05 is considered significant.

Table 6.3: Donation regression

Variable	Relationship and p-value
Non-current average	+ (0.000)
Sadness	+ (0.000)
Pressure	+ (0.000)
Blame	- (0.788)
Res. to get out	- (0.000)
Suffering	- (0.254)
Similarity	- (0.407)
Marginal impact	+ (0.278)
Urgency	+ (0.057)

Other than the testing of the distinct relationships of sadness and pressure, a number of other hypotheses were tested with these results. Before discussing the hypotheses, it is worth noting the significance of the non-current average. This is the average amount donated by an individual to other cases. It suggests that, the more an individual donated in general, the more they were likely to donate to any specific case. This in some way measures generosity. If the examination were concentrated on who gives, then this would be the key variable to explain. The interest here, however, is on the relative importance of different cases and how feelings and assessments are related to these differences, so it is necessary to control for non-case-specific factors such as generosity.

As was discussed in the methods chapter, the follow-up questions were transformed before being included in the regression. As different individuals may have a tendency to use higher numbers to respond, making comparisons across individuals difficult, the difference of each variable from the average response of the individual to that variable for other cases was calculated. The variable then measures the relative responses, which are more readily comparable. The above regression is, therefore, measuring the association between relative responses and, having controlled for generosity, relative donations.

The degree of blame, and of responsibility for the recipient to get themselves out of their situation, were both expected to be negatively related to donations. Both were, but only the responsibility to get out coefficient was significant. As mentioned above, it appears that respondents brought feelings into line with decisions, so blame and recipient responsibility may well be related. Indeed they are, with a correlation of 0.5. Despite this, removing recipient responsibility from the regression did not lead to the blame relationship becoming significant.

None of the other variables were associated with a significant relationship, only urgency coming close and with the correct sign.

The above results would appear to suggest that the general level of giving (generosity), feelings of sadness, pressure, and perceived responsibility of the recipient are the key factors linked to donations. That is not to say that the others do not have a place in determining donations, but rather that their influence may be felt through their impact on sadness and felt pressure. This was the argument behind similarity and suffering being hypothesised to have no relationship to donations independently of sadness; indeed, this relationship is what was observed. It was thought that urgency, marginal impact, and the recipient's responsibility might have independent impacts associated with a

rationalisation of the response to the case in an experimental setting, but this was only observed for the last of these variables, participant responsibility.

The following two regressions examine the hypotheses related to the relationships between the variables. The first deals with the sadness variable and what relationship there was with blame, suffering, and similarity. As with the donation regression, it was necessary, in order to control for the general level of the rating, to remove individual characteristics not associated with the case by including the non-current average of the sadness variable in the regression.

Table 6.4: Sadness regression

Variable	Relationship and p-value
Non-current average	+ (0.000)
Blame	- (0.000)
Suffering	+ (0.000)
Similarity	+ (0.000)

The results here were in line with expectations. A negative relationship was observed between blame and reported sadness and a positive relationship with both suffering and similarity. None of these variables had an impact independent of sadness in the donation regression, but they do appear to be associated with the level of sadness. Similarly, the relationships between pressure and other responses were examined.

Table 6.5: Pressure regression

Variable	Relationship and p-value
Non-current average	+ (0.000)
Res. to get out	- (0.017)
Suffering	+ (0.000)
Similarity	+ (0.003)
Marginal impact	+ (0.000)
Urgency	+ (0.000)

Again, all the relationships were as expected and significant. Individuals reported that they felt relatively less pressure to help when they attached higher levels of responsibility to the potential recipient. They indicated more pressure when they felt the individual was suffering more, was more similar to them, when their help would have more impact, or when the situation was perceived as relatively more urgent.

6.5 Case ten

The final case each participant was asked to consider was one of three versions of case 10: Set one having two versions and set two having one. The details of the versions are provided in the previous chapter but to recap:

Set one version I involved a choice between option A and option B; option A had the higher pay-off for the participant and a low pay-off for the charity Oxfam. Option B was for a reduced pay-off for the participant and an increased pay-off for Oxfam. In version II of set one, another participant was linked and they faced the following pay-off matrix, where the first figure in each set would go to them, the second to their partner and the last to Oxfam:

		Partner Chooses		
		A	or	B
You choose	A	8, 8, 1		8, 6, 5
Or	B	6, 8, 5		6, 6, 5

Either individual could opt to select B, ensuring Oxfam received more, or option A ensuring that they themselves received more. This is a test of the hypothesis that, as more potential helpers are involved, the individual will help less and the probability of help being provided at all will be reduced.

The set two version was identical to the set one version II, except that the participant was given the option of finding out what their 'partner' had selected. The partner was hypothetical and, if the information was taken, it always reported that they had not helped. The following responses were observed. The numbers on the left of each column reflect the number of respondents and those in parenthesis are the percentage of respondents faced with the version being examined who responded in this way.

Table 6.6: Case 10 responses

Option	Set 1		Set 2	
	Version I – alone	Version II – paired	Did not take info. 11 (26)	Did take info. 32 (74)
A (did not help)	6 (27)	12 (75)	7 (64)	23 (72)
B (helped)	16 (73)	4 (25)	4 (36)	9 (28)
Total	22 (100)	16 (100)	11 (100)	32 (100)

The table shows that 22 people were faced with version I. This was the alone version and a majority of 73% opted to give up some of their pay-off so that

Oxfam received more. A total of 16 people dealt with version II, with the bulk opting not to help in this paired situation.

Forty-three people were presented with the set two version. The majority decided to take the information and, having found out that their partner did not help, opted themselves not to help. Some (9) did help after finding out that their partner had not helped; some even helped and ignored the information.

If option A responses are defined as 'did not help' and option B as 'helped' then the case can be discussed in more traditional terms. The first question to ask is whether the differences in the probability of helping were significant across the versions and choices on information. This requires that the responses in each column in the above table be compared to the responses in every other column. The comparison involved a chi-squared test and the results are presented in the following table. A p-value of less than 0.05 suggests a significant difference.

Table 6.7: Significance tests of case 10 responses

Chi-squared tests	P- values
Set 1 version I / Set 1 version II:	0.004
Set 1 version I / Set 2 did not take:	0.044
Set 1 version I / Set 2 did take:	0.001
Set 1 version II / Set 2 did not take:	0.525
Set 1 version II / Set 2 did take:	0.818
Set 2 did not take / Set 2 did take:	0.608

Version I prompted significantly different responses to all other versions, regardless even of the set two information choice. None of the other responses were significantly different from each other.

A number of key observations can be taken from these results. Firstly, the probability of a participant helping did significantly fall as soon as others became involved. This is what has been traditionally found. Also in line with traditional results is that not only does the individual's probability of helping fall when more people are involved, but so too does the total probability that help will be provided. In version I, the probability that help would be provided was 0.73. In version II, the probability that any help at all would be provided was equal to 0.44 $[1-(0.75)^2]$.

The above replication of results in an anonymous situation is interesting, but what is more interesting are the responses to the set two version. The situation was constructed to see whether respondents would refuse information and avoid the possibility of being put in a situation where the case reduces to version I. The majority of participants, however, opted to take the information. What is interesting is that, even when a participant found out that their partner had not helped and that the choice was entirely theirs – essentially the same situation as in version 1, they still helped significantly less than did those who were faced with version I. In fact, they helped with pretty much the same probability as if they had not had the option of knowing their partner's response.

Some participants opted to ignore the information and not help, showing that they did not intend to help in any case. What is difficult to explain is why anyone would refuse the information but still help.

6.6 Discussion

The experiment was designed and implemented with a dual purpose in mind: the collection of practical data and the examination of theoretical arguments. The practical data were collected in order to feed into the discussion of when

and why people help, while the theory is more concerned specifically with the question of why people help.

From a practical viewpoint, the above results suggest that people help others when they feel saddened by their situation and feel a pressure to help. What that pressure is, is the subject of the theoretical debate. These associations with sadness and pressure to help are far from insightful and are rather obvious. A more substantial observation was the apparent importance of these feelings in mediating the influence of other variables. The data support an argument that similarity, suffering, blame, marginal impact and, to some extent, urgency influence help only by way of these feelings and not independently of them.

If the relationships observed in this sample are causative, it suggests that the individuals included in the study tend to help those who suffer more, are less to blame, can be helped more, and whose situations are more urgent. All of these would appear to be reasonably objective grounds on which to base allocations, but there are other factors. Firstly, they also appeared to be more saddened and pressured, and as a result helped more, when they felt the potential recipient was more similar to them. This type of association was also apparent with Christian respondents helping more when a Church was mentioned. Secondly, the level of suffering and blame appear not to be assessed on a common scale. There is at least some suggestion that respondents brought their feelings into line, with the attitudes towards the lawyer and the man returning from the bar being the main examples. These examples also showed the important role of assuming, and attaching, blame in determining levels of support.

Consideration of the question of why people help was always going to be more difficult than examining when they help and what feelings are associated with that help. Previous discussions have noted the variety of models that have

been employed to explain why people help. These models work to varying degrees to explain the above results.

The argument that people only help themselves, and are only interested in material pay-offs, is useless in explaining the behaviour observed in this study, in any dictator game where participants share, and in many other situations. Also, any arguments that are based on reciprocity are similarly useless in this context, as are those that argue that giving is associated with public acclaim and image. The inability of these models to explain any sharing of the payment in this experiment was clear even before it was conducted. What is of interest is whether the results shed any light on the relative merits of the other models. It must again be noted that the search is not for a single model to explain helping in all circumstances, but rather to see which models have merit in any circumstances and, in particular, if there are any situations in which the responsibility formulations hold sway.

The incorporation models are based on the assertion that people can incorporate others' welfare into their own welfare functions. As a result of this incorporation, when they then maximise their welfare function, they share resources between themselves and the other as appropriate to achieve the desired outcome. The strong role played by sadness in the above results would be supportive of this type of model. The knowledge of another's plight reduces the individual's welfare and, as a result, prompts help. That said, there are still many problems associated with using this model to explain helping behaviour, particularly when trying to explain the helping of strangers. These have been discussed at length in previous chapters.

A number of the problems with the incorporation models are highlighted in the results, aside from any discussion of how the individuals in these cases came to be in the welfare functions of these respondents, and the possibility for temporary presence. Firstly, if an individual incorporates another in their

welfare function, even only temporarily, the model would require that the coefficient on that recipient's welfare be variable – dependent on a host of factors – if it is to hope to explain observed behaviour. For example, how would the model explain less help when the recipient is more to blame? The more blame the individual perceives to be associated with the potential recipient the lower the coefficient on that recipient's welfare outcome in the individual's welfare function. If the coefficient varies according to context, it must be asked if it is the welfare of the particular recipient that the individual is interested in or the welfare of recipients in that situation and context. If it is the latter, this would seem to be more associated with values than welfare incorporation and be better explained by the other models.

Secondly, if the welfare of the recipient is incorporated, what motivation would the individual have to assume fault or the ability of the recipient to help themselves? Such assumptions appear to be made in an effort to avoid helping. If the potential helper would receive benefits from helping, there is no reason for them trying to interpret the situation so that helping would provide them with fewer benefits.

Imagine that an individual was faced with two scenarios, identical except that in the first the potential beneficiary is clearly at fault. In each scenario the individual divides resources between themselves and the recipient. So, in scenario one, they allocate themselves an amount and the potential recipient another amount so as to maximise their own welfare function, which includes the welfare outcome of the recipient. If blame is associated with a lower coefficient on the recipient's welfare in the helper's welfare function, then when faced with scenario two, where there is less blame, the helper would allocate more to the recipient to maintain their maximisation. As a result, the potential helper would generate greater welfare for themselves in scenario two. If in doubt, it would pay for the potential helper to assume less fault, or more suffering, or

less ability to get out of the situation, or greater marginal impact. This is not what was observed in the cases presented above.

The final problem with the incorporation models highlighted by these results is associated with case 10. If an individual's interest in helping is because they incorporate the welfare of those they help, you would not expect the outcome observed for this case.

The reduced personal, and total, probability of helping is in line with the predictions of this model of behaviour. Predictions based on the incorporation model would also suggest that anyone considering help would take the information available to them in the set two version; indeed, the majority of participants did take the information. The incorporation models would predict helping as much in set two, once the information was taken, as in set one version I. This is because, once the information is taken and the participant is aware that the allocation to Oxfam is up to them, if they are considering only their pay-off and the pay-off to Oxfam, the situation would be identical to set one version I, where they were alone in the decision. The counter to this is that helping alone and helping when another has refused are different things. Indeed they are, but the only way to integrate this into the incorporation models is again to argue for a context-specific coefficient, that is to value the welfare of others differently depending on the context.

It would appear that the incorporation models can only cope if the variability of coefficients is allowed. This shifts the discussion to what shapes the determination of the coefficient of inclusion. Such a discussion changes the focus from incorporation models to private benefit of action models, such as warm glow versions, and private benefit of maintenance models, such as the maintenance of a positive self-image model.

Private benefit of action models essentially argue that helping behaviour is shaped by the seeking of psychological benefits and the avoidance of psychological costs. Similarly, maintenance of self-image models are based on the individual's desire to maintain a positive self-image; acting in a positive way and avoiding negative behaviours can be seen as efforts to do this.

Both these types of model provide a much more powerful framework within which to understand the above results. As has been discussed in previous chapters, they open the way for context effects and self-manipulation. It is not the recipient that they value; it is their actions towards recipients within a context that they value. Moreover, it is the way they interpret these actions within their interpretation of the context.

It is much easier to provide an explanation for the case 10 responses with these models. The benefits to yourself, in terms of how the action might make you feel, are arguably different when you help or refuse help as the lone helper, as opposed to helping or refusing to help as the second potential helper to do so. Similarly, the benefits of helping someone who is to blame could easily be argued to be different from helping someone who is not.

Both these models also deal well with individuals bringing feelings and assessments into line with their decision. If the individual decides that they will feel best by not helping, they can then interpret the decision so as to maximise the psychological benefit of their decision even further. Or, if they value maintaining a positive self-image, they could interpret the context so that their not helping is seen in the best light, and so does minimal damage to self-image.

The strength of these models is also their weakness. By explaining behaviours in such broad terms, an explanation of almost anything can be thought up. There is a need to narrow down the models if they are to be useful. Efforts within the framework of self-image models have been made in this direction,

such as valuing being fair (Dana et al, 2007). This narrowed version has been used previously in experiments to explain seemingly altruistic behaviour. At a stretch, it could be used here to explain the case 10 results, but it would battle with the case variations.

An alternative self-image model would be that discussed in the preceding chapters, a model where individuals value feeling that they are a responsible person. This model would do well in explaining the above results. An individual feels less responsibility to help when others have not, as the resultant lack of help is not only their doing. It is easy to provide arguments, and the results do, as to why individuals may feel more or less responsible to different individuals in different circumstances and contexts.

The theoretical discussions of this topic outlined how this model of responsibility-based action could be constructed in three different ways. Firstly, the individual could value responsible action, but this would not fit well with self-manipulation to reduce the need for help. The apparent evidence that individuals take prompts to suggest fault would not be easily explained as such models would suggest a motivation to do the opposite for the same reasons described with regards to the incorporation models. Secondly, an individual could value feeling that they act in a responsible way. In the case of self-image maintenance, it is not the action but the feeling that is important, thus opening the way for a role to be played by self-manipulation and motivated assessments of contexts. The third approach is to argue that individuals feel a pressure to act responsibly and that this pressure is open to self-manipulation. From the perspective of making predictions, the difference between these latter two approaches makes little difference. There are, however, differences in terms of implications, which will be discussed in Chapter 8. More importantly, there are differences in the way in which the problem will be approached, depending on which framework is used.

The experiment reported on above was based on the balance model framework. The use of this framework led to the focus on the mediating role of emotions and the links with self-manipulation. The application of this framework, it is argued, has allowed for the collection of data in a manner which is useful in examining the four research questions outlined in the introduction of the previous chapter, to recap:

1. What characteristics of distant others and their need influence responses to them?
2. Are there emotional responses which place a pressure to help or not help?
3. Is there a pressure to act responsibly distinct from the pressure associated with relieving one's own feeling of sadness?
4. Do individuals engage in self-manipulation to avoid helping and how does this relate to attitudes towards information?

In response to question two, the results suggest the importance, within this group of respondents, of sadness, felt pressure to help and perceived responsibility of the potential beneficiary to get themselves out of the situation.

In relation to the first question, and building on the results of the second, it is suggested that within the group these emotional responses were linked to characteristics of the respondents and their contexts. When the potential beneficiaries were seen as less to blame, were suffering more and/or were more similar, respondents reported feeling sadder. This sadness was in turn related to higher donations. Similarly, when respondents felt the potential beneficiary was suffering more, that they were more similar, that they would benefit more from their donation and that there was more urgency, they felt a greater pressure to help that was itself related to higher donations. Critically, in response to question three, the results imply that this felt pressure is independent of sadness related to helping.

A clear cut yes/no answer to question four is difficult. Many of the results are suggestive of self-manipulation, but proof is hard to come by. That respondents appeared to be ever ready to read fault into situations, that they brought feelings into line with their decision, that they were seemingly happy to share blame for not helping Oxfam, all suggest the possibility of self-manipulation.

There was not sufficient variability to take this examination to the next level of asking what specific characteristics and contexts were linked to perceptions of urgency, or similarity, or other responses. To some extent, this was covered in the within-case analysis, but it could be an area where further research would be helpful.

At the level of individual motivation, it would appear that the balance model has provided an interesting and potentially useful framework within which to consider decisions relating to the helping of distant others. The discussions here have focused on the individual, so as to isolate as far as possible individual responses. If, however, the framework is to be used to consider individual actions in relation to real-world events, the social context must also be considered. The following chapter returns to the process of developing the model by adding in this consideration of the broader environment.

Chapter 6 Appendix: Detailed results

Chapter 6 was kept deliberately as non-technical as possible to focus discussion on the implications. The tables provided in this appendix detail results of the regressions conducted and discussed in the chapter. All regressions reported on were conducted using STATA version 8⁸. A description of the results reported on in each table is provided below; this is followed by the results themselves with notes to the tables where appropriate to draw attention to specific results.

Table A6.1 presents a detailed summary of the ordered probits conducted for each case, comparing the set one version to the set two version in terms of amount donated.

Table A6.2 presents a summary of the regressions conducted, comparing versions of each case in terms of donations and all follow-up questions. For each dependent variable three regressions were conducted: an ordered probit, a tobit and an OLS regression. For each of these, the sign and significance of the coefficient on the set one dummy is reported. A significant coefficient on this variable signifies a significant difference in the average amount donated or feeling reported between the two versions of the case.

For the ordered probits, a test of the parallel regression assumption was conducted. This is what the log likelihood test refers to. If the assumption can be rejected at the 0.05 level, it may not be appropriate to use the ordered probit model.

If the coefficient on the set one variable is significantly different from 0 at the 0.05 level and neither of the models based on alternative interpretations of the

⁸ Stata Corporation, Texas: www.stata.com

variables is different in terms of sign or significance, a LEVEL 1 confidence is noted.

If the difference is significant in all three models but the parallel assumption is violated, then only a LEVEL 2 confidence is attached. Any other combination is treated as insignificant and indicated by a 3.

It is important to note that the tobit and OLS are not seen as replacements for the ordered probit if the parallel regression assumption is violated; this would be inappropriate as they involve stricter restrictions (Long and Freese, 2006). Rather, they are viewed as the appropriate model under a slightly different interpretation.

Table A6.3 summarises the results presented in Table A6.2. It highlights only those variables which were identified as being significantly different across the versions. The level of confidence in these results is also summarised and the direction of the relationship.

Table A6.4 summarises the regression results where 'donation' was the dependent variable and the follow-up questions the explanatory variables. In these regressions all cases were considered simultaneously. The results are reported for each of the three regression models used. The table also reports the final log likelihood for the ordered probit and the probability of all the coefficients being simultaneously zero, the results of the test of the parallel regression assumption, as well as goodness of fit measure for the OLS regression.

Tables A6.5 and A6.6 report the same as A6.4, but with the sadness variable and pressure variable as dependent variables respectively.

Table A6.1: Donation differences across sets by case

Explanatory Variable	Dependent Variable: Amount donated								
	Coefficient (P – value)								
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9
Set one dummy (1=faced set one)	-0.56 (.043)	+0.97 (.001)	-1.07 (.000)	+0.09 (.743)	-0.71 (.014)	+0.53 (.065)	+0.87 (.002)	+0.44 (.117)	+0.37 (.192)
Non-current average	+0.89 (.000)	+0.72 (.000)	+0.62 (.000)	+0.73 (.000)	+0.92 (.000)	+0.75 (.000)	+0.49 (.000)	+0.91 (.000)	+0.87 (.000)
Location dummy	+1.14 (.001)	+0.38 (.259)	-0.19 (.579)	-0.18 (0.591)	-0.99 (.005)	+0.38 (.275)	+0.72 (.032)	-0.19 (.581)	-0.43 (.229)
Religious dummy	-0.69 (.016)	-0.04 (.851)	+0.45 (.118)	-0.11 (.688)	+0.55 (.051)	-0.36 (.213)	-0.47 (.093)	+0.61 (.031)	-0.31 (.279)
Age	-0.04 (.233)	+0.01 (.766)	-0.02 (.360)	+0.01 (.705)	-0.00 (.918)	-0.11 (.002)	+0.02 (.586)	-0.04 (.223)	+0.02 (.524)
Sex	+0.26 (.323)	-0.05 (.866)	+0.21 (.422)	-0.01 (.956)	-0.07 (.789)	-0.22 (.413)	+0.15 (.574)	-0.15 (.588)	+0.21 (.457)
Asia ties dummy (1=yes)	+1.14 (.004)	+0.02 (.958)	+0.15 (.715)	-0.45 (.250)	+0.54 (.166)	-0.68 (.097)	-0.24 (.536)	+0.48 (.223)	-0.22 (.571)
Africa ties dummy	+1.12	+1.58	-1.24	+0.59	-0.45	-0.65	+0.65	-0.58	-0.14

(1=yes)	(.117)	(.028)	(.084)	(.406)	(.538)	(.371)	(.359)	(.429)	(.852)
Eastern Europe ties dummy (1=yes)	-.57 (.371)	-.19 (.747)	-.50 (.397)	+.12 (.833)	-1.23 (.044)	+.49 (.439)	+.37 (.530)	+.23 (.714)	-.66 (.287)
Financial status 2 dummy	-.03 (.955)	-.93 (.074)	-.17 (.759)	+.81 (.127)	+.41 (.438)	-.52 (.306)	+.70 (.182)	-.11 (.838)	-.31 (.5570)
Financial status 3 dummy	-.60 (.240)	-.01 (.982)	-.07 (.900)	+.62 (.215)	-.29 (.575)	-1.12 (.024)	+1.29 (.011)	-.32 (.538)	-.03 (.960)
Financial status 4 dummy	-.23 (.722)	-.41 (0.53)	-.37 (.582)	+.61 (.352)	+.88 (.185)	-.19 (.074)	+.52 (.423)	-.11 (.868)	-.13 (.850)
Financial status 5 dummy	-.72 (.421)	-.172 (.058)	-1.78 (.046)	-.23 (.796)	+.96 (.324)	+.86 (.374)	+.108 (.231)	+.98 (.295)	-.81 (.410)
Log likelihood	-115.96	-116.22	-119.98	-113.81	-107.90	-115.22	-130.04	-109.53	-121.522
Probability of Chi ²	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes:

The final two rows of the above table indicate that all nine regressions were significant as the hypothesis that all coefficients are simultaneously equal to zero can be rejected at the 0.0000 level for each of them. The relationships which are significant at the 5% level are highlighted although the primary interest is in the significance and sign of the coefficient on the set one dummy. If this coefficient is significant the hypothesis that the small change in wording led to a significant change in amount donated cannot be rejected. A negative sign indicates that the set one version received on average less and a positive sign that it received on average more.

Table A6.2: Case version differences in donations and follow-up questions

	Sign on set one dummy or value of test (p – value)								
Donated	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9
Oprobit	-(.043)	+(.001)	-(.000)	+(.743)	-(.014)	+(.065)	+(.002)	+(.117)	+(.192)
Likelihood ratio test	148 (.0006)	145 (.0062)	113 (.0230)	97 (.0682)	134 (.0023)	109 (.2260)	96 (.4400)	101 (.6071)	118 (.1206)
Tobit	-(.067)	+(.002)	-(.000)	-(.969)	-(.003)	+(.078)	+(.002)	+(.142)	+(.166)
OLS	-(.051)	+(.007)	-(.000)	-(.763)	-(.001)	+(.181)	+(.004)	+(.386)	+(.138)
Confidence level and sign	-3	+2	-2		-2		+1		
Q1 – sadness									
Oprobit	-(.000)	+(.003)	-(.001)	-(.023)	-(.383)	+(.108)	+(.003)	+(.242)	+(.000)
Likelihood ratio test	107 (.0872)	149 (.0003)	121 (.0078)	143 (.0001)	130 (.0019)	104 (.355)	122 (.0076)	132 (.0014)	121 (.0072)
Tobit	-(.001)	+(.003)	-(.001)	-(.015)	-(.415)	+(.104)	+(.005)	+(.231)	+(.000)
OLS	-(.001)	+(.008)	-(.007)	-(.015)	-(.369)	+(.188)	+(.013)	+(.361)	+(.000)
Confidence level and sign	-1	+2	-2	-2			+2		+2

Table A6.3: Summary of significant differences across versions of each case

	Sign (set 1 compared to set 2) and confidence level								
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9
Donated		+2	-2		-2		+1		
Q1 – sadness	-1	+2	-2	-2			+2		+2
Q2 – pressure		+1	-2				+1		+2
Q3 – fault	*	*	+2				-2		
Q4 – Res. to get out			+2	+1			-1		-1
Q5 – suffer		+2					+1		+2
Q6 – responsible	-2	+2		-2			+1		+2
Q7 – similar						-2		+2	
Q8 – impact							+2		
Q9 – urgency		+1							

*Insufficient variation to estimate regressions

Notes:

Tables A6.2 and A6.3 examine the robustness of the results relating to within-case variation to model selection. A confidence level of 1 or 2 was considered acceptable to discuss the results as significant in the text of the chapter.

Table A6.4: Donations regression

Explanatory variable	Oprobit	Interval regression	OLS
Non-current average	+0.709(0.000)	+1.196(0.000)	+0.973(0.000)
Sadness	+0.136(0.000)	+0.203(0.000)	+0.185(0.000)
Pressure	+0.158(0.000)	+0.207(0.000)	+0.227(0.000)
Blame	-0.008(0.788)	-0.013(0.788)	-0.002(0.963)
Res. to get out	-0.102(0.000)	-0.175(0.000)	-0.139(0.001)
Suffering	-0.030(0.254)	-0.047(0.298)	-0.040(0.276)
Similarity	-0.016(0.407)	-0.038(0.257)	-0.029(0.321)
Marginal impact	+0.029(0.278)	+0.044(0.337)	+0.034(0.392)
Urgency	+0.027(0.057)	+0.040(0.124)	+0.045(0.011)
Constant	-	-0.930(0.000)	+0.097(0.017)
Log likelihood	-1127	-	-
Probability	0.000	0.000	0.000
Likelihood test	67 (0.87)	-	-
Goodness of fit			0.75

Notes:

The results indicate that model selection makes little difference in terms of the results of interest. The sign and significance varies only for the coefficient on reported urgency which is significant in the OLS regression and not the others.

The likelihood test indicates that the parallel regression assumption does not appear to be violated.

The high R^2 for the OLS regression indicates the strength of the relationships observed.

Table A6.5: Sadness regressions

Explanatory variable	Oprobit	Interval regression	OLS
Non-current average	+0.579(0.000)	+1.019(0.000)	+0.911(0.000)
Blame	-0.095(0.000)	-0.163(0.000)	-0.152(0.000)
Suffering	+0.315(0.000)	+0.559(0.000)	+0.503(0.000)
Similarity	+0.114(0.000)	+0.204(0.000)	+0.177(0.000)
Constant	-	-0.063(0.753)	+0.416(0.000)
Log likelihood	-1328	-	-
Probability	0.0000	0.0000	0.0000
Likelihood test	43(0.0868)	-	-
Goodness of fit			0.58

Notes:

The results suggest that model selection makes little difference. All three regressions are significant, the relationships are strong and the parallel regression assumption is not violated.

Table A6.6: Pressure regressions

Variable	Oprobit	Interval regression	OLS
Non-current average	+0.635(0.000)	+1.030(0.000)	+0.938(0.000)
Res. to get out	-0.077(0.017)	-0.122(0.019)	-0.113(0.013)
Suffering	+0.229(0.000)	+0.369(0.000)	+0.340(0.000)
Similarity	+0.090(0.003)	+0.151(0.002)	+0.130(0.003)
Marginal impact	+0.090(0.000)	+0.145(0.000)	+0.128(0.001)
Urgency	+0.085(0.000)	+0.138(0.000)	+0.129(0.000)
Constant	-	-0.214(0.240)	+0.211(0.041)
Log likelihood	-1278	-	-
Probability	0.0000	0.0000	0.000
Likelihood test	68(0.0310)	-	-
Goodness of fit			0.60

Notes:

The results suggest that model selection makes little difference and that the relationships are strong. All three regressions are significant. The parallel regression assumption is, however, violated which suggests that the order probit may not be appropriate.

Chapter 7: The individual in context

7.1 Introduction

Responding to distant others involves, by definition, distance between those that need help and those who can provide it. The individual responding does not experience the need first-hand nor, typically, do they have the means to respond directly. They rely on others both for the experience of the need and the facilitation of a response. Thus far, the discussion has focused on the individual's internal response: how they feel and value different outcomes. This chapter seeks to place that discussion of the individual in a social context, so as to better appreciate the implications in real-world scenarios.

In the preceding chapters, an argument has been made in favour of the balance model. It has, however, been repeated throughout that the results of the balance model can be replicated with a maintenance of self-image model, under the more conventional utility-maximisation framework. In this chapter, the aim will be to place the balance model within a social context; it is considered sufficient this time to note that the same could be done with the self-image model with largely the same predictions resulting. The reason, which has been given a number of times, for demonstrating that both the balance model and the self-image models can generate the same predictions is so that the acceptance of the predictive conclusions of the argument are not contingent on the rejection of utility maximisation that is implicit in the balance model. While not predicatively different, the models have different implications in terms of the approaches to applied research each suggests, and because of the rejection of utility maximisation in the balance model. This latter difference is the topic of the next chapter.

Hauser, in his recent (2007) work on morality, repeats a variant of a common set of scenarios. In the first, an individual bears great cost to help a child in need; in the second, the individual refuses minimal cost to do the same. The difference in the scenarios being that, in the first, the individual is physically confronted with the need when they come across the child; in the second they are informed of it through a charity appeal. Hauser argues that we all possess a moral grammar from which we construct our moral attitudes, and that this was developed when all human contact was only with those immediately around us, so it was not developed to deal with distant need. Hauser may well be right, although it is unclear what distant means in this case: over the hill could be distant to some at some points in history. What is clear is the difference in being confronted by need first-hand or having the information on the need relayed through other parties. It is one thing to note the difference, but what is of interest here is that more can be said about the nature of the system that leads to these differences.

This chapter will argue that we do possess a response to the needs of others. Such a response must result from a reading of the situation and an assessment of the implications of that reading. It is not important at this point whether this is a conscious or unconscious process; what is important is that it is an assessment and conversion process, assessing observed/gathered/received information on a situation and converting it into a feeling/motivation/desire. Such assessment and conversion opens up the possibility for alternative assessments and resultant conversions; again, it does not matter at this stage whether it is conscious or unconscious. The process would be difficult to explain, in the same way as the possibility of denial has had to be battled with. While the process of denial has been difficult to pin down, its existence has frequently, although not universally, been accepted. The balance model suggests that there is a possibility for self-interested motives to interfere with the assessment and conversion processes, so that the implications of information may be interpreted favourably.

When physically confronted with a child in need, it is more difficult not to assess and convert the information received to demands on you to respond. At a distance, with limited information received, and then only through others and where a response can only be made through others, the possibility of alternative assessments of the situation is great. If then there is a motivation for the demands not to be generated – as they would then have to be addressed – the selection of assessments and the associated conversions, where this is possible, will likely favour lower demands on the self.

This chapter is concerned then with how the flow of information on the needs of distant others to the individual allows for, or influences, the selection of assessments. From this discussion it will become clear that the alternative assessments available will, to some extent, be determined by the selection and presentation of information. To understand the impact of the social context is then to understand the shaping of information flows. In the context of distant others, this relates directly to the selection of media and, in particular, news content. Indirectly, the entire social context plays a role in shaping information flows and interpretations. The discussions in this chapter will concentrate on the direct influences, but it is important to keep the broader social context in mind.

The flow and translation of information to the individual is, however, only one side of the context. The other is to consider the impact of the response of individuals to this information. There is, of course, the possibility of charitable contributions. There is also the possibility of the influence of public opinion on government policy. This is an area that has attracted much academic attention; some insights from this literature will be reviewed, so as to provide a more evenly-weighted picture of the social context.

Social contexts are complicated and any attempt to apply a reductionist approach to isolate certain relationships is bound to be flawed. The previous six chapters have focused on the individual, but understandings of the individual are of little use apart from the context in which they operate. So, while flawed, the following discussions are necessary to begin to consider the implications of the arguments outlined in the preceding chapters.

7.2 Information preferences interaction and the role of the media

In economics it is often assumed that preferences are stable or even fixed. The problem is that it is often assumed that individuals have preferences over objects and states, as opposed to internal feelings. An individual values an object or a state because of the way it makes them feel. Information on an object or state may well change the way in which an individual feels about it. The context, or the needs of an individual at a particular time, may well interact to generate different feelings towards the same object at other times. It is not that the individual's preference for an object changes but that their response to it may change. In this way of thinking, the individual does not even have a preference for the object in and of itself. The balance model suggests by its name that the individual has a preference for balance of their various different and non-substitutable feelings. The feelings that come into play in any particular decision relate to what is considered appropriate, which is determined by the information available. To place the balance model in a social context, it is therefore very important to consider how information on distant others reaches individuals and how this might affect different feelings, depending on how it is presented. Fortunately, research on charitable giving specifically, and the media in general, have considered related issues and insights can be drawn from these studies.

In reports involving distant others, two issues are relevant for extending the model to consider the information-flow aspect of social context: firstly, how the

nature and representation of information can alter the response of the receiver of that information; secondly, that information is only received on those events that are considered newsworthy. Studies of response to distant need, and other areas of consideration that do not involve first-hand experience, have consistently identified these issues (Kinder, 1998). When distances are great, the latter issue is particularly important, as it is the media that turns a local event into an international one, if it is considered newsworthy (Cohen, 2005; Bennett and Kottasz, 2000). It is, however, not only a question of what information reaches individuals, but also of the form in which it reaches them that is important. For example, reports involving highly-emotive imagery have been linked to higher likelihood of donations, while reports concerning inequitable distribution of aid or welfare would reduce support (Bennett and Kottasz, 2000).

As this chapter intends to expand on the balance model, which has focused on the individual, by introducing elements of the social context, the discussions will work from the individual outward. This section will consider how the presentation of information has been linked to different responses. The following section considers what information is selected by the media. The literature specifically on distant others is limited, but insights can be drawn and considered from other areas of related research, such as reporting on the poor, social welfare and political contests. These discussions deal with the way in which mass communication can influence individuals who are distant from the events and who are faced with a complex environment. There is too much going on and individuals have to find ways to make sense of the situation in order for them to make related choices. While there are similarities in these topics, there are also significant differences. Where appropriate these similarities and differences will be discussed.

Before moving away from the individual, it is worth noting the treatment of the individual in this area of research, as it will help in understanding the influences

of information that are suggested. Two conceptualisations of the individual decision making process are commonly used in the literature: memory-based models and online processing (Druckman and Lupia, 2000).

The most extreme form of memory-based conceptualisations found in the literature are comprehensive memory-based models, which essentially assume that, when considering a preference, individuals recall all information stored on the object under consideration (Druckman and Lupia, 2000; see for example Enelow and Hinich, 1984). Many authors argue that individuals do not have the means or inclination to conduct such thorough searches all the time and simpler alternatives have been proposed (Druckman and Lupia, 2000; Kinder, 1998; van der Pligt and Eiser, 1984).

An example of a memory-based approach that is frequently discussed in the literature, and which does not rely on a comprehensive search, is the use of stereotypes (Kinder, 1998). This line of thinking divides memory into two parts: the working memory and the long-term memory. The suggestion is that, when individuals are presented with information, they scan their long-term memory to identify a satisfying, although not necessarily optimal, schema that helps to identify the appropriate response. This then opens the way for the media to shape the debate by alternative framings. The cues contained in framings prompt the application of a particular schema that influences the conclusions by influencing the agenda of how the subject is considered. Other non-comprehensive models include those based on accessing information that is more accessible (Fazio, 1995), possibly because it has recently been used (Zaller, 1992).

Online models, on the other hand, suggest that individuals receive information on issues or political candidates, for example, process it, use it to update their perception on the issue or candidate and then often forget the information, remembering only the updated perception (Kinder, 1998). Online models

essentially assume a running counter: they argue that individuals develop preferences and then might well find it difficult to explain why they prefer one option to another (Druckman and Lupia, 2000; see for example Bassili, 1989). Asking why a preference is held is then believed to prompt rationalisation (Lodge and Steenbergen, 1995). Such a conceptualisation casts preferences as far more stable than those based on non-comprehensive, memory-based models (Krosnick and Brannon, 1993).

It is, of course, possible that individuals could use both online and memory-based evaluations. Arguments have been made that suggest that individuals use online evaluations when they expect to be asked, and memory-based when they are presented with an unexpected call (Druckman and Lupia, 2000; Hastie and Park, 1986). The frequency of use of each type of process is, however, not agreed upon (see for opposing views McGraw and Pinney, 1990 and Zaller, 1992).

Using such models as those described above, authors have been able to examine the link between how information is represented and the response of the individual to it. Their interest has, however, typically not been in the individual alone but in public opinion. More specifically, the interest has been in how alternative ways of providing information influence the formation of public opinion. In these examinations, three topics have emerged as important and have attracted the most attention: framing, priming and persuasion.

Regarding framing, Iyengar provides a number of important contributions to the debate and it is useful to discuss his work in some detail. He argues that perceptions of causal and treatment responsibility – that is, who is to blame for the situation and who has the power to rectify it – play important roles in the formation of public opinion (1989). He suggests that individuals use responsibility constructions, which they spontaneously form, to make sense of distant and complex political issues and use these constructions as the primary

consideration in opinion formation. He goes on to argue that these are not firm formations and that the presentation of issues in the media can influence the attribution of responsibility and so influence the opinion formed. Iyengar takes this argument further, examining the case of poverty in particular (1990). Much of the research in this area relates to the presentation of the poor – the “deserving” and “undeserving” poor (Katz, 1989) – and the impact this has on public opinion and policy. While not the same as responses to distant others, research on responses to the poor have generally focused on those who do not have direct experience of poverty; in this way the issues are similar. The issues are, however, different in that, while individuals may not have direct experience, the poor may not be so distant. Proximity may well play a role in motivation to avoid engagement or to admit responsibility. This issue is discussed in more detail in regard to constructions of responsibility in the following chapter.

In his work on poverty, Iyengar argues that the way in which poverty is framed in the media or survey questionnaires influences the opinions of respondents as to the nature of the solution and the level of help that should be provided (1990). This approach presents opinions as flexible, not as innate traits based solely on dispositional factors. This is not to argue that dispositional factors are not important, or that opinions are weak preferences, but rather to suggest the importance of context in the decision-making and opinion-forming process. Iyengar found that, where the focus was on individual situations of poverty, the likelihood of supporting help was weaker. Moreover, and with obvious similarities to the results presented in the previous chapter, he found that respondents were likely to infer the worst case scenario and would readily accept cues relating to laziness or low morality.

Iyengar’s work on poverty provides a good example of framing. The basic argument is that many issues are complex and can be seen from different angles. How an issue is framed defines the nature of the problem and, by implication, often the nature of the solution. Presenting an individual with the

same information, but with a different frame, can elicit a different response (Gamson, 1992).

Priming relates to attention being given to a particular issue, resulting in that issue being considered more often when evaluating options. In a review of the literature, Kinder cites a number of studies that have shown that the presentation of an issue prompts that issue to become the focus in later decisions (1998). The more frequently an issue is mentioned in the media, the more attention individuals will give it in related decision making. This has been noted, for example, in studies linking television news content with evaluations of political candidates (Iyengar and Kinder, 1987). In the Iyengar and Kinder study, participants, after being exposed to news focusing on a particular issue, were asked to rank presidential candidates. If inflation was stressed in the news then economic policy was a focus of the evaluation; if defence issues were stressed in the news, they were the focus of the evaluation. Similar evidence of the importance of priming has also been noted outside of the experimental setting (Kinder, 1998; see for example Krosnick and Kinder, 1990; Stoker, 1993).

Priming and framing together are seen to play an important role in political debate, with arguments being made that it is not confrontation that is the key, but attention. It is important to note that both of these work better with memory-based models, particularly the framing. They suggest how individuals may form views and select alternative responses with the information at hand. The situation may well be different with strongly-held views – an issue which is considered when examining the possibilities for persuasion.

The evidence on the role of the media in terms of persuasion is more mixed (Druckman and Lupia, 2000; Kinder, 1998). There is a lot of interest in this area, particularly in political science, given the importance of persuasion in political processes. There are similarities to the focus on distant others in that

advocates for distant others try to persuade people to respond. The difference again is the proximity. While domestic political issues may be complex, and many of the decisions made have no direct affect on many members of the general public, they are tied up with everyday life and identity in a way in which the needs of distant others may not be.

With reference to the political process, a number of studies have shown the limited impact of, for example, presidential campaigns, as it appears that decisions are largely made even before the campaigns begin. This may just be because one campaign neutralises the other (Bartels, 1992) but, even if this is not the case, there does still appear to be a role for such campaigns to reinforce decisions (Kinder, 1998). There are areas in which the evidence suggests campaigns can be more successful, but it does require some searching to find this evidence, as it appears that this is not often the case (Kinder, 1998).

One very interesting conclusion relating to persuasion, which comes out of the political science literature, is that individuals who are highly involved rarely change their views, while those not at all involved hardly ever receive messages (Popkin and Dimock, 2000; Zaller, 1992). Druckman and Lupia note that this finding is increasingly common (2000). Those heavily involved have strong views and are not easily swayed from them, while the uninvolved are not linked into information channels, leaving those in between to be persuaded.

The above point in regard to the invariability of the highly involved and informed is particularly important, as a number of authors have argued that individuals are informed by the decisions made by well-informed others or of those perceived as leaders (Kinder, 1998). It is not always that they follow a leader; it may be that they conclude that, if a particular group is in favour of an option, then they should be against it (Kinder, 1998). Druckman and Lupia's review identifies a whole range of characteristics of the information source that play a part in the persuasive influence of the message; these include trustworthiness,

public approval, likeability and ideology (2000). Related to following leadership prompts, group identity is also argued to provide a short cut to conclusions. If a particular view is associated with a particular group, such as a political party, then individuals link their support to their loyalty to the party (Kinder, 1998).

The above-mentioned literature indicates that priming and framing are both important in determining what to think about and how to think about it. While the evidence on persuasion is less robust, the importance of opinion leaders and group identity is clearer. Drawing much of this together is Zaller's attempt to develop a theoretical system for examining the formation of public opinion (1992), which is perhaps one of the most important contributions in this field. He argues that people have a variety of opinions on any particular issue and that these may well be conflicting. Moreover, individuals do not take all these into account every time they consider the issue. The considerations that are taken into account are those that happen to be on the individual's mind; this links into the priming and framing arguments. In terms of how individuals will respond to messaging, Zeller proposes two axioms: the first is that individuals receive information on an issue proportionate to the level of engagement that they have in the issue; the second is that individuals will resist information that is inconsistent with their currently-held views. Zaller's conception of the individual has many similarities to the balance model, so his theoretical system, linking the individual to information flows, is very useful. The arguments relating to resistance of information are particularly relevant.

Priming and framing can both be linked to the balance model more directly. Priming suggests that information selection can affect what becomes the focus, thereby influencing what is considered. Framing, on the other hand may, to some extent, influence which responses are activated and the nature and magnitude of these responses, particularly when linked to apparent motivation to infer scenarios that limit responsibility to respond. Also interesting with

regard to the balance model are the apparent limited impact of persuasion and the strong impacts of group identity and leadership. Both of these suggest that individuals take a position and then seek to hold it. As mentioned, a number of studies argue that communications do not influence the heavily involved, because they have already made up their mind, but they also do not influence the very uninvolved, as they never encounter them. This leaves the moderately informed to be influenced. Having selected a position, moving from it may well put an individual out of balance on a variety of fronts. There is, therefore, an apparent motivation not to seek the truth but simply to be comfortable with the facts. A motivation to defend your construction of an understanding of the world, and your place in it, is an issue that is discussed in more detail in the following chapter.

When responding to information flows in a social context, it is important to consider the role of social norms. Social norms, which have been mentioned in previous chapters, may be important when considering responses to information on different issues. For example, norms may provide guidance on which schema to apply or shape, and whether certain behaviours or situations are good or bad. In certain circumstances, cultural understandings may dominate the way in which information is responded to. Cohen's (2005) discussions of cultural denial are central here: when everyone knows and doesn't know and everyone knows that everyone else knows but doesn't know. The integration of social norms into the balance model framework is a little involved and is taken up in detail in the next chapter. What is important here is that individual responses to information may well be linked with other individuals' responses, or at least what they believe others responses will be.

The ideas of priming and framing, and the motivated interaction with information suggested in the Zaller approach, all help in expanding the balance model. All of these points highlight the importance of what information is available and how it is represented. As few individuals are likely to be deeply involved in the

problems of distant others, and while they may well hold views, these are likely to be general. The important question in terms of thinking about responses to distant others is what gets coverage in the media and how those involved choose to present it.

7.3 Media selection and presentation

There are various forms of media that can carry information about what is happening elsewhere. This includes, but is not limited to, the news media. Documentaries, movies, fiction and non-fiction books, television series and many other forms of media, aside from news media, can contain information relating to the situation and needs of distant others. And there is, of course, the deliberate awareness raising efforts of activists, interest groups and non-governmental organisations. The discussions here will centre on the news media, but most of the points raised can be transferred to other media forms. The question then is how news content is selected; it is not, however, just a question of what is considered newsworthy, but also of how information is collected and presented.

A number of interesting areas covered in the literature are relevant to the current discussion: one relates to the portrayal of natural disasters, as these are often distant and involve need; similarly the reporting of human rights abuses; while another relates to the portrayal of the poor, which is a well-studied area involving need and excuses. The issue of poverty is, as mentioned in the previous section, somewhat different, given the sometimes close proximity of the poor, and the inter-relationships between poverty and other domestic issues. In terms of drawing attention to the role of the media in supporting particular constructions of a problem, the studies of poverty are, however, extremely useful and will be discussed here in some detail.

In his discussion of the reporting of human rights abuses, Cohen identifies three models of news selection: correspondence, arbitrariness and pattern (2005). The correspondence model involves the rational selection of news content based on criteria such as the seriousness of the events being reported. Cohen argues that this model does not appear to be applied. He similarly argues that there is little evidence at the other extreme where the arbitrariness model is located. On the basis of this model, events will be reported on largely by chance, such as the chance availability of a reporter at the scene.

The pattern model suggests that political, cultural and organisational factors may result in a systematic selection of news content. When considering selection, filtering and framing, it is important to keep in mind the role of the media in political processes and even, as Cohen points out, in the maintenance of cultural denials. These patterns may well be linked to seriousness but will also be affected by a range of other factors, including audience interest, clarity and focuses on particular types of events, such as those involving violence or sudden crises.

Interlinked with these patterns is what Cohen argues is the most important determinant of selection: whether or not the story is already in the news. The media are self-referential (Cohen, 2005: 172).

As mentioned, much of the news selection process is caught up in political processes and cultural constructions; it is unclear, therefore, to what extent those involved will be entirely open about their selection processes. This is particularly true for human rights abuses, as they are typically highly-charged political issues. Natural disasters, while often having a political element, can often be reported without it; examining selection in this field may, therefore, be more beneficial. Bennett and Daniel discuss findings from having interviewed editors of newspapers and senior journalists about what makes disasters newsworthy, how they go about collecting information on them and how they

rate the information they receive from major charities (2002). In terms of determining newsworthiness, a number of factors were identified. Key among these were: human interest, individual bravery, enormity and some of the victims being British nationals (it was a study conducted in the United Kingdom). In terms of which aspects of disasters are of interest, the possibility of highlighting corruption in response efforts was seen by many as an area of interest. A number also felt that it was important to show the suffering. The responses would lend weight to the pattern model, showing how seriousness may come into it, but only along with other factors.

The above example suggests a patterned selection of events to report, but the media's potential impact, as discussed in the previous section, goes beyond determining content and associated priming. The manner of reporting, particularly in terms of framing, is also critical. Events are not simply considered important, selected for coverage and then reported on. The news is not the conveyance of some objective reality from site to recipient.

Information gets distorted and shaped during the transmission process. To some extent, this may be a by-product of the collection process; alternatively, it may be a result of deliberate efforts to shape content in particular ways. At times it is difficult to distinguish the distortion that is a by-product of the process from more deliberate efforts to shape the news. For example, Campbell argues, on the basis of a Scottish study, that, given the nature of the news process and the tight time constraints involved, journalists favour the easiest information source; this is more the case for television than for print (1997). Given this approach, she suggests that both the meaning and structure of information are distorted in the course of the news formation process, as it is simplified and condensed. This could be an example of process distortion, but Campbell goes on to note that the simplification and condensing are conducted in a manner seen as appropriate for the audience. Campbell suggests that information collection by journalists follows operational rules, as does the

editorial process. The process, she argues, means that the news cannot be considered objective, particularly when you think of the easiest route being taken for information collection, combined with a picture of the needs of the audience in mind.

The media are often not alone in the deliberate framing of messages. Returning to Campbell, she found that reporters favoured seeking expert opinions rather than searching through printed material and media libraries. This, she argues, often occurs if the nature of the information is complex, given the ability of experts to summarise. She further argues, however, that the selection of the initial expert to talk to is shaped by audience considerations and that this initial selection, to a large extent, shapes who else is spoken to, as the journalists often follow the recommendations of the initial expert, thus potentially biasing the views presented.

The importance of relationships between those with information and the press is critical. There is obviously distance involved when considering responses to distant others. Couple this with the complexity of situations which often surround the needs of distant others, and the importance of relationships between the press and those with information is magnified. This is something that charities are becoming increasingly aware of, and more strategic about, given the increasing reporting on disasters in the media (Payne, 1994; Smallman, 1997) and the perception by charity fund-raisers that the media focuses on the dramatic and sensational, the perpetuation of myths and stereotypical viewpoints, and even racism (Bennett and Kottasz, 2001).

These informational relationships have been seen as particularly important in disaster reporting. There appears to be a reliance on traditional, often formal, sources of information and this can result in distorted reporting, as information is not gathered on all efforts; smaller and local responses are excluded (Quarantelli, 1996).

The time pressures on reporting disasters may then lead to somewhat understandable distortions and could be linked with process distortions. The problem, however, seems also to be present in the reporting on poverty, which has been around for a long time, with plenty of time to check stories. An examination of just a few of the important studies in this field highlights the possibility of distortions that cannot simply be explained away as by-products of the information collection and condensing process.

A study that has generated a great deal of debate in this area has been Gilens'(1996) analysis of reporting on poverty on television news, and in weekly news magazines, in the United States of America. Examining the images of poverty presented, he found that African-Americans were over-represented, in that they appeared in a higher percentage of the images of the poor than they are a percentage of the poor. He also found that groups, such as the elderly, who were more likely to evoke sympathy, were under-represented. This under-representation was even more pronounced for African-American poor. African-Americans were over-represented in negative images of the poor, such as the unemployed and the under-class, but not so in the portrayal of the working poor, the elderly and stories on Medicare. This, he argues, fuels the belief, held by many, particularly white Americans, in the over-representation of African-Americans among the poor. A belief that, he argues, is linked to opposition to welfare. He also noted the greater willingness of respondents to support help to the elderly, particularly the disabled, who were not at fault for their poverty. This similarly applied to children, although there were complications associated with supporting the parents of children so that the children could be supported.

Gilens goes on to discuss the influence of media representation of public opinion. He cites evidence that portrayals, and particularly images, influence perceptions and opinions. He also notes that changing public perceptions of

the composition of the poor are related to the depiction of the poor in the media, although he does note that the relationship may not be causal. Gilens also discusses an alternative explanation that the personal experience of poor people may drive perceptions, but he found that, even in States with very low African-American representation among the poor, respondents still over-estimated their representation nationally.

One of the most interesting aspects of Gilens' discussion, and the most relevant for this section, is his efforts to explain the media misrepresentation. He first examined the availability argument: that poor African-Americans are more available to be photographed. The argument that this would be the case because news media are based in urban areas and that the urban poor are disproportionately black was discredited. To reach the levels of representation in the media, the photographers would have had to focus only on the pockets of extreme and concentrated poverty that tend to be much more African-American. Even if this were the case, it would not explain the selection by photo editors and, particularly, not of the over- and under-representation across groups such as the working and non-working poor. Gilens argues that the journalists and the editors also suffer from a misconception of the true racial composition of the poor and this may be reflected in their reporting; though this alone would not be enough. He suggests that they may also indulge the public's misconception, so as to make stories more acceptable.

More recently, Clawson and Trice continued Gilens' work on presentation of the poor in media photographs (2000). They found similar racial bias in the presentation of the poor, as well as images that reinforce negative perceptions of the poor, such as laziness. They again note that, when the poor are presented positively, they are more likely to be white. The authors argue that this presentation influences public opinion and in turn has an influence on policy. They also point out that such inaccuracy is not only present in the news media, but also in other forms of media, such as television sitcoms.

One of the most interesting issues raised in these studies is the question of why the media would misrepresent the situation. The possibility that there are influences over the media will be discussed in the next section, linking the media, public opinion and policy in more detail. That editors may indulge their viewers' misconceptions is an intriguing argument, similar to the argument mentioned previously that journalists may select experts with their audience's views in mind. Why would they do that? If individuals do not want information that suggests that they have a responsibility to help, and they have a motivation to avoid such information, then editors would not want to report it, as they do not want their outlets avoided. It may be that individuals construct complex understandings of the world that keep the responsibility they feel in check. Media that challenges these constructions may well be avoided or resisted. Consider Cozzarelli et al's interesting study examining the relationship between individuals' perceptions of the poor, their belief in stereotypes and the beliefs they hold (2001). They found that individuals who had negative views of the poor, in terms of their attributes, were also more likely to hold to stereotypes that highlight individual causes of poverty. More strongly held work-ethic beliefs, and beliefs in a just world, were similarly related to attribution of responsibility to individuals. This all suggests the construction of a view and, indeed, most studies show that Americans favour individual explanations of poverty over those that focus on the role of the system, which might challenge their constructions (Cozzarelli et al, 2001), giving editors a possible reason to favour such. It would appear that there is a balancing act between responding to interest in suffering – the entertainment value of suffering – and framing responsibility.

Evidence that the media do not want to challenge constructions can be found in how reporting changes as the importance of such constructions changes. Bullock et al provide a review that concentrates on the research in the US on media images of the poor, prior to the major policy shift in the welfare system in

the mid-1990s (2001). They note a range of authors who suggest that the media helps perpetuate the belief among many in the US that the society is largely classless; they are all middle class. This is achieved by focusing on middle-class interests, such as financial reports, and the lack of attention paid to structural economic concerns. The poor are then left either invisible or portrayed according to negative stereotypes. The literature that they review notes this tendency, not only in news coverage but more broadly, such as in the absence of the poor in most sitcoms and their presence in reality crime shows.

In relation specifically to news coverage, Bullock et al note the lack of coverage, given the scale of the problem. What coverage there is tended towards negative framing, at least in the build-up to the reforms. The authors report on research that suggests that, during this period, welfare recipients were one of the most disliked and disrespected groups in the country, even though the majority of recipients were children; but, as the children were rarely the focus of stories, this factor may not have been considered. Mothers, in particular, were negatively portrayed, with a focus on their low morals, negligent behaviour towards their children and the damage they have done to the nuclear family. This framing keeps the debate centred on birth control and parenting skills rather than on jobs and support. The authors conducted their own research on the post-reform period and noted that the focus had shifted somewhat, with a growing focus on the success of reforms. Success, however, was measured by the reduced burden on the welfare system, rather than by reduced poverty. They note that the images of the poor in the media do appear to reflect the interests of the political elite. Such bias may help in maintaining individual and collective constructions, constructions which support a process favoured by the elite. This possibility is considered in more detail in the following section.

The literature on news selection and portrayal suggests that there is some evidence to suggest that the news is frequently biased in very particular ways. There may be patterned selection of events to cover, errors in collection, and

systematic distortions. Add to that the possibility of playing to the audience: television news needs viewers, newspapers need readers. Reporting news that they have a motivation to avoid would seem motivation to reframe or omit it.

The previous two sections taken together suggest that the media can influence public opinion, while at times it might play to it; priming and framing have been shown to influence decisions. Combine this with patterned selection, systematic distortion and shaping news content, and the possibilities for influence are clear. If the media influences an individual in relation to the needs of distant others, what impact can this influence have? It may prompt or reduce donations or other pledges of support to charities and international non-governmental organisations. Prompting donations is indeed the aim of charity appeals through the media. But this is an individual impact, what of the impact of public opinion? The key question becomes the extent to which public opinion can influence policy.

7.4 Opinion and policy link

The question of the link between opinion and policy has been examined from two related, but different, standpoints. One has been to ask the extent to which the media, through its impact on public opinion, can influence policy. The second is the more general question of the extent to which public opinion in general can influence policy. The discussions presented here will, picking up on issues raised in the last section, run from the first to the second standpoint.

The literature concerned with the first standpoint often relates to foreign policy. This focus appears essentially to assume a central role for the media in relation to informing domestic audiences of international events. Given the central role assumed for the media, the impact of the media on opinion and opinion on policy are treated collectively.

A good example of this approach of collective treatment is what has become known as 'the CNN effect': the argument that the media can drive foreign policy by influencing public opinion (Robinson, 1999; 2000). A major commentator on this approach is Robinson, who reviews existing work and notes that the availability of technology, that allows the media to provide global real-time news, has led many to discuss this in relation to its impact on policy. Furthermore, he outlines how some have argued that such interference is problematic, interfering as it does in the policy-formation process and leading to rushed and poorly-conceived responses. Such arguments are based on the assumption that the CNN effect does indeed occur and that the press does impact on policy. This, Robinson notes, is not easily reconciled with the manufacturing consensus line, which asserts that the media are mobilised, and even manipulated, by policy makers to support the government view. Having reviewed a number of key studies in the area, Robinson argues that the two theories need not be entirely at odds. Studies of the media's influence have suggested that it is possible that policy may be influenced, but this is generally the case when the policy line is unclear. This would suggest that the media have an impact when there is conflict among the elite relating to what policy should be. When policy is clear and largely free of conflict, the state can influence the media. Indeed, Robinson cites examples of where media action, far from only prompting action, appeared to excuse inaction, with the framing of news stories clearly being an important factor. This debate centres on the question of who controls the content of the news and how this control varies, given the political situation.

In summary, Robinson argues that it is when there is policy uncertainty – that is, when there is no consensus among the elite – there is the potential for media impact on policy. If the media present empathetic coverage of victims, along with critical coverage of the policy response, at a time when there is policy uncertainty, then there may well be a policy impact. If, however, there is policy certainty, Robinson hypothesises that there will not be scope for the media to

influence policy, even if it is critical. This model highlights the importance of not only framing and priming, but also of the policy context in shaping the potential impact of the media.

Whether public opinion and policy are related, independent of the media, is itself a hotly-debated issue. Page notes that there is substantial evidence, in the US at least, that public opinion and policy are related (1994). That is, that people's responses to survey questions show a relationship to policy stances. While this evidence is strong, there remain many unanswered questions, key among them being the question of causality. Does opinion influence policy, or policy influence opinion? Or is there some third force, such as elite manipulation, that influences both? Policy makers, too, can take note of events around them and may anticipate public opinion changes and so respond before the swing. In reviewing the evidence and the methods used, Page concludes that there is much that supports the view that opinion influences policy, but that the results are uncertain, given the methodological difficulties in identifying causality. He also notes the importance of asking what else influences policy.

In reviewing the literature, Manza and Cook identify two contrary positions: one arguing that there is a strong link and the other arguing that there is a weak influence of opinion on policy (2002). Similar to Page, they argue that much of the difference stems from the lack of appreciation of the interaction of public opinion with other factors, so that in some circumstances there is a large impact, in others a small one. For example, highly salient issues lead to greater cost to politicians if ignored, so they are more likely to be heeded. Other contextual factors identified in their review of the literature include the role of interest groups in a particular domain, the budgetary impact of the decision and the flexibility of existing policy

In his review/meta-analysis of studies on the links between public opinion and policy, Burstein argues that most social scientists studying the links between

public opinion and policy in democratic countries agree that: public opinion does have an influence; that the influence is greater the more salient the issue; and that interest groups can interfere with this relationship (2003). His concerns are the lack of agreement on the strength of these relationships, how they have altered over time, and their generalisability.

He suggests that there are few who believe public opinion has no impact, and none who believe it always determines policy. The debate is certainly about a question of degree. In the majority of studies he reviewed, the relationship was significant and was often substantive.

The heavy influence of interest groups, and political and economic elites, is often cited as a contextual factor that limits the role of public opinion. An alternative argument has, however, been put forward. There are those who suggest that the impact may not be negative and that such groups can act as amplifiers or intermediaries. Burstein's meta-analysis of the data, he argues, shows that the role of interest groups may be important but does not negate the influence of public opinion entirely and that there is evidence that they may at times increase the impact. (Burstein, 2003)

There does appear to be wide agreement in the literature that there is a link, but not about its strength. There is also agreement that other contextual factors interfere with this link. It is not, however, just the context but also the issue itself that may determine the strength of the relationship. Given the complexity of the issues, and the lack of understanding of the general public, a number of authors argue that only a few issues are important and that these, particularly if they are simple, are the ones that are open to influence. Burstein's review found that the few studies that had examined salience suggested that it deserves the theoretical attention that it has been given.

The bulk of the work linking public opinion to policy has been done with specific reference to the US context; other contexts have, however, been considered. Hobolt and Klemmensen, for example, discuss how the conclusions might differ between political systems, with specific reference to Britain and Denmark (2005). They argue that there is the possibility of policy leading opinion and opinion leading policy, but that it is the latter that dominates, particularly in political systems based on proportional representation. They find that, as in the US, there is a relationship, but they suggest that institutional arrangements do influence the nature and strength of this relationship.

While useful, the above literature linking opinion to policy does not often deal specifically with policy concerning distant others. Knecht and Weatherford, however, discuss the link between public opinion and presidential decision making in the US in relation to matters of foreign policy (2006). They argue that the influence of public opinion is strong when attentiveness is high. With this in mind, they then distinguish between crisis and non-crisis situations. They argue that there are different phases in attentiveness, which means that policy that is in line with public opinion can be announced when the public is attentive, but implementation can differ considerably from the stance taken as attentiveness fades. In crisis situations, attention peaks at the point of implementation, so responsiveness to public opinion is high. During non-crisis situations, the attention peaks at the point of the decision. This is where the opportunity for implementation to vary comes in. This is counter to the usual argument that the public are not interested in foreign policy that does not involve military action, arguing rather that their interest comes at a different point.

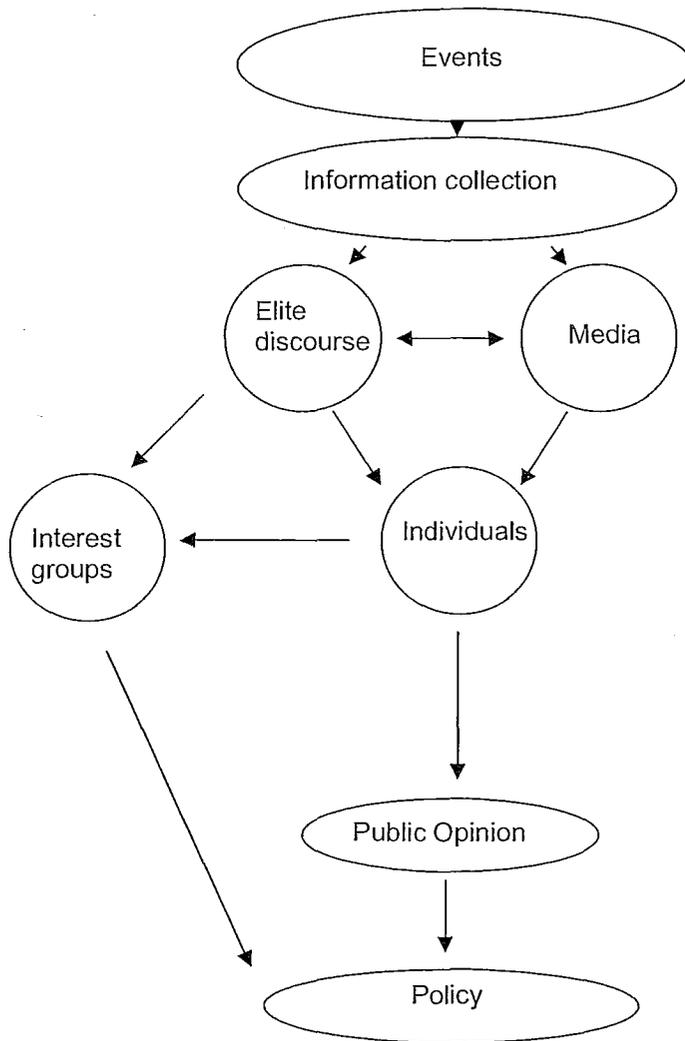
From the above, it appears that the possibility of public opinion influencing policy receives general support. The importance of policy context, elite unity, salience of the issue, the role of interest groups, the type of political institutions and various other contextual factors, were also noted. Considering how reports of the needs of distant others may influence policy through public opinion is,

clearly, a highly context-specific discussion. If the additional, related complexity of how the media select for coverage and choose to represent events is considered, it is clear that each individual situation will need to be examined in turn. What, however, the review of the literature in the above sections does provide is an understanding of some of the social context factors with which the individual interacts in determining responses, both personally and collectively. The following section draws on the insights of this literature to expand the balance model so as to consider the social context in which the individual operates.

7.5 Expanding the model to consider context

The research discussed above allows for the placement of the individual within a social context and for the associated linking to policy outcomes. The following figure outlines a simple model linking events affecting distant others with policy. To clarify, this refers to the influence on policy of individuals who themselves are not directly affected by events, because of their distance from them.

Figure 7.1: Expanded model, simple version



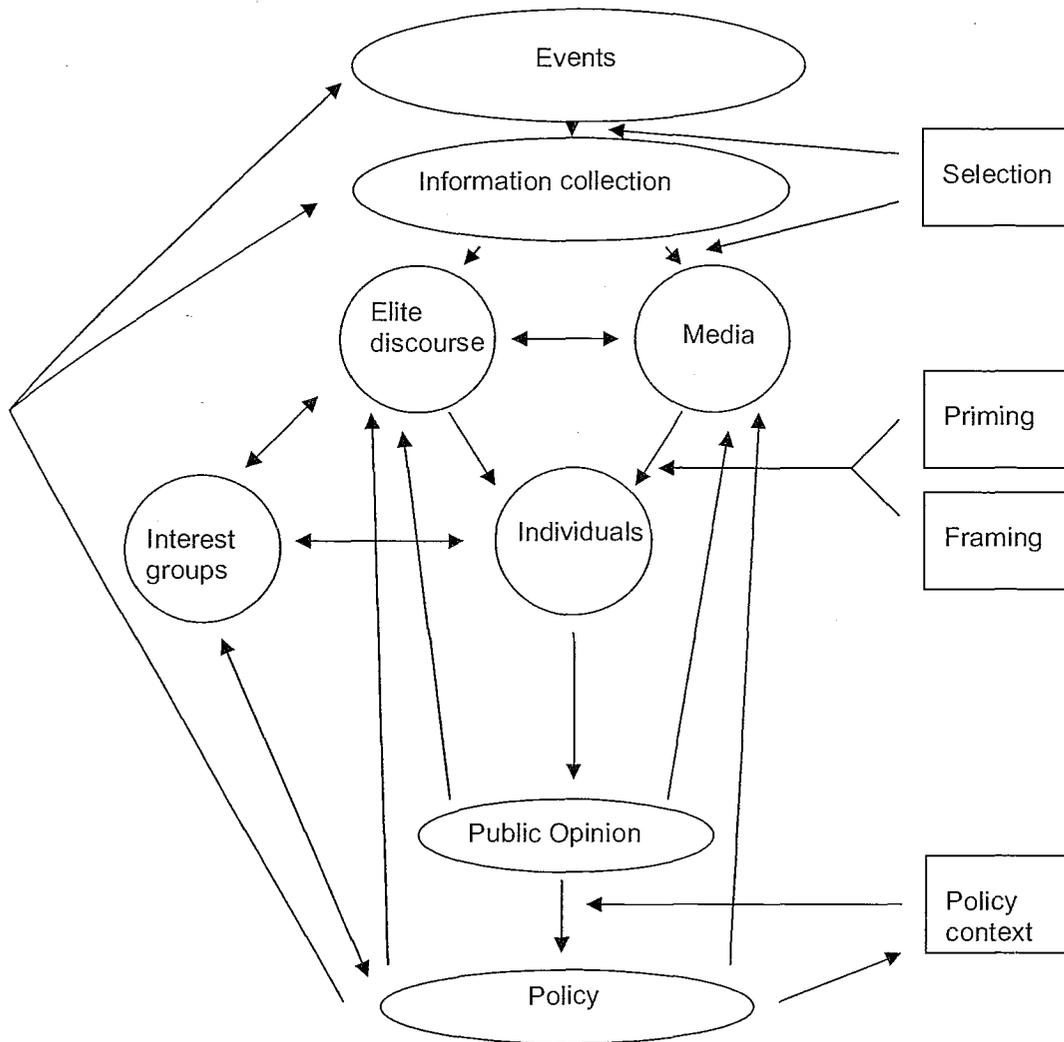
The model suggests that, as individuals are distant from events, they do not have direct access to information on them. Information is first collected, then transmitted by way of the media to individuals. It will also be passed on to the elite, which includes government policy makers. As outlined in the previous sections, the elite may then influence what the media report and the manner of that reporting. Individuals may also follow the lead of particular members of groupings from the elite. On the other hand, the elite may be influenced by what the media choose to focus on. The relationship is likely to be highly specific to the event and the policy context.

Collectively, these routes of the transfer of information influence individuals' views on events. This then is the impact on public opinion. Individual views may also influence interest groups, which can also be influenced by the elite. The relative strengths of these impacts on interest groups, and the power of interest groups to influence policy, determine whether they dampen or amplify the impact of public opinion on policy.

The above figure is highly simplified. It says little about what is known to determine the strength of relationships and what impact feedback has. The literature outlined above suggested that there is evidence that the strength of the impact of media on the individual was shaped by priming and framing effects but, even before that, by the selection of what is covered. Similarly, something is known about how the policy context, in terms of, for example, policy uncertainty and public attentiveness, influences the impact of public opinion on policy. It is also obvious that policy changes will themselves influence public opinion and events.

Figure two below adds in these feedback loops and some notable markers of what influences the strength of relationships.

Figure 7.2: Expanded model, full version



The above, somewhat more complex, figure is intended to be suggestive of the dynamic nature of the interactions. As one aspect is influenced by another, so a feedback loop is initiated and that factor in the end is influenced by itself. The formulation draws on the literature discussed above and highlights a number of relationships that have been identified. What, however, is clear from the literature is that, while these relationships have been identified, their strengths are unclear and the strengths differ according to a range of contextual factors. There is clearly much more to be done in this field, but the insights do allow the discussions of the individual to be placed in a social context, which, from the

above, is clearly important in determining the impact of their responses to the needs of distant others.

7.6 Explanatory power

From the very beginning of this discussion, it has been clear that individuals will respond to the needs of distant others for a range of reasons. They may, for example, feel good about helping and about what it tells them about themselves. It has, however, been argued that a major motivation in helping is the degree to which the individual feels they ought to help, possibly in terms of the degree to which they feel some responsibility to these other people. What has also been argued is that this feeling of 'ought' is not immune from self-manipulation and, if individuals can convince themselves that they do not have a responsibility, it reduces helping.

Combining the arguments made in relation to individual responses with the understanding of the social context that can be drawn from the literature and is summarised above, allows for the examination of the possibilities of explaining responses to real-world needs.

As has been discussed above, the literature makes clear that the impact of public opinion on policy is only part of the story; States certainly have latitude to act according to other concerns. Almost all real-world instances of need involve some political and economic interests. States' responses are, therefore, likely to differ, at times substantially, from public opinion as they are also shaped by strategic concerns. Examining state responses, therefore, involves more than examining the impact of public opinion.

Strategic and economic concerns are evaluated by state and other institutions that themselves consist of individuals. It is interesting to ask how institutional decisions can at times be so different from individual responses. A

responsibility construction, within the balance model, may well assist in answering such a question. It could be argued that individuals can act differently from their own values, if they feel that they have a responsibility to do so. Allocated responsibility may have the potential to alter feelings of 'ought' - think of a ship's captain going down with his vessel. Individuals in institutions charged with strategic and economic concerns, may well feel a different action ought to be followed from what they would do if they were selecting preferred actions as an individual, independent of their assigned responsibility. The issue of work-related responsibility is discussed in more detail in the following chapter.

Identifying strategic and economic concerns of real-world cases is in itself a complex task. For many of the major situations involving human suffering, numerous research projects could be undertaken on these issues alone, especially as the concerns are often not openly acknowledged.

Compared to the difficulty of disentangling strategic and economic concerns and their impact on responses, examining differences in charitable responses may be simpler. To some extent this is true but, as outlined above, individual responses are likely to be influenced by elite responses and policy, not to mention the impact of these two on the media. Individuals may at times take their lead from their governments, while others may respond because they feel their governments have not done enough.

The balance model and the understanding of the social context do provide an opportunity for examining real-world situations. What is clear, however, is that any examination would itself be a major project. Such a project would need to be multi-levelled and to investigate individual responses and social contexts relating to the specific problem. In relation to a real-world situation where there is a need for help, a research project would need to consider:

- The differences in individual responses based on different framings of the need.
 - Including the identification of opportunities to reduce felt responsibility.
- The likelihood of media selection of the event and what interests may shape this.
- The common frames of presenting information adopted by the media.
- The strategic and economic interests of governments and strong interest groups in relation to the need.
- The policy context and its level of certainty with regard to the need for help.

These difficulties aside, it is still worth at least discussing how some of the theory can be useful for looking at real-world events. Explanations can at least be hypothesised, even if testing them would be a separate project.

This thesis was originally motivated by a desire to understand the international response to HIV and AIDS. Despite the enormity of the human suffering associated with the HIV and AIDS epidemic, the international community appeared to be very slow in responding on anything like the scale one might reasonably have expected. Busby's analysis of the donor response notes the disjunction between the scale (tens of millions of deaths) and the donor response up until 2001, which had been relatively small in scale (2006). Since 2001, the Global Fund to Fight AIDS, TB and Malaria and The President's Emergency Plan for AIDS Relief (PEPFAR), as well as increases in bilateral funding, have led to increased international support. Why did it take close to 20 years from the onset of the epidemic for a large-scale response to materialise?

There are probably myriad reasons for the change. Among these would be the development of more effective treatments and the eventual fall in the cost of providing them. Examining the problem, with the expanded balance model as a

framework, may also generate some hypotheses relating to the initial low level of support and the subsequent change.

The response of international players to the epidemic was initially concerned with their own epidemics. One could imagine that the interest in the possibility of a growing epidemic within the country was considered far more newsworthy, and policy relevant, than the epidemic in Africa. The issue of domestic risk was most likely more important for policy makers, the media and the domestic public, leading to low responsiveness to international need.

The above argument may account for some of the delay but, as mentioned, even by 2001, the response was still at a relatively low level. By that time, the epidemics in the wealthy world were seemingly under control and treatment was widely available. What could explain the continued delay?

Firstly, for events to be considered by the public they have to make the news. HIV does not result in dramatic public deaths, such as would be associated with wars or natural disasters. It is a slow-moving, unfolding crisis not a catastrophic event; AIDS deaths may not have the entertainment value of an earthquake. As a result, AIDS did not get the press coverage that it would have had if the same numbers of people had been dying as a result of war or a natural disaster.

The model could be used to suggest more sinister hypotheses in regard to the media. Governments may have been well aware of the difficulty of responding and the costs involved; there were also the interests of the drug manufacturers to consider. It is plausible that international media could have been influenced by these interests. This may have reduced coverage or promoted framings that suggest individual fault.

A less sinister hypothesis of media elite interaction relates to the degree of policy certainty. Until relatively recently, the widely-held view was that treatment was neither affordable nor feasible in developing countries. As regimens became simpler and possibilities around generic, or reduced price, drugs started to be discussed, agreement on policy weakened. The elite in donor governments and the international health community started to divide on the issue of the possibility of delivering treatment in the developing world. With policy uncertainty, and the elite divided, the possibility of the media and public opinion playing a role arguably also increased. Further, from a news selection perspective, a story of people dying because they can't get treatment may be a better story than people dying and there not being much that can be done about it. The treatment story provides a different frame: individuals die, not because of their own doing but because of international drug companies protecting their profits. This situation changed what information was received by individuals, how it was framed, and the potential for public opinion to have an impact.

This change in information, and its framing, can be examined in more detail at the level of the individual, who is at the centre of the model, by exploring the situation through a responsibility formulation. HIV is sexually transmitted and it is easy to blame individuals for their situation. If it were simply the well-being of distant others that was the concern, then this would not matter. Particularly in the early stages of the epidemic, the disease was associated with commercial sex work, promiscuous migrant labourers and homosexuals, all of which were frames that the media arguably supported. It is easy to imagine how these characteristics could be used at the individual level to excuse inaction, along the lines that they deserve it for their 'immoral behaviour'. Applying the balance model would suggest that simply not wanting to help would not be enough, that individuals need excuses to diffuse the pressure to help and "AIDS", as constructed at that stage both by the media and epidemiologically, readily provided them with such. What's more a particular frame could link in with other views, for example presenting prevalence data by race rather than socio-

economic group could be used to feed stereotypes (Barnett and Whiteside, 2006).

Just as with poverty, there are individual and systemic explanations. Consider an example of sex for cash. The same situation could be described as a commercial sex worker and her gold miner client, or as a poor discriminated against woman accepting cash for sex from a man separated from his family by an oppressive migrant labour system. The selection of media portrayals and the influence of an elite, themselves looking to avoid pressure to respond to systemic inequities, may well have played a role in what those distant from the problem were led, and chose, to believe. As the frame changed to one of drug companies being the cause of deaths, so the individual response was also likely to change. The potential beneficiaries became victims of corporate greed rather than of their own moral failings. The model would suggest that, on receiving information in the corporate greed frame, as opposed to those which stressed choices relating to sexual behaviour, individuals would be more likely to support efforts to help.

There is some suggestion that policy activists and charity fund-raisers are well aware of the importance of framing in the HIV and AIDS debate, particularly in regard to portrayals of fault. Activists played a role in shifting the frame to one of corporate greed. Their efforts may, however, stretch further, as there appears to be a growing emphasis in the HIV and AIDS field on the high percentage of women infected, with a particular emphasis on married monogamous women. This suggests an effort to break down a perception of fault.

Even a casual consideration of the changing response suggests the importance of framings and assessments of fault, changing policy context and the role of interest groups. PEPFAR, for example, does not fund needle exchange or the promotion of condoms; nor does it work with commercial sex workers. If the goal of PEPFAR was solely the prevention and mitigation of the epidemic, these

exclusions would be nonsensical. It is as if they implicitly assume that having a role in prevention for people who put themselves at risk is not their responsibility.

Much of the increase in funding for HIV and AIDS has been directed at treatment. Treatment is a high-profile response and may, for the reasons mentioned, also be favoured by the media and possibly by certain interest groups that do not wish to get involved in the issue of sex. Moreover, the numbers of people on treatment can be counted and responsibilities considered met. Treatment is also good, as it diverts attention from prevention and the systemic problems that have created the epidemic, which would take far greater effort to address. There may be other reasons, aside from saving lives, that have influenced the policy shift. In that case, considering the situation through the extended balance model framework would suggest that there may be need for some concern. As pledges are made, attention may well fade; implementation may not be so closely monitored and, as a result, may well deviate from what was promised. What is more worrying is that treatment provision may be presented as 'responsibility met', leaving prevention efforts and large-scale social change unsupported.

It took millions of deaths and the tireless efforts of activists to change the scale of the response to AIDS. Why then such an outpouring of support for the victims of the tsunami?

Again, this is all hypothetical, but the application of the model does suggest some possible explanations. The tsunami was dramatic, unexpected and occurred in a region where many from the wealthy world holiday. It therefore generated huge media attention and, as a result of self-reference in the media, the attention spiralled upward. The visual nature of the devastation also made it hard to deny the scale of the event. Drawn to watch out of curiosity and amazement, viewers may have been unable to deny the call that such images

made on them to respond. Further, responding to a tsunami can be conceived of as a once-off event. Responding to AIDS admits a responsibility and ties responders into future calls. A once-off act of charity may easily restore an individual to balance. When, however, that act is associated with admitting responsibility, which may be hard to disassociate from the need for an ongoing commitment, the implications for future help are far greater and the individual may want to avoid these.

The above examples are obviously oversimplified and there are likely to be a whole range of explanations for responses and the differences between them: in time, surely more research will be done to explain such variations. The point here is that there does seem to be a possibility for a responsibility formulation and the balance model, expanded to consider context, to explain aspects of these responses and the differences between them.

To close this section there is one more example worthy of mention. The Rwandan genocide was one of the most horrific events of recent times. Close to a million people were killed in the space of a few months. Media reports and comments from powerful countries appeared to resist the term “genocide” because of the obligations that came with it. The focus on a civil war was more acceptable. Could it be that Rwandans, living in the heart of the “Dark Continent”, are somehow so foreign to Northerners that they did not warrant the same empathetic response as others might have? Or was it a case of a particular framing? The long history of negative images of Africa, war-torn and corrupt, may have provided a frame to cast the events as nothing out of the ordinary, not needing special attention. The possibility of the media being in collusion with the elite, or at least being manipulated by them to avoid public outcry, appears very real.

7.7 Conclusions

Every event of human suffering is embedded in a complex historical, social and economic context. Even natural disasters link in with past inequities. One need look no further than New Orleans. To comment on the responses of those not present at the site of the suffering, requires investigation specific to the problem. The aim here has been to develop a theoretical framework of the individual and then to expand that framework to consider the individual within a social context. It is hoped that this development will shed some light on the issues and assist in case studies. Human nature is an old topic of study and one far from complete. This is intended only as a small contribution.

It was mentioned in the introduction that the discussions in this chapter could be replicated with the use of a maintenance of self-image model. The model and the conclusions would be as they were with the balance model. This chapter has, however, focused on the context and the explanatory/predictive capabilities, but not the implications. There are surely different implications from conclusions based on individual balancing, as opposed to individual maximising. These have already been alluded to frequently and examined to some extent in Chapter 2, but will be discussed more fully in the following chapter.

Chapter 8: Extensions and implications

8.1 Introduction

This chapter aims to stretch the discussions and will raise more questions than it provides answers. Thus far, the discussions and arguments of this work have flowed between theoretical models of human behaviour and the practical issue of how their application to the question of helping distant others compares with what is observed. Given that the motivation for this thesis was the explanation of outcomes, the theoretical discussion has always been guided by these specific practical goals. This practically-guided, theoretical discussion has led to a focus on two alternative types of model: the balance model and maintenance of self-image models.

It has been argued in previous chapters that the predictions of the two model types are very similar, if not the same; but, while they are the same as each other, it was also argued that they were different from other models. This situation leads to the possibility of extending the discussions into two interesting areas characterised by the following questions:

- Are there areas of behaviour, other than the helping of distant others, in which the application of either of these two models would lead to different conclusions than would be reached from applying more conventional models?
- Why are these models so similar in their predictions and, given this similarity, how do they differ in their implications?

This chapter discusses both of these questions in the above order. The examination of the first question offers conceptual discussions of how these types of models might be applied to different areas of behaviour and what hypotheses or explanations might come from such an application. There is no

intention here to go in search of evidence to test these hypotheses, or to locate them within the literature; rather, they are simply suggestions of possibly interesting avenues of research in other areas, given the theory developed here. Possible applications will be considered in the areas of more general constructions of responsibility and the need to defend them, the integration of social norms, including violence and sexual behaviour, and payment for responsibility.

The second question is theoretical and, to some extent, philosophical. It was touched on in Chapter 2 and is essentially a discussion of the very old question – which will not be resolved here – of what constitutes self-interested behaviour. There is no proof to offer, only argument and intuition. It is, however, an important debate as, depending on which side of the argument you fall, there are serious implications for your interpretation of welfare economics and your understanding of human nature more generally.

8.2 Constructions of responsibility

This is simply a direct extension to other aspects of well-being of the arguments presented in previous chapters with regard to health. If there is motive to avoid challenges to your self-image that would require costly actions to correct, or if there were motive to avoid the triggering of feelings of responsibility, then constructed understandings of the contexts we live in, which keep these in check, need to be defended.

Assuming an individual has developed an understanding of the world around them that keeps feelings of responsibility in check, then they would want to protect this understanding. If the situation changes, or new information is happened upon, they would want to limit the exposure to responsibility that this change may prompt, especially if admitting responsibility in one regard could

threaten the stability of the construction that defends them from other responsibilities. An example may help.

Let us return to issues surrounding HIV and AIDS in the high prevalence Eastern and Southern African context, but this time with a specific focus on children. Children are affected by HIV and AIDS in a number of ways. They can become infected at birth, during breastfeeding, or when they become sexually active. They can be affected through living with infected parents, other caregivers or other adults in the household as these people's health deteriorates. The effects of adult illness can come from the psychological difficulties associated with watching a person close to you die, through increased demands on children's time to replace labour lost through adult illness, or to provide care, or because of the economic impact of the illness on the household. They then suffer the loss, and the possibility of moving to a new home, which may not always be as supportive as where they were prior to the loss and may even be discriminatory or, in the worst cases, abusive.

As a result of living through these situations, children face the possibility of a range of physiological, psychological and material impacts. These impacts can include deteriorating health status or even death, depression, decreased access to education, and many other difficulties (Richter et al, 2004). Many of the negative experiences are a result of increased poverty associated with the illness, death, or the placement of children after the death of the caregiver. The impacts occur along a continuum from HIV specific to HIV non-specific. Being infected with HIV is specific to the epidemic, while impacts of poverty are similar for equally poor children regardless of the cause of their poverty. Given the non-specific nature of many of the impacts, there has been considerable work discussing whether orphans are indeed worse off than other children (for examples see Arnab and Serumaga-Zake, 2006; Bishai et al, 2003; Case et al, 2004). Then there are impacts that are semi-specific, such as discrimination and stigma. These may occur as a result of HIV and, when they do, they may

take on a specific character, but they may also occur for other reasons - making them only partially specific.

Against this backdrop, it is orphans who have been the primary focus of efforts to respond and advocacy efforts to promote further response (Richter and Desmond, 2008). Among orphans there is a particular focus on child-headed households and what are called 'skip generation households' where there are only elderly caregivers and children present (Desmond et al 2003; Richter and Desmond, 2008).

It is interesting to consider the response through the lens of the models discussed. Firstly, emotive images (visual or mental) of orphans, children without family, make it difficult to ignore the need and, given the models' assumptions regarding motivation to avoid responsibility, this difficulty in ignoring may be important. Furthermore, it is also difficult to allocate responsibility for meeting that need to the individuals affected, as they are children. This may well be why advocates and respondents chose to focus on orphans in the first instance. This in itself is not a radical suggestion: people saw an obvious need and responded and tried to get others to do the same, but the models point to more than this.

Focusing on orphans, or even more so on narrowly-defined subcategories of orphans, could arguably be a means of limiting exposure – responsibility is admitted, just not too much of it. If it is recognised that other poor children may be facing similar challenges, then the need to respond more widely may become more difficult to deny. But these children have families. It is the family's responsibility to respond, not ours. But what if the family does not have the resources or the capacity to respond? We then need to find a way to help the child, given the failings of their family. This way of thinking can also be seen as limiting exposure, as it leads to efforts to help the child independently of the family; at times these efforts may even be designed to avoid 'leakage' to

the family. The exposure is of responsibility to the child not to the poor in general.

It is interesting to ask why there appears to be a need to cast the family in a negative light. Currently, there is much debate in the international community about the possibility of responding to HIV and AIDS and child poverty by way of cash transfers to the family (DfID, 2005). In my own experience of the international policy debates on these issues, the response of those involved to the suggestion of cash transfers is often that the family will drink away the money or spend it on themselves in some inappropriate way. This line of argument has seemingly led many to suggest the need to provide the support in kind or cash with conditions and monitoring. No doubt not all families (rich or poor) are supportive of their children and the need for mechanisms to promote child protection is crucial, but it would appear to be a very negative view to hold that this is true of large proportions of poor families. Why then does the view appear, at least anecdotally, so pervasive and the resistance to supporting the family so common?

The two models focused on in this thesis provide a possible explanation, again related to limiting exposure. Both models suggest motive to assess situations so as to limit responsibility. If you have developed an understanding of the world around you that successfully limits feelings of responsibility, you would not want to interfere with it. For example, if you admit that most poor families do care for their children as best they can, it becomes more difficult to maintain the view that the poor are poor simply because they are stupid, lazy and immoral. If this latter belief starts to falter and you admit they are poor, at least in part because of an unjust system, then you have to begin to realise that you are, at least in part, richer than them because of the same unjust system. Your position is not simply a result of your abilities and hard work. If, in order to keep feelings of responsibility in check, it is important to you to believe that the poor are poor because of their own doing, while you are not, because of our own

doing, then this would be a dangerous road to go down and it may be best to battle the unravelling of your construction before it even frays at the edges. As discussed in earlier chapters, this battle may well not even be conscious: lying to ourselves is seemingly a well-developed art.

This application of the models is no longer only about distant others. The above example includes researchers and advocates who may well have personal experience of the impacts of HIV and AIDS on children and the contexts in which these impacts occur. Thinking back to the introduction of this thesis, the motive for selecting distant others was so as to be able to differentiate helping that could not be easily explained by relatedness or reciprocity. There is no reason to think that the models only apply to the helping of distant others; they may in fact be highly applicable, alongside other explanations of behaviour, in closer relationships.

When thinking of distant others it may be relatively easy to deny responsibility but, when the need is much closer, such complex constructions may well take on more bizarre forms. In the 1980s, during a period of great unrest in the South African Apartheid state, with racial tensions running high, one of the most popular television programmes on South African television, which had a predominately white audience, was 'The Cosby's', a sitcom about an African-American family. How was this possible? Anecdotally, at least, it could be argued that many white South Africans believed that 'their blacks are different from our blacks.' Admitting the humanity of black people in South Africa would be too much of a strain on the mental construction necessary to live guilt-free in white South Africa.

These examples come back to the argument that no one seeks the truth; they rather seek only to be comfortable with the facts. It is interesting to think about this in relation to the research community. Researchers should, at least at some level, be seeking "truth". But researchers too, have constructed their own

understanding of the world; they have taken up positions that may well shape what they look for and how they interpret their findings. Moreover, it may well influence the degree to which they will resist findings that are contrary to their understanding. It is easy to tear holes in the methods used in research that is not in line with your view, while being forgiving of failings in the methodologies of supportive studies. Perhaps the clearest example, and one which possibly links closely with the maintenance of self-image, is the extent to which many researchers will cling to, and defend, their own results.

Arguing for constructions that make you look better, could be explained within more traditional utility maximisation models. It is the arguing with oneself and the believing of constructions that require the introduction of the maintenance of self-image component, or the crossover to the balance model. In these models, it is what you believe to be true that is valued not truth itself; this would seem to be closer to the truth of the matter.

8.3 Integration of social norms

There is a substantial literature on social norms and their integration into individual decision-making processes; some of this literature was discussed in previous chapters. There follows a sketch of a theoretical framework for understanding the role of social norms in individual decision making, using the self-manipulation arguments present in the maintenance of self-image and balance models. This framework is discussed in relation to two sets of behaviour: violence and sexual risk taking.

Both of the models can be used to examine the question of the role of social norms in determining average and individual attitudes to violence and resultant behaviour. A more traditional model would include decisions relating to violence as simply preference-based; social norms would only be important to the extent that they influenced preferences. This is a very simplistic understanding with

little predictive power; the application of both models suggests a potentially more useful way of looking at the issue.

A maintenance of self-image model can be constructed based on the assumption that individuals wish to maintain a self-image of being not overly or excessively violent, or possibly non-violent. In the balance model, it may be hypothesised that humans have negative feelings towards violence that inhibit them from being excessively violent, or violent at all. As these are both assumed to be mentally-generated images or feelings – subconsciously generated or otherwise – they are subject to manipulation. For example, violence in self-defence does not have the same impact on self-image, or on restraining feelings, as the situation is interpreted differently. This example could be generalised to individuals being constrained by their need for a positive self-image or the need to keep negative internal responses to violence in check from unjustified violence. This is what was discussed in Chapter 4, but how does the argument extend to social norms?

Some societies and social groups are more violent than others, but not all individuals within a particular group are likely to be equally violent. What may be considered excessively violent by one group may not be considered so by another; arguably, therefore, social norms surrounding the use of violence differ, as does the degree of internalisation by members within any one group.

Suppose a social norm specifies a level of what is considered justifiable violence, or perhaps a formula of sorts that considers the context in terms of self-defence, provocation etc and the justified response. This is not to suggest that there is some objective social norm. Rather it suggests that individuals' perception of a norm is not independent of their social group. This raises the issue of what they consider to be their social group, but that is not critical at this point. What is critical is the assumption that, for a social group, norms of

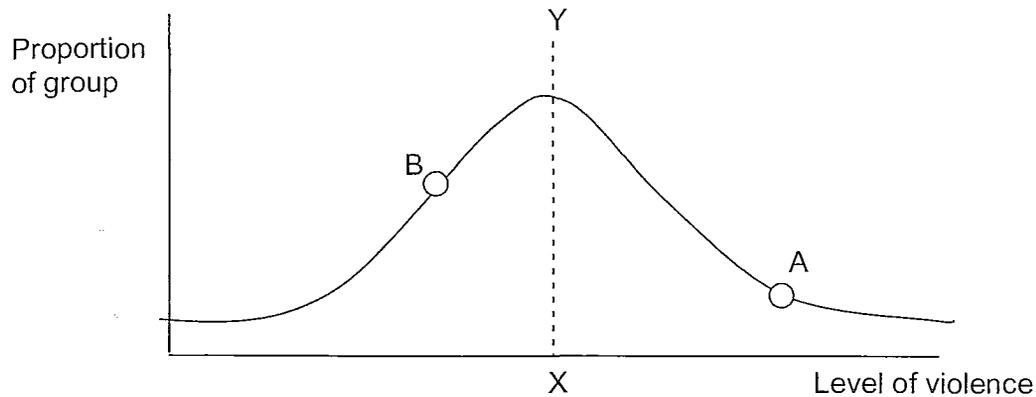
behaviour exist that are associated with that group. The question then becomes: if there is a group norm, how does this affect individual behaviour?

Assume the existence of a norm regarding violence and that individuals internalise this social norm. No one would suggest that they then all adhere to it in the same way; some will be less violent in a given situation and some more. What the models provide is a framework for understanding how this variation might operate.

Once internalised, the norm shapes the strength of restraining feelings or the level of damage to self-image associated with a particular act. Recall, however, that the degree to which these feelings are felt is, to some extent, open to manipulation. If they had reason to, people might well try to reduce negative feelings associated with a particular violent act.

Certain individuals may have more motivation, or more ability, to self-manipulate themselves into believing that their more violent responses are justified, while other individuals may find themselves unable to interpret the norm so loosely, so it constrains them more than others. This would suggest that violent responses would be distributed in some manner around the social norm. The following diagram provides a graphical representation of this hypothesis, if the distribution were assumed for illustrative purposes to be normal.

Figure 8.1: Individual relationship to social norm



In the figure above, the line XY denotes the social acceptability of violence, with most individuals clustered around this norm. Individual A, however, is to the right and, while they set themselves relative to the norm, they are likely to be more violent in a given situation. For individual B the situation is the opposite.

That individual behaviour is distributed around the social norm is far from an original argument. What the models suggest, however, is more than this. They suggest that the distribution is linked to the norm by way of a process of varying degrees of internal control. If we add to this framework, the assumption that the ability to self-manipulate is similarly distributed across populations, then the implications of this argument become clearer. The assumption puts aside the argument that members of some societies may just be better at avoiding the restraining effects of social norms and that is why these societies are more violent.

If the potential to manipulate is similar across populations, then the reasons for one population or society being more violent than another, within this framework, would be attributed to a more violent social norm or to the context providing more fuel and motivation for self-manipulation. If this is the case, it

has important policy implications, but before coming to these a practical example may help.

South Africa is well known around the world for its high level of violent crime. What would be interesting to investigate is the degree to which the general acceptance of violence as a response to a given situation is higher in South Africa than elsewhere. From my own experience, South African society does appear to be more accepting of violent responses than other societies, such as the UK, in which I have lived. This is obviously not a basis for drawing conclusions, but there is an interesting hypothesis to test. Furthermore, South Africa's history of violence, oppression and the promotion of defiance of social institutions may not only lend itself to a more violent norm; it may also provide the basis for self-manipulation. Having suffered violence and oppression, or even just seen it around you, it may become easier to justify your own violence. Combine this with continuing high rates of poverty and inequality and you may be on the way to explaining how criminals justify their violence to themselves, thereby weakening their internal constraint systems, and what motivates them to do this. Anecdotally, many victims of crime say that it was as if the perpetrators, whom they had never met, hated them. Did they hate them or did they need to cultivate that hatred in order to free their actions of constraints, thereby giving themselves the chance to pursue their other motives relating to material acquisition?

In terms of policy implications, this framework would suggest that a violent response to violence would have limited success, particularly in the long run. If the South African state opted to respond to violent crime in the country with a physically aggressive response, this might well shift the norm to a more violent one. The old adage that violence begets violence would hold with these models. Analysis with these models would point to the need to find ways to lower the social acceptability of violence generally and to work on reducing the availability of excuses and motivations. This is not to suggest a small role for

policing. Considering the potential costs and benefits of violent crime would, within this framework, feed into the level of motivation to weaken the constraint. There would not be much motivation to weaken the constraint if criminals believed they would be caught and imprisoned. What this conceptualisation would suggest, with regard to enforcement, is the need for improved investigative policing and response time rather than more 'strong arm' policing.

The central difference in the above framework, when compared to more traditional models, is that the social norm has a greater role than simply shaping preferences. The idea of a social norm is obviously stylised, as no objective norms are observable. But the argument that the general level of acceptability of violence influences the behaviour of those who have motive to be more violent may have potential to inform policy. Moreover, the argument links individual decision making to the context in which it is occurring, again with policy implications. It is the linking to context that motivated the application of the framework to the questions of sexual behaviour.

Colleagues⁹ and I have recently tried to apply this framework of social norm integration to the question of sexual risk behaviour in the context of HIV and AIDS, with an emphasis on young people. The framework is used to link together two approaches to prevention, as two distinct arguments around HIV prevention have developed: one that focuses on the individual and the other on the context. The former focuses on the importance of individual decision making regarding sexual behaviour and seeks to provide the individual with information on risks and how to protect themselves. The latter focuses on how the social and economic context shapes risk and how there is a need to address these structural factors – such as migration patterns and unemployment – if risk is to be reduced.

⁹ Harrison, D and Richter, L. This is ongoing work and will eventually be submitted for publication with the three of us as authors. The application to youth is what is being jointly undertaken while the social norm theory and its application to HIV prevention are entirely independently developed.

Within the social norm framework developed in relation to violence, as discussed above, these two arguments can be conceptually linked; this may well assist in linking them in policy terms. In this instance, the social norm relates to sexual behaviour. Again this norm is internalised but constrains behaviour to varying degrees, depending on the strength of the individual's internal constraints and their level of motivation to weaken them. Just formulating it in this way raises a frightening possibility regarding individuals' relationships with reality. In the context of HIV and AIDS, a norm may relate to acceptable risk. To weaken the hold of this norm, if an individual has motive – such as wanting to enjoy risky behaviour – to do so, one way of doing this would be to deny the risk. This would mean manipulating oneself into a false sense of security. This is different from an argument that the individual simply places a greater value on the present than on an uncertain future outcome¹⁰. This may well be the case, but what this suggests is that individuals wanting to take this course of action will need to convince themselves of a lower probability of a negative outcome. The problem is that only they are convinced, as probabilities of infection, of course, cannot be persuaded.

Leaving the denial of reality to one side for the moment and returning to the linking of the two arguments: if the framework is correct, there are three ways of trying to reduce individual risk.

- Shift the individual's position within the distribution.
- Change the nature of the distribution.
- Shift the entire distribution by shifting the norm to which behaviour is linked.

¹⁰ This argument need not be in place of arguments that they do value the present far more than the future even without self-manipulation. The two possibilities could well work alongside one another. What this suggests then is that it is important to ask what the relative influence of each is, as the responses to one may well need to be different to the responses to the other.

Provision of information to individuals is essentially an attempt to shift their position within the distribution, unless perhaps the numbers involved are sufficiently high to shift the norm. Information may help strengthen the influence of the norm on individual behaviour by removing excuses and misconceptions, and by doing so pull them closer to it. Improving the social context may shift the distribution and may even change its nature. If these changes are positive then the risk of the entire group is reduced.

Improving the social context and shifting the distribution would seem to be highly important in the long term. The fact that it is a long-term response should not, however, be used as an excuse for not pursuing such improvements; it is, however, a reason for not pursuing them alone. Linking the approaches in the above framework may help in thinking through short- and medium-term responses that may reduce risk. Understanding the pathways through which the social context influences the mean and the distribution may help in this regard.

In terms of the issue of young people and sexual risk, the evidence in South Africa suggests a major increase in risk after leaving schooling (Shisana et al, 2005; Pettifor et al, 2004). The majority of South African youth leave school only to enter a period of prolonged unemployment. It is during this period that many of the infections happen. Using the above framework, there are a number of issues that emerge. Has the social norm become a more risky one for this group? Has the control of the social norm become weaker, or has the motive to weaken it increased?

Going back to the question of the denial of reality in terms of risk, could be useful here. If this is what is occurring, then it would be helpful to try to understand what motives there could be for engaging in such self-endangering weakening of a social constraint. A range of possibilities come to mind: it could, for example, be that, for young women, the financial benefits associated with

sexual favours are a major motive; it could be a heightened need to feel good about oneself when one is unemployed and facing a difficult future. These are just suggestions, but they are made because they show the intersection between the individual decision-making process and the social context. Social context prevention should aim to improve the transition between school and work and so shift the distribution, or change its nature. Understanding how this context plays out might point to ways to respond in the short term. For example, efforts aimed at promoting positive peer-group pressure, and programmes to support youth during the transition phase, may help in reducing the weakening adherence to social norms. The framework does not pretend to provide answers, but it possibly offers a useful way of looking at things.

8.4 Payment for responsibility

This topic is useful as it provides a bridge between the two themes addressed in this chapter. The above discussions have raised some issues relating to the possible use of these models in examining other areas of human behaviour, besides the helping of distant others. Another area that they can help to explain is payments for assuming responsibility. While the models again come to the same predictive conclusions, the implications are somewhat different. The nature of this difference is the topic of the following section.

It is common for jobs that involve greater responsibility to be paid better, irrespective of whether these jobs involve the commitment of more time or physical effort. To some extent this can be explained within standard utility models. Accepting more responsibility in some ways increases the uncertainty of employment. The more responsibility a position involves, the more potential there is to fail and to look bad or to lose one's job. Therefore, it would make sense that the job be better paid to compensate for this increased risk. If it was simply a case of costs imposed by others, then individuals would act responsibly in their jobs only to the extent that such actions could be observed

by others and the standard they would adopt would be that of their employers. Going above and beyond the call of duty would make no sense, unless they just enjoyed doing so.

Using the maintenance of a responsible self-image model can help explain a wider range of behaviours. The added responsibility of a job could be assumed to be incorporated as part of the individual's self-image. As a result, protecting their self-image becomes more difficult; knowing that this will be the case, could prompt prospective candidates for a job to require higher pay to accept positions with greater responsibility attached. This would explain responsible behaviour even when not observed and would explain individuals holding themselves to their own standard, not only the standard of others. Going above and beyond the call of duty would be explained by them having interpreted their responsibility as being greater than that of others associated with the position. This, however, is where things start to become difficult.

A similar argument could be made with the balance model: taking on a job with more responsibility results in the response system incorporating prompts that it would previously not have considered at all, or not to the same extent. In this way, the individual assumes the responsibility and knows that feelings of responsibility, which they will have to respond to, will be more frequent and will require them to divert attention from meeting other needs and they require compensation for this possibility.

As always, the same predictive conclusions: individuals will need payment or other reward to assume responsibility; they will internalise the responsibility and hold themselves to this new standard, based on their interpretation of events. This interpretation of events is subject to manipulation to avoid responsibility being felt.

What are arguably different are the implications. In the maintenance of a responsible self-image model, every action done to maintain the positive self-image is still done for private benefit. That is, the actions are self-interested. Individuals undertake the action because, if they don't, the damage to their self-image will be greater than the costs associated with their selected action.

In the balance model, the motive need not be self-interested, only the response. The question is whether an action that is motivated for others, and responded to for oneself, is self-interested or not. This is the subject of the next section. For many everyday situations the distinction may seem trivial, but there are times when the difference would seem far more important. A fire officer is paid for their added responsibility (arguably not enough but that is another issue). When faced with a high-risk situation some officers will go 'beyond the call of duty', and risk, possibly even lose, their lives to save others. It would seem strange that they do this only to avoid feeling bad about themselves.

8.5 Different implications

Throughout this thesis, the balance model and the self-image models have been discussed as two different types of model. Their similarity has been in their predictions, but it has been suggested that they are fundamentally different and that the implications differ as a result. Before moving on to examine the differences in implications, it is worth examining the root of the similarity.

The maintenance of self-image models are based on the argument that individuals want to generate and maintain a positive self-image and that they generate utility from doing so. This self-image may, for example, be of a fair person, or a responsible person, or a person who is not inappropriately violent.

The balance model, on the other hand, suggests that individuals have a range of feelings that put pressure on individuals to act. These feelings are not

substitutable and some examples of these may relate to feelings of responsibility or feelings against violence or unfair actions.

The similarity in the models comes, to a large extent, from them both being concerned with internal states, not external consequences. External consequences may influence the internal state, but it is the internal state that guides actions. The two models could be merged to some extent. It could be argued that the need for a positive self-image is what prompts the feelings in the balance model.

There are, however, three fundamental differences that cannot be merged away:

- In the maintenance of self-image models the self-image is an end in itself.
- In the maintenance of self-image model the utility from maintaining a positive self-image is of the same type as is gained from any other good or outcome.
- Only in the balance model is there a difference between a 'want' and a 'have'.

In terms of the first difference, the positive self-image is the end that is sought in the maintenance of self-image models, whereas, in the balance model, it could be a pathway through which certain feelings may be prompted, but the end goal is balance. Linked to this, in the balance model, the need for a positive self-image – as being, for example, responsible – may be a reason for certain actions prompting a feeling of responsibility, but it may not be the only one.

The second difference arises because the positive self-image in the maintenance of self-image models is substitutable with other utility-generating outcomes. If a sufficiently valuable option that involves a serious loss of self-image is offered, it will be taken. The individual will try to reduce the cost by

self-manipulation, but in terms of the net effect they will be better off having maximised their utility. In the balance model, feelings can be traded but they are not substitutable. That is, an individual may accept some level of negative feelings in order to gain some other good, but this leaves them feeling good in some other way, while still feeling bad in terms of the negative feeling they accepted.

The difference here is subtle. In the self-image models, benefits from any good can cancel out costs of any bad. The greater the pay-off from an alternative, the greater the loss in self-image an individual would be willing to accept in order to gain such an alternative. Put another way: you would sell your own mother if the price were right. As mentioned in Chapter 4, some heavy assumptions would have to be made for this not to be the implication.

In the balance model, the greater the pay-off linked with a negative feeling, the greater the motivation to manipulate yourself into believing that the level of feeling is inappropriate. You can sell your own mother, but only if you can convince yourself that it is not such a bad thing to do: remembering that she dropped you on your head as a child would help here. The greater the amount offered for your mother, the more motivation to think of ways to manipulate yourself, but if you can't sufficiently reduce the negative feeling then you can't sell, no matter the price.

In the maintenance of self-image models, the self-manipulation reduces costs. In the balance model, there are times when self-manipulation is necessary if the action is to be seen as an alternative. This brings us to the third difference, which is closely related to the second. In the balance model, as mentioned in chapters 2 and 4, 'have to' and 'want to', as well as 'can't' and 'don't want to', are possible, whereas in the self-image model only 'wants' and 'don't wants' make sense. In the maintenance of self-image models, it is possible to have very strong 'wants', but they are still just 'wants'. Of the two model types, only

the balance model, with its added assumptions relating to feelings at their extreme, gives meaning to 'have to' and 'can't'.

The three differences described are only important in so far as they lead to different implications, as it has already been shown that they do not lead to differences in terms of predictions. More than that, they are only important if there is some suggestion that the balance model may be more appropriate than the maintenance of self-image models.

Borrowing from Hauser (2007), who has been mentioned in previous chapters, could be useful here in examining the appropriateness question. Hauser describes three moral creatures: the Kantian, the Humean and the Rawlsian, characterising the arguments of Kant, Hume and Rawls respectively. The Kantian is a rational creature. She evaluates the alternatives and thinks through the consequences and draws judgements about what is right. The Humean is an emotional creature and is guided by her emotions as to what is right or wrong. The Rawlsian is guided by her emotions but these emotions are first generated by the application of principles, often at a subconscious level.

Hauser's thesis is that we all possess a universal moral grammar and, while this is shaped by context, it sets certain boundaries. If the moral grammar argument is correct then we are all Rawlsian. It would be difficult to see how innate guidelines shaping responses would work if we were Kantian and consciously thought through all moral decisions. It would likewise be difficult for a Humean, as the application of grammar requires a mental process not just an emotional response.

As part of his argument in favour of the Rawlsian conception, and the important point for this discussion, Hauser argues that morality is not entirely linked to rational conscience calculus. This is a major deviation from previous arguments and so requires Hauser to take issue with authors, such as Kohlberg, who

argue for stages of moral development matched to children's ability to rationalise. Hauser argues that we have a moral system that develops over time with interaction with the context, but begins with certain universals. As a result, children's sense of morality is, to some extent, innate and not totally linked to children's ability to think through the consequences of their actions.

The maintenance of self-image model is essentially a form of Kantian creature, not in the sense that they are looking for universal principles, but in the sense that the decisions are linked to rational conscious assessments. The balance model is closer to the Rawlsian in that conscious or unconscious processes have a role, but the decisions are then guided by the feelings that result. The balance model is not Humean; such creatures do not allow space for self-manipulation, as the feelings are the guide and are not linked to cognitive processes.

Hauser presents a wealth of data on early childhood development, which suggests that children have a sense of certain moral principles long before they are able to argue through them. He also points out that, even in adulthood, many moral decisions are made with almost gut reaction and, if asked, the decision makers often struggle to explain their choice. Both these sets of evidence suggest that a purely conscious model of rational decision making is difficult to defend. If individuals struggle to explain their decisions, it is difficult to argue that they carefully weighed up the costs to their self-image of alternatives and then decided. As discussed in Chapter 4, Hauser also points out that psychopaths are often well aware of what is considered right or wrong; they may even have the same gut reactions. The difference is that their awareness and gut reactions don't prompt the feelings that guide behaviour. In the self-image model, psychopaths simply have different preferences regarding self-image; in the balance model, something is considered abnormal about their decision-making process.

So there is evidence that humans may have a system of morality that develops with time and that this system involves a combination of conscious and unconscious processes, which prompt feelings that guide behaviour. This sounds awfully like the balance model and points to another important difference between this and the maintenance of self-image models, as well as other models. The balance model need not be based on the purely conscious assumptions of rational calculation models.

The balance model also provides a framework for grammar-type arguments such as Hauser's. The existence of grammars could be conceptualised as specifications within which aspects of the human system are developed. Grammar arguments could probably be made for a number of different aspects of the system. To some extent, the biological explanations of how physiological systems operate are very similar to the functioning of the moral system suggested by Hauser. The balance model then provides an argument for how individuals manage across aspects, with the drive for balance taking centre stage.

The above discussion suggests then that the balance model is different and that there is evidence to suggest that it may be more appropriate. The next question is: so what? Does the application of the balance model have different implications? One could argue that the balance model is simply a more complicated representation of utility maximisation. This argument would be that, in the balance model, utility comes from balance and that in the end it is still utility maximisation. This, however, misses an important point: that equal balance may not always mean equal individual welfare. To show this, it is helpful to consider in more detail the difference between motivation and response.

Consider a simple decision-making process with two parts. The individual feels a motivation to act and then responds to it. In the interest of discussing self-

interested behaviour, consider two types of motive and response: A motive for the benefit of others and a motive for the benefit of self, and a response for others and a response for self. The four combinations are outlined in the following table.

Table 8.1: Combinations of motivation and response

Motivation	Response
Self	Self
Other	Other
Self	Other
Other	Self

Responses to internal psychological pressures can only be for self. The motive may conceivably be for others, but the response is simply an internal process. If this is accepted, then two rows of the above table fall away, leaving the following.

Table 8.2: Possible combinations of motivation and response

Motivation	Response
Self	Self
Other	Self

The fact that a response to an internal pressure must, at some level, be for the individual is used by some to conclude that all behaviour is self-interested. This line of reasoning, however, ignores the possibility of other-regarding motives. Recall the discussion on commitment in Chapter 2, where attempts were made to discredit it by arguing that, once another's goal is taken on, it becomes your own goal. What the balance model suggests is that the goal remains for the benefit of others, but it links into the individual's system, prompting feelings that are responded to for the benefit of self.

The key question in determining the difference in implications between the balance model and the self-image model is: is there a difference between decisions stemming from motivation for self-benefit and response for self-benefit (self-self) when compared to decisions based on motivation for the benefit of others and response for self-benefit (other-self)? The self-image model, and all traditional utility maximisation models, consider only self-self, or consider other-self as essentially the same as self-self.

It could be argued, for example, that responding to motives characterised by 'ought' is self-interested in the response but not necessarily in the motive and that they represent a different type of action than self-self actions. It is only with this distinction that actions of self-sacrifice can be seen as categorically different from purely self-interested actions. Sen's (1985) concept of commitment is essentially an example of other-self decision making. He sees it as different and the implications flow from this difference. He argues that welfare economics is based on the assumption that all actions are self-interested and that individuals will maximise their utility; if this assumption is removed, a major reevaluation is necessary.

A strong argument could be made that other-self decisions are indeed different from self-self decisions, but a more difficult question is: are they a different category of self-interest or can they not be considered as self-interest at all? If they are not self-interested, it suggests that there are situations when individuals could move to a position of greater welfare for themselves but do not do so because of concern for others. Some would argue that, if the outcome is what they choose, this must be the best outcome for them, but this argument is based on a focus on the response, not the motive.

The argument that the other-self decisions are not self-interested appears a little paradoxical: the individual is worse off for having considered their motives to benefit others, but better off than they would have been if they didn't heed

them. Having other-regarding motives does not change balance being the goal, but to achieve balance requires a consideration of others.

In the end it comes down to the point of judgement. Judged by motives, individuals are arguably not always self-interested. Judged by the decision process then selection of outcomes is always self-interested.

Before taking this argument back to the balance model, it is worth noting the possibilities for weakening some of the starting assumptions. Motives were discussed above as either for the benefit of self or the benefit of others, leading to two categories of decision: self-self and other-self. These, however, need not be seen as two categories but rather as the ends of a continuum, as depicted in the figure below.

Figure 8.2: Continuum of choice



Motives could well be assumed to have elements of benefit for self and others and may not necessarily be at the extremes.

Returning now to the balance model and how these discussions play out, the balance model hypothesises a range of pressures that create imbalance, which the individual tries to address. Some of these pressures are purely self-interested and, many would argue, only temporarily satiable (Nettle, 2005), which may be why they are considered so often. Some of these pressures are mixed and some may be for the benefit of others. The greater the pressures related to actions which will benefit others, the less attention and resources available to respond to the pressures related to self-benefit. Consider two

scenarios, A and B, and assume that, in both situations, there is no external pressure to act in a particular way. In scenario A, an individual comes across no information that prompts any pressures relating to others and they devote all their attention and resources to responding to purely self-interested motives. In scenario B, the individual comes across information that prompts other-regarding motives. To achieve balance, the individual gives some attention to these motives, leaving less for the purely self-interest motives.

If other-self decisions are a subcategory of self-interest, then the individual's welfare in scenario B is greater than in scenario A: the individual shifts resources away from purely self-interested motives to meet the needs of other-regarding motives, because responding to them relieves greater pressure, resulting in more balance. If other-self motives are not self-interested, then the individual chooses a lower welfare outcome in scenario B compared to A, even though the welfare outcome of A is available in scenario B. They do this because they seek balance, not welfare maximisation, but this is only true if you believe other-self decisions are different from self-self decisions. If they are not different, the balance model collapses to a rather involved utility maximisation framework.

It was made clear in the introduction to this chapter that this question of self-interest would not be resolved here. What the balance model offers is a framework for examining actions for those who think that other-self decisions are categorically different. For those who hold that they are simply a subcategory, the model adds very little to the maintenance of self-image models.

8.6 Conclusions

This chapter aimed more to raise questions and issues than to draw conclusions. It has highlighted a number of possible areas of study to which these types of approaches might be extended. The first half of the chapter was intended to outline a number of hypotheses in different fields and link these to policy implications. The policy implications rest on the validity of the hypotheses, which remain to be tested.

The second half of this chapter has endeavoured to expand on the discussions in Chapter 2, relating to the differences between the balance model and the self-image models. Some differences were noted and some arguments in favour of the balance model were provided. The predictions of the two approaches are very similar, but the balance model provides space for non-self-interested action. It is the possibility of non-self-interested action that alters the implications. Arguments can be made, but the question as to what extent humans are motivated only by self-interest and what constitutes self-interest will be debated for some time to come yet.

Chapter 9: Summary and final question

9.1 Introduction

With all of this discussion of self-interested behaviour, I should be clear that this thesis was a self-interested exercise. I wished to understand why so few people have felt the urgency of the situation that results when HIV and AIDS combine with poverty. I wanted to understand how it is that millions of people could die, while the world remained so quiet. Was it that they just did not matter enough to those with the power to do something? Or was it that this call on the world's humanity was simply too easy to ignore?

This final chapter seeks to summarise the process followed in trying to develop an understanding of what the reasons were. The argument I developed to do this is similarly summarised here, as are its implications. Possibilities in terms of further research and potential implications for policy in different areas have been identified during the course of this thesis. While these outcomes were, I hope, useful, they were not the aim. The aim was simply to try to understand.

We all have our constructed understandings of the world and our place in it. Arguably these are necessary, as otherwise we would forever be pulled one way and then the other. We need some foundations. There are, however, times when it might be a good idea to examine the validity of some of these foundations. This thesis concludes by asking what the implications would be, in regard to our felt responsibility to others in need, if we did not seek only to be comfortable with the facts, but admitted the truth.

9.2 The process

This thesis has attempted to outline a process of identification, development, attrition and application. The need to understand the responses of individuals

to the needs of others was identified, as were existing efforts to explain these responses. Deficiencies in these arguments were noted and an additional model and formulation were developed. The balance model provided an alternative explanation of behaviour and the responsibility formulation offered a way of formulating alternative models to address the issue at hand. This identification and development, however, resulted in too many models being available to proceed to their practical application. A process of attrition was then undertaken by examining these alternatives against evidence and theory available in the psychology literature. This process resulted in the identification of two models: the balance model and a maintenance of self-image model, both of which fit well in a responsibility formulation. These models could then be applied to consider their implications at the individual level, by way of experiments, and at the societal level by way of expanding them to consider social context.

The identification of the problem was initially centred simply on when and why people will help others. Decisions to help others can result from a complex interplay of motives. The shift to a focus on distant others was so as to be able to remove some possible motives from the discussion. Arguments relating to reciprocity, for example, are not easily applicable when distance is involved. Similarly removed as potential reasons are other motives for helping, where the help is shaped by concerns not directly related to it, such as might be argued to occur in local policy making. When determining access to health services, for example, the policy maker may have in mind political motives in terms of gaining support, and not simply the provision of help without material benefit.

The focus on distant others means that explanations of helping behaviour, if they are to be useful, must be able to explain help when there is no potential material reward. Without a material reward, the models of behaviour quickly become concerned with psychological benefits.

A review of the economics literature highlighted a number of different models. These models were placed into four groups.

- Incorporation models: Individuals help because others' welfare is incorporated in their own welfare function.
- Private benefit of action models: Help is provided because the action makes individuals feel good.
- Private benefit of maintenance models: Individuals help in order to maintain a belief about themselves.
- Multiple-self models: Individuals have motives to help that are distinct from other motives.

In the process of development, two additions were made. Firstly, an alternative model was suggested: the balance model. The balance model argues the existence of different types of non-substitutable motives, similar to the multiple-self models. The model also contains an assumption relating to the pursuit of balance. This assumption provides a basis for decision making that is lacking in the multiple-self models. There will be more on the argument of the balance model in the next section.

The second addition was the responsibility formulation. Many of the models are very broad and so difficult to apply. Individuals get a private benefit from an action, they want to maintain some image of themselves etc. What selecting a formulation allows is the narrowing of these models so that they may be more easily applied. The consideration of responsibility allows for such a narrowing. Individuals value acting responsibly; they value maintaining a responsible self-image etc. The responsibility formulation was introduced, as it was argued to provide a useful means of narrowing the models when explaining the helping of distant others.

With these two additions, and the four groups of models, there were too many options to progress directly to applying the theory to the practical questions. Following on from this process of development, all the options were taken through a process of attrition. Re-examining the data and arguments present in the psychology literature, with a view to selecting between the economics models, proved to be a useful exercise. A number of problems with some of the models and the strengths of others were identified. The exercise was not to identify one model, as there are different reasons for helping in different situations. The exercise was to identify the most useful theory to apply to the practical questions.

The incorporation models fared poorly; it becomes difficult to explain when the welfare of others is included and when it is not. For the model to be able to fit with the evidence, others' welfare has not only to hop in and out of the decision maker's welfare function but to slide in and out to varying extents. If incorporation models are to be useful, then they require a theory of when others' welfare is included and to what extent. Introducing this theory essentially changes the nature of the model into one that falls into one of the other groups.

The private benefit of action models fared better than the incorporation models. They do easily allow for explanations of why help will be affected by context. Taken within the responsibility formulation, for example, individuals would be argued to benefit from helping actions when they feel it is their responsibility to help. While potentially useful in some situations, the model implies that individuals would want to know when it was their responsibility to help. They would have a motive to infer responsibility and to seek it out, as they gain benefit from doing so. This attitude towards information does not appear to be present and suggests a problem with this type of model.

The private benefit of maintenance models do not only easily accommodate the responsibility formulation; they also deal well with both the problem of differential helping depending on situation, and with the possibility of information avoidance. It has been argued that this is the fork with which mainstream economics can eat sushi. It is, however, still a fork, as it is still an approach based only on self-interested motives.

Multiple-self models do not only consider self-interest. Most of the models in this group, however, do not provide a reasonable theory of how motives are weighed together when making decisions. A hard constraint model does provide a decision theory and can accommodate situational differences in helping and avoidance of information. The problem with such a model, however, is that it does not explain helping behaviour within a constraint. It only explains helping when helping is the only option the individual is not constrained from.

The balance model is more flexible than the hard constraint models. It deals with why helping would be shaped by the situation and why individuals may wish to avoid information or its implications. Its predictions are very similar to the maintenance of self-image models, but it provides for the possibility of non-self-interested behaviour.

The first phase of the attrition process highlighted the strengths of the maintenance of a responsible self-image model and the balance model. These two models were then discussed in relation to the literature on harming. This process did not result in a clear 'winner', but it did highlight the conceptual differences between the models. In the balance model, an individual cannot harm unless they can justify it to themselves. In the maintenance of a responsible self-image model, they cannot harm unless the pay off is high enough to compensate for the loss to self-image; they may try to justify their actions to increase the net pay-off, but this is not necessary.

Once the process of attrition had reduced the number of possibilities, it became possible to apply the theory to the practical questions. The application primarily involved the balance model. This shaped the approach used but it is noted that the predictions with the maintenance of a responsible self-image model would have been very similar. The application was initially done at the individual level. While social contexts are obviously important, the discussions have dealt primarily with the individual, so first examining individual helping behaviour was appropriate. The consideration of individual responses highlighted the possibility of using the balance model to examine helping behaviour.

The experiment lent weight to the balance model argument and provided input into how motivated assessments of situations play out and how perceptions of responsibility are formed. The sample did not allow for generalisations to be made, but did suggest the importance of similarity, suffering and perceived fault, among other things, in influencing felt pressure to help. The result also supported the argument that this felt pressure to help is distinct from sadness, the relief of which might be considered self-interested. Finally, it provided additional evidence of the importance of context and the individual's response to information in determining decisions relating to providing help.

Expanding from the individual level was made possible by drawing on insights from the literature on media, public opinion formation and the opinion-policy link. Using this literature, it becomes possible to consider the balance model in a social context. The importance of media selection and framing in determining what information reaches individuals, and in what form, becomes obvious. The importance of the policy context, and the possibilities for public opinion to influence policy, were also noted as central in determining what impact individual responses to others' need will have.

Once this framework was developed, it was possible to return to the initial motivation for this thesis: explaining the slow response to HIV and AIDS. Considering the global response to HIV and AIDS, or any crisis, is clearly a complex task. The expanded balance model did, however, appear to provide a number of strong hypotheses. The model suggests reasons for the slow response and the change in recent times, and allows for predictions to be made with regard to likely changes in the response.

Considerable effort has been made within the thesis to stress the possibility for the predictions of the thesis to be based on the maintenance of a responsible self-image model, or the balance model. The former model would suggest motivated reading of situations and the possible resistance or avoidance of unwanted information. As such, the explanations offered with regard to the delayed response to HIV and AIDS could be repeated with this model.

While it is possible to replicate the predictions of the balance model with the maintenance of a responsible self-image model, the two are different. They differ in terms of their characterisation of actions and their implications for welfare economics and research design. The maintenance of a responsible self-image model was included so that those who had difficulty accepting an alternative formulation that seeks to highlight the importance of balance, the possibility for non-self-interested action and the importance of motivations relating to others, would not reject the predictive conclusions.

To conclude, however, it might be best to repeat the underlying balance model argument without always pointing to how other models, in particular the maintenance of a responsible self-image model, can produce the same results.

9.3 The argument

In the end, with all the disclaimers and alternative explanations aside, the argument is simple. We all experience a range of feelings which pressure us to act to address them. The more we feel any one of them, the more attention that one demands. Each feeling is to some extent distinct and you cannot reduce the pressure of one feeling completely by seeking to address others. Comfort eating may help, but the potential for the relief of hunger to reduce other negative feelings or generate positive ones is limited. What is sought is balance that, given the non-substitutability, requires attention to be paid to whichever emotions are prompted. Individuals make choices so as to keep the negative feelings down, the positive feelings up, and to maintain balance.

The increasing attention, non-substitutability and need for balance are what gives meaning to the words 'have' and 'can't' as being different from 'want' and 'don't want'. As feelings approach their extremes, it becomes more and more difficult to ignore them, and more and more beneficial, in terms of balance, to address them. Choices that will not assist in addressing the particular feeling are effectively ignored, as attention is only given to potential relief of the feeling which is approaching its extreme. It comes to a point where you feel you just have to respond to restore balance, as you can't think of anything else. As the feelings are not perfect substitutes, you have to respond to the feeling that is approaching its extremes.

While distinct and not completely substitutable, these feelings can be grouped. For example, many of these feelings relate to private 'wants' and 'needs' and balancing between them within this group of 'wants' is essentially maximising your welfare.

It does, however, seem that there is more than personal welfare maximisation going on during the individual's decision-making process. Some of these

feelings that pressure your actions are of a distinctive character and differ from the group characterised by 'want'. They feel different. A feeling of 'ought', or a feeling of right and wrong, has a different character than feelings of hunger or desires for status. 'I want to do this' or 'I should do this' are not simply different wordings of the same statement.

Among these feelings of 'ought', there is a sense of responsibility. When this feeling is prompted, it creates a pressure to be addressed like any other feeling. Individuals take steps to address this feeling in order to maintain balance. What makes this, and other feelings related to 'ought', different is that it is not a totally self-interested motive. Addressing it is self-interested, but having it suggests that the individual's system is not purely so.

The knowledge of others being in need can prompt this feeling of responsibility and by doing so lead to a response. This is not the only reason for helping, but it is arguably an important one.

Responses to feelings of responsibility or potential feelings of responsibility do not, however, have to be in the form of helping. Individuals have the potential to manipulate, to an extent, their assessment of the context and alter the feelings of responsibility that are generated by such an assessment. As their aim is not to help but rather to maintain balance, they have motive to avoid the invoking of the feeling.

When trying to understand the responses of people to others in need, be it because of ill health or some other problem, then there is a need to understand what prompts feelings of responsibility and what manipulations are available to reduce them. Perceived similarity, degree of need, possibility of fault, urgency and the possibilities for manipulating, may each be crucial to understanding individual responses driven by feelings of responsibility.

To understand the implications of this individual response, it is necessary to consider how individuals interact with their environment. Where information comes from, what guides are available in suggesting what implications different types of information should have, and how context affects other needs, all influence the individual response to the needs of others. Furthermore, it is necessary to consider the policy context to understand the possibility for individual responses to collectively influence policy.

9.4 The implications

There are many behavioural models in the economics and psychology literature. The important difference of the above argument from many of the others is that it is an internal state that is the focus, not actions or external outcomes. The individual does not value the act of helping in and of itself, nor do they value the increase in another's welfare in and of itself. They value the way these outcomes make them feel. It is the feeling they value; if an alternative way of achieving that feeling is available then they may well take that option. But, as they are seeking balance, then once a feeling of responsibility is evoked they have either to help or to convince themselves of the inappropriateness of the feeling, so that they can reduce the pressure. This difference in what is valued is what drives the possibility for avoidance of information and self-manipulation, which are argued to be very important. If individuals do indeed seek to avoid feeling a responsibility to help, and if they try to manipulate this feeling away, there are implications for when others will be helped. If it is not the action or the outcome of help which is valued, then when the possibility for avoiding responsibility is high, individuals will try to do so.

There are other models which could be constructed that would generate very similar predictions. These models, however, are argued to fail to capture the distinctive nature of feelings of responsibility or other feelings of 'ought'. This has important implications in terms of the conceptualisation of individuals as

self-interested or not. More than this, using the above argument suggests a particular approach to applied research. In terms of applied research, the model suggests the usefulness of examining responses in terms of emotional prompts and in turn linking these to perceptions of the context. This may assist in understanding how contextual factors come into play in the decision-making process. The model also highlights the importance of examining how information is interpreted and how individuals might do this in a motivated way. If the model is correct, they are motivated to assess situations in such a way as to minimise the responsibility they feel. This possibility suggests the importance of examining what excuses are used and what type of information lends itself to the constructions of understanding which minimise felt responsibility.

Considering a social context and how an individual, with feelings of responsibility that they can manipulate, interacts with it should provide a powerful lens through which to examine everyday responses to real-world need. Applying the lens to the question of the slow response to HIV and AIDS certainly suggested some strong arguments as to the reasons for the delay and the recent change. The changing policy context, and the shift from a perception of individual fault to victims of greed, could help explain the change in response. More than explaining past events, the model, when applied to real-world situations, has predictive potential and suggests the need to watch for a dropping off of support, particularly for needs other than treatment. Treatment may be framed in the media and by individuals as responsibility met. When considering prevention, the frame goes back to one of individual fault; as a result, there is an opportunity to conclude that there is no responsibility to help.

While developed with the view to examining the response of individuals to the needs of distant others, the above argument could potentially be applied to other areas of interest. The distance was introduced so as to have more chance of isolating responses not motivated by reciprocity, or other motives, that are relevant when the person in need is close by. While these responses

may come back into play when considering close need, it is possible that feelings of responsibility are also important in these contexts, alongside other motives. To what extent does the nurse in the under-resourced clinic, or the local policy maker with a tight budget, for example, take on the responsibility of the job? Do their decisions reflect the interpretation of responsibility of the institutions they represent or their own feelings of responsibility or some combination? Related to the question of taking on responsibility, the model has potential in the examination of corruption. Most behavioural models would suggest the importance of monitoring, catching and punishing the corrupt. While these would still be seen, through the balance model lens, as important, more could be added. The balance model would suggest a need for the corrupt to have justification. Trying to understand how individuals justify their corrupt activities and how these justifications link to social norms, which may make it easier, may assist in developing responses beyond capture and punishment. Capture and punishment reduce the motivation to find justification. If this is combined with efforts to combat justifications, the impact may be greater.

Unlike purely self-interested constructions, the balance model suggests a strong role for self-regulation, as opposed to the 'rational psychopath' of utility maximisation. Applying the model suggests the importance of ways in which to ensure these self-regulatory mechanisms have more effect in the areas of interest. This line of thinking would suggest the importance of identifying in what areas of behaviour these self-regulatory mechanisms exist, and ways to strengthen them when there is a social benefit from doing so.

9.5 Final question

The above argument was originally developed with a view to understanding what shapes responses to others' need. As discussed in the introduction to this thesis, it is not typically those most in need that make decisions about what need is addressed. It is the poor who suffer the most, while often it is the rich

who decide who among them lives and dies. Given this situation, it is important to understand the basis of these decisions.

As far as feelings of responsibility play a role, it does not appear that the efforts of those with means to help are directed so as to generate the greatest benefit. Simply note the millions of people who die each year from preventable conditions, while millions of pounds are raised for Siamese-twins to be separated. Feelings of responsibility clearly interact with contexts and constructed understandings of the world to generate a pattern of help quite distinct from the pattern of need. So long as the situation is as it is, with those who have deciding for those who need, then this mismatch between help and need will continue.

The world is a closed system. If someone has more than the average, then someone else has less than the average. The implications of this are uncomfortable outside of a constructed understanding that seeks to avoid them. If the poor were so because of their own fault, or because of their own corrupt governments, or because of famine and other nature disasters, then seeing our responsibility to them being met through charity and aid would seem appropriate. The situation is, however, not so simple. The poor of the world are poor, at least in part, because we are not. They die because we have more than we need and they have too little. If we were honest with ourselves what responsibility would this truth bring?

References

Akerlof, G. & Dickens, W. 1982, "The economic consequences of cognitive dissonance", *The American Economic Review*, vol. 72, no. 3, pp. 307-319.

American Association of Fundraising Council 1994, *Giving USA*, American Association of Fundraising Council, New York.

Anand, P. & Wailoo, A. 2000, "Utilities versus rights to publicly provided goods: arguments and evidence from health-care rationing", *Economica*, vol. 67, no. 268, pp. 543-577.

Anderson, C. & Huesmann, R. 2003, "Human Aggression: A Social-Cognitive View," in *The Sage Handbook of Social Psychology*, M. Hogg & J. Cooper, eds., Sage Publications Inc, CA, pp. 296-323.

Andreoni, J. 1989, "Giving with impure altruism: applications to charity and Ricardian equivalence", *The Journal of Political Economy*, vol. 97, no. 6, pp. 1447-1458.

Andreoni, J. 1990, "Impure altruism and donations to public goods: a theory of warm glow giving", *The Economic Journal*, vol. 100, no. 401, pp. 464-477.

Andreoni, J. & Miller, J. Giving according to GARP: An experimental study of rationality and altruism. 1998.

Ref Type: Unpublished Work

Arnab, R. & Serumaga-Zake, A. 2006, "Orphans and vulnerable children in Botswana: the impact of HIV/AIDS", *Vulnerable Children and Youth Studies*, vol. 1, no. 3, pp. 221-229.

Arrow, K. 1951, *Social Choice and Individual Values* Wiley, New York.

Arrow, K. 1972, "Gifts and Exchanges", *Philosophy and Public Affairs*, vol. 1, no. 4, pp. 343-362.

Bandura, A. 1999, "Moral disengagement in the perpetration of inhumanities", *Personality and Social Psychology Review*, vol. 3, no. 3, pp. 193-209.

Barnett, T. & Whiteside, A. 2006, *AIDS in the 21st Century: Disease and Globalisation* Palgrave Macmillian, London.

Bartels, L. 1992, "The impact of electioneering in the United States," in *Electioneering*, D. Butler & A. Ranney, eds., Oxford University Press, New York, pp. 244-277.

Bassili, J. 1989, *On-Line Cognition in Person Perception* Lawrence Erlbaum, Hillsdale.

Batson, C., Dyck, J., Brandt, J., Batson, J., Powell, A., McMaster, M., & Griffitt, C. 1988, "Five studies testing two new alternatives to the empathy-altruism hypothesis", *Journal of Personality and Social Psychology*, vol. 55, pp. 52-77.

Batson, C. 1991, *The Altruism Question: Toward a Social Psychological Answer* Routledge, NY.

Batson, C. 1993, "Experimental tests for the existence of altruism", *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 1992, Volume Two: Symposia and Invited Papers, pp. 69-78.

Batson, C. & Ahmed, N. 2001, "Empathy-induced altruism in a prisoner's dilemma II: what if the target of empathy has defected?", *European Journal of Social Psychology*, vol. 31, no. 1, pp. 25-36.

Batson, C., Lishner, D., Cook, J., & Sawyer, S. 2005, "Similarity and Nurturance: two possible sources of empathy for strangers", *Basic and Applied Social Psychology*, vol. 27, no. 1, pp. 15-25.

- Becker, G. 1962, "Irrational behaviour and economic theory", *The Journal of Political Economy*, vol. 70, no. 1, pp. 1-13.
- Becker, G. 1974, "A Theory of Social Interactions", *The Journal of Political Economy*, vol. 82, no. 6, pp. 1063-1093.
- Bennett, R. & Kottasz, R. 2000, "Emergency fund-raising for disaster relief", *Disaster Prevention and Management*, vol. 9, no. 5, pp. 352-359.
- Bennett, R. & Kottasz, R. 2001, "Advertising imagery employed by disaster relief organisations and media stereotyping of the recipients of aid", *Developments in Marketing Science*, vol. 24, no. 1, pp. 225-230.
- Bennett, R. & Daniel, M. 2002, "Media reporting of Third World disasters: the journalist's perspective", *Disaster Prevention and Management*, vol. 11, no. 1, pp. 33-42.
- Bennett, R. 2003, "Factors underlying the inclination to donate to particular types of charity", *Voluntas*, vol. 8, no. 1, pp. 12-29.
- Bennett, R. & Gabriel, H. 2003, "Image and reputational characteristics of UK charitable organisations: An empirical study", *Corporate Reputation Review*, vol. 6, no. 3, pp. 276-289.
- Bihm, E., Gaudet, I., & Sale, O. 1979, "Altruistic responses under conditions of anonymity", *The Journal of Social Psychology*, vol. 109, pp. 25-30.
- Bishai, D. 2003, "Does biological relatedness affect child survival?", *Demographic Research*, vol. 8, no. 9, pp. 261-278.
- Brekke, K., Kverndokk, S., & Nyborg, K. 2003, "An economic model of moral motivation", *Journal of Public Economics*, vol. 87, no. 9-10, pp. 1967-1983.
- Buchanan, J. 1954, "Individual choice in voting and the market", *The Journal of Political Economy*, vol. 62, no. 4, pp. 334-343.

Buchanan, J. 1975, "The Samaritan's Dilemma," in *Altruism, Morality, and Economic Theory*, E. Phelps, ed., Russell Sage Foundation, New York, pp. 71-85.

Bullock, H., Wyche, K., & Williams, W. 2001, "Media images of the poor", *Journal of Social Issues*, vol. 57, no. 2, pp. 229-246.

Burstein, P. 2003, "The impact of public opinion on public policy: a review and an agenda", *Political Research Quarterly*, vol. 56, no. 1, pp. 29-40.

Busby, J. From benign neglect to moral awakening: Donor responses to HIV/AIDS. 2006.

Ref Type: Unpublished Work

Cacioppo, J., Losch, M., & Petty, R. 1986, "Attribution of Responsibility for helping and Doing Harm: Evidence for Confusion of Responsibility", *Journal of Personality and Social Psychology*, vol. 50, no. 1, pp. 100-105.

Campbell, F. 1997, "Journalistic construction of news: information gathering", *New Library World*, vol. 98, no. 1133, pp. 60-64.

Case, A., Paxson, C., & Ableidinger, J. 2004, "Orphaned and vulnerable to infection, under-nutrition and early death: increasing threats to infants and children", *Demography*, vol. 41, no. 3, pp. 483-506.

Clawson, R. & Trice, R. 2000, "Poverty as we know it: media portrayals of the poor", *Public Opinion Quarterly*, vol. 64, no. 1, pp. 53-64.

Clotfelter, C. 2001, "Who are the alumni donors? Giving by two generations of alumni from selective colleges", *Nonprofit Management and Leadership*, vol. 12, no. 2, pp. 119-138.

Chatterjee, D. 2004, "Introduction," in *The Ethics of Assistance: Morality and the Distant Needy*, D. Chatterjee, ed., Cambridge University Press, NY.

Cialdini, R., Brown, S., Lewis, B., Luce, C., & Neuberg, S. 1997, "Reinterpreting the empathy-altruism relationship: when one into one equals oneness", *Journal of Personality and Social Psychology*, vol. 73, no. 3, pp. 481-494.

Cohen, S. 2005, *States of Denial: Knowing about Atrocities and Suffering* Polity Press, Cambridge.

Collard, D. 1975, "Edgeworth's proposition on altruism", *The Economic Journal*, vol. 85, no. 338, pp. 355-360.

Cookson, R. & Dolan, P. 2000, "Principles of justice in health-care rationing", *Journal of Medical Ethics*, vol. 26, pp. 323-329.

Cozzarelli, C., Wilkinson, A., & Tagler, M. 2001, "Attitudes towards the poor and attributions for poverty", *Journal of Social Issues*, vol. 57, no. 2, pp. 207-227.

Dana, J., Cain, D., & Dawes, R. 2006, "What you don't know won't hurt me: costly (but quiet) exit in dictator games", *Organizational Behavior & Human Decision Processes*, vol. 100, no. 2, pp. 193-201.

Dana, J., Weber, R., & Kuang, J. 2007, "Exploiting moral wriggle room: Behavior inconsistent with a preference for fair outcomes", *Economic Theory*, vol. 33, no. 1, pp. 67-80.

DFID 2005, *Social Transfers and Chronic Poverty: Emerging Evidence and the Challenge Ahead*, Department for International Development, London.

Desmond, C., Richter, L., Makiwane, M., & Amoateng, A. 2003, "Dissecting our understanding of who is at risk", *ChildrenFirst* no. 43, pp. 56-58.

Dietz, T. & Stern, P. 1995, "Toward a theory of choice: socially embedded preference construction", *Journal of Socio-Economics*, vol. 24, no. 2, pp. 261-279.

Dolan, P., Cookson, R., & Ferguson, B. 1999, "Effect of discussion and deliberation on the public's views of priority setting in health care: focus group study", *BMJ*, vol. 318, no. 7188, pp. 916-919.

Doob, A. & Ecker, B. 1970, "Stigma and Compliance", *Journal of Social Psychology*, vol. 14, no. 4, pp. 302-304.

Driver, E. 1987, "Self-Evaluations and altruistic behaviour", *The Journal of Social Psychology*, vol. 124, no. 4, pp. 393-394.

Druckman, J. & Lupia, A. 2000, "Preference Formation", *Annual Review of Political Science*, vol. 3, pp. 1-24.

Dutton, D., Boyanowsky, E., & Bond, M. 2005, "Extreme mass homicide: from military massacre to genocide", *Aggression and Violent Behavior*, vol. 10, no. 4, pp. 437-473.

Eagly, A. & Crowley, M. 1986, "Gender and helping behaviour: a meta-analytic review of the social psychological literature", *Psychological Bulletin*, vol. 100, no. 3, pp. 283-308.

Eckel, C. & Grossman, P. 1996, "Altruism in anonymous dictator games", *Games and Economic Behavior*, vol. 16, no. 2, pp. 181-191.

Eco, U. 2001, "When the Other Appears on the Scene," in *Five Moral Pieces*, U. Eco, ed., Harcourt, Italy.

Edelmann, R., Childs, J., Kellock, I., & Strain-Clark, C. 1984, "The effect of embarrassment on helping", *The Journal of Social Psychology*, vol. 124, no. 2, pp. 253-254.

Edgeworth, F. 1881, *Mathematical Psychics* C. Kegan Paul & Co, London.

Elster, J. 1989, *The Cement of Society* Cambridge University Press, Cambridge.

Enelow, J. & Hinich, M. 1984, *The Spatial Theory of Voting: An Introduction* Cambridge University Press, New York.

Etzioni, A. 1986, "The case for a multiple-utility conception", *Economics and Philosophy*, vol. 2, pp. 159-183.

Etzioni, A. 1988, "Normative-Affective factors: towards a new decision-making model", *Journal of Economic Psychology*, vol. 9, pp. 125-150.

Fazio, R. 1995, "Attitudes as object-evaluation associations: determinates, consequences and correlates of attitude accessibility," in *Attitude Strength: Antecedents and Consequences*, R. Petty & J. Krosnick, eds., Lawrence Erlbaum, Mahwah.

Fischer, P., Greitemeyer, T., Pollozek, F., & Frey, D. 2005, "Unresponsive bystander behaviour: are bystanders more responsive in dangerous emergencies?", *European Journal of Social Psychology*, vol. 36, no. 2, pp. 267-268.

Folbre, N. & Goodin, R. 2004, "Revealing Altruism", *Review of Social Economy*, vol. 62, no. 1, pp. 1-25.

Frank, R. 1987, "If Homo Economicus could Choose his own Utility Function, would he want one with a conscience?", *The American Economic Review*, vol. 77, no. 4, pp. 593-604.

Frank, R. 1988, *Passions Within Reason: The Strategic Role of the Emotions* W.W.Norton, New York.

Frank, R. 1989, "If Homo Economicus could choose his own utility function, would he want one with a conscience? Reply", *The American Economic Review*, vol. 79, no. 3, pp. 594-596.

Fudenburg, D. & Tirole, J. 1991, *Game Theory* MIT, Boston.

- Gamson, W. 1992, *Talking Politics* Cambridge University Press, Cambridge.
- Gilens, M. 1996, "Race and Poverty in America: Public misperceptions and the American news media", *Public Opinion Quarterly*, vol. 60, no. 4, pp. 515-541.
- Greene, W. 1997, *Econometric Analysis*, 3rd edn, Prentice-Hall, New Jersey.
- Hammond, P. 1975, "Charity: Altruism or Cooperative Egoism?," in *Altruism, Morality, and Economic Theory*, E. Phelps, ed., Russell Sage Foundation, New York, pp. 115-131.
- Harrington, J. 1989, "If Homo Economicus could Choose his own Utility Function, would he want one with a Conscience?: Comment", *The American Economic Review*, vol. 79, no. 3, pp. 588-593.
- Harris, M. & Samerotte, G. 1975, "The effects of aggressive and altruistic modelling on subsequent behaviour.", *The Journal of Social Psychology*, vol. 95, pp. 173-182.
- Harris, M., Benson, S., & Hall, C. 1975, "The effects of confession on altruism", *The Journal of Social Psychology*, vol. 96, pp. 187-192.
- Harris, M. & Samerotte, G. 1976, "The effects of actual and attempted theft, need, and a previous favour on altruism", *The Journal of Social Psychology*, vol. 99, pp. 193-202.
- Harsanyi, J. 1955, "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility", *The Journal of Political Economy*, vol. 63, no. 4, pp. 309-321.
- Hastie, R. & Park, B. 1986, "The relationship between memory and judgement depends on whether the task is memory-based or on-line", *Psychological Review*, vol. 93, pp. 258-268.

Hauser, M. 2007, *Moral Minds: The Nature of Right and Wrong* Harper Perennial, New York.

Hausman, D. 2005, "Sympathy, Commitment and Preference", *Economics and Philosophy*, vol. 21, pp. 33-50.

Hobolt, S. & Klemmensen, R. 2005, "Responsive government? Public opinion and government policy preferences in Britain and Denmark", *Political Studies*, vol. 53, no. 2, pp. 379-402.

Hoffman, M. 1991, "Is empathy altruistic?", *Psychological Inquiry*, vol. 2, no. 2, pp. 131-133.

Iyengar, S. & Kinder, D. 1987, *News that Matters: Television and American Opinion* University of Chicago Press, Chicago.

Iyengar, S. 1989, "How citizens think about national issues: a matter of responsibility", *American Journal of Political Science*, vol. 33, no. 4, pp. 878-900.

Iyengar, S. 1990, "Framing Responsibility for Political Issues: The case of poverty", *Political Behavior*, vol. 12, no. 1, pp. 19-40.

Johansson-Stenman, O. & Martinsson, P. 2003. "Are some lives more valuable?", Department of Economics, *Goteborg University, Working Papers in Economics*, no. 96.

Johansson-Stenman, O. & Svedsater, H. 2003. "Self image and choice experiments: Hypothetical and actual willingness to pay". *Goteborg University, Working Papers in Economics*, no. 94.

Jones, A. & Posnett, J. 1991, "Charitable donations by UK households: evidence from the Family Expenditure Survey", *Applied Economics*, vol. 23, no. 2, pp. 343-351.

Jones, A. & Marriott, R. 1994, "Determinants of the level and methods of charitable giving in the 1990 Family Expenditure Survey", *Applied Economics Letters*, vol. 1, no. 11, pp. 200-203.

Katz, M. 1989, *The Undeserving Poor* Pantheon, New York.

Khalil, E. 1999, "Sentimental fools: a critique of Amartya Sen's notion of commitment", *Journal of Economic Behavior & Organization*, vol. 40, no. 4, pp. 373-386.

Khalil, E. 2004a, "What is altruism?", *Journal of Economic Psychology*, vol. 25, no. 1, pp. 97-123.

Khalil, E. 2004b. "Integrity, shame and self-rationalization". Vassar College, Department of Economics Working Paper Series.

Kinder, D. 1998, "Communication and Opinion", *Annual Review of Political Science*, vol. 1, pp. 167-197.

Knecht, T. & Weatherford, M. 2006, "Public opinion and foreign policy: the stages of presidential decision making", *International Studies Quarterly*, vol. 50, no. 3, pp. 705-727.

Konecni, V. & Ebbesen, E. 1975, "Effects of the presence of children on adults' helping behaviour and compliance: two field studies", *The Journal of Social Psychology*, vol. 97, pp. 181-193.

Konow, J. 2000, "Fair shares: accountability and cognitive dissonance in allocation decisions", *The American Economic Review*, vol. 90, no. 4, pp. 1072-1091.

Konow, J. Mixed feelings: theories and evidence of warm glow and altruism. 2005.

Ref Type: Unpublished Work

Kothari, A. 1999. "Soliciting Public Preferences for Health Care Priorities: a Critical Review of Methodologies", McMaster University for Health Economics and Policy Analysis Research Working Paper 99-07.

Kottasz, R. 2004, "Differences in the donor behaviour characteristics of young affluent males and females: empirical evidence from Britain", *Voluntas*, vol. 15, no. 2, pp. 181-203.

Krosnick, J. & Kinder, D. 1990, "Altering the foundations of popular support for the president through priming: Reagan and the Iran-Contra affair", *American Political Science Review*, vol. 84, no. 2, pp. 495-512.

Krosnick, J. & Brannon, L. 1993, "The impact of the Gulf War on the ingredients of presidential evaluations: multidimensional effects of political involvement", *American Political Science Review*, vol. 87, no. 4, pp. 963-975.

Kruger, D. 2003, "Evolution and altruism: combining psychological mediators with naturally selected tendencies", *Evolution and Human Behaviour*, vol. 24, no. 2, pp. 118-125.

Latane, B. & Darley, J. 1968, "Group inhibition of bystander intervention in emergencies", *Journal of Personality and Social Psychology*, vol. 10, no. 3, pp. 215-221.

Latane, B. & Darley, J. 1970, *The Unresponsive Bystander: Why doesn't he help?* Appleton Century Crofts, New York.

Lenaghan, J. 1999, "Involving the public in rationing decisions. The experience of citizen juries.", *Health Policy*, vol. 49, no. 1-2, pp. 63-74.

Levine, M. 1999, "Rethinking bystander non-intervention: social categorization and the evidence of witnesses at the James Bulger murder trial", *Human Relations*, vol. 52, no. 9, pp. 1133-1155.

Lodge, M. & Steenbergen, M. 1995, "The responsive voter: campaign information and the dynamics of candidate evaluation", *American Political Science Review*, vol. 89, no. 2, pp. 309-326.

Long, S. & Freese, J. 2006, *Regression Models for Categorical Dependent Variables Using Stata* Stata Press, College Station.

Lynne, G. 1999, "Divided self models of the socioeconomic person: the metaeconomics approach", *Journal of Socio-Economics*, vol. 28, no. 3, pp. 267-288.

Macaulay, J. 1975, "Familiarity, Attraction, and Charity", *The Journal of Social Psychology*, vol. 95, pp. 27-35.

Manza, J. & Cook, F. 2002, "A Democratic Policy? Three views of policy responsiveness to public opinion in the United States", *American Politics Research*, vol. 30, no. 6, pp. 630-667.

McGraw, K. & Pinney, N. 1990, "The effects of general and domain-specific expertise on political memory and judgement", *Social Cognition*, vol. 8, no. 1, pp. 9-30.

Micklewright, J. & Wright, A. 2003. "Private Donations for International Development", World Institute for Development Economics Research Discussion Paper.

Mill, J. 2002 [1863], *Utilitarianism and On Liberty* Blackwell Publishers, Oxford.

Minkler, L. 1999, "The problem with utility: toward a non-consequentialist/utility theory synthesis", *Review of Social Economy*, vol. 57, no. 1, pp. 4-24.

Monroe, K., Barton, M., & Klingemann, U. 1990, "Altruism and the theory of Rational Action: rescuers of Jews in Nazi Europe", *Ethics*, vol. 101, no. 1, pp. 103-122.

Murray, C. & Lopez, A. 1997, "Mortality by cause for eight regions of the world: global burden of disease study", *Lancet*, vol. 349, pp. 1269-1276.

Nettle, D. 2005, *Happiness: The Science Behind your Smile* Oxford University Press, Oxford.

Neuberger, J., Adams, D., MacMaster, P., Maidment, A., & Speed, M. 1998, "Assessing priorities for allocation of donor liver grafts: survey of public and clinicians", *BMJ*, vol. 317, no. 7152, pp. 172-175.

OECD 2005, *The DAC Journal Development Co-operation Report 2004*, OECD.

Oliner, S. & Oliner, P. 1988, *The Altruistic Personality* The Free Press, New York.

Page, B. 1994, "Democratic Responsiveness? Untangling the links between opinion and policy", *Political Science and Politics*, vol. 27, no. 1, pp. 25-29.

Payne, C. 1994, "Handling the press", *Disaster Prevention and Management*, vol. 3, no. 1, pp. 24-32.

Penner, L., Dovidio, J., Piliavin, J., & Schroeder, D. 2005, "Prosocial behaviour: multilevel perspectives", *Annual Review of Psychology*, vol. 56, pp. 365-392.

Pettifor, A., Rees, H., Steffenson, A., Hlongwa-Madikizela, L., McPhail, C., & Kleinschmidt, I. 2004, *HIV and Sexual Behaviour among Young South Africans: A National Survey of 15 – 24 Year Olds* Reproductive Health Research Unit, Johannesburg.

Pettit, P. 2005, "Construing Sen on commitment", *Economics and Philosophy*, vol. 21, no. 1, pp. 15-32.

Pharoah, C. 1996, *Individual giving: Seesaw Evidence, Static Trends, Dimensions of the Volunteer Sector*, Charities Aid Foundation, London.

Piliavin, J. & Charng, H. 1990, "Altruism: A review of recent theory and research", *Annual Review of Sociology*, vol. 16, pp. 27-65.

Popkin, S. & Dimock, M. 2000, "Knowledge, trust and international reasoning," in *Elements of Reason: Cognition, Choice and the Bounds of Rationality*, A. Lupia, M. McCubbins, & S. Popkin, eds., Cambridge University Press, New York.

Quarantelli, E. 1996, "Local mass media operations in disasters in the USA", *Disaster Prevention and Management*, vol. 5, no. 5, pp. 5-10.

Rabin, M. Moral preferences, moral constraints, and self-serving biases. 1995.
Ref Type: Unpublished Work

Richter, L. & Desmond, C. 2008, "Targeting AIDS orphans and child-headed households: A perspective from national surveys in South Africa, 1995-2005", *AIDS Care*, in press.

Richter, L., Manegold, J., & Pather, R. 2004, *Family and Community Interventions for Children Affected by AIDS* HSRC Press, Cape Town.

Ridley, M. & Dawkins, R. 1981, "The Natural Selection of Altruism," in *Altruism and Helping Behavior*, J. Rushton & R. Sorrentino, eds., Lawrence Erlbaum Associates, Hillsdale, pp. 19-40.

Robinson, P. 1999, "The CNN effect: can the news media drive foreign policy?", *Review of International Studies*, vol. 25, pp. 301-309.

Robinson, P. 2000, "The Policy-Media Interaction Model: measuring media power during humanitarian crisis", *Journal of Peace Research*, vol. 37, no. 5, pp. 613-633.

Rushton, J. 1980, *Altruism, Socialization, and Society* Prentice-Hall, Englewood.

Russell, G. & Mentzel, R. 1990, "Sympathy and altruism in response to disasters", *The Journal of Social Psychology*, vol. 3, no. 130, pp. 309-317.

Ryan, A., Scott, D., Reeves, C., Bate, A., van Teijlingen, E., Russell, E., Napper, M., & Robb, C. 2001, "Eliciting public preferences for health-care: a systematic review of techniques", *Health Technology Assessment*, vol. 5, no. 5.

Sargeant, A. 1999, "Charitable giving: towards a model of donor behaviour", *Journal of Marketing Management*, vol. 15, no. 4, pp. 215-238.

Schlegelmilch, B. & Love, A. 1997, "Responses to different charity appeals: the impact of donor characteristics on the amount of donations", *European Journal of Marketing*, vol. 31, no. 8, pp. 548-560.

Sen, A. 1977, "Rational fool: A critique of the behavioural foundation of economic theory", *Philosophy and Public Affairs*, vol. 6, no. 4, pp. 317-344.

Sen, A. 1985, "Goals, commitment and identity", *Journal of Law, Economics and Organization*, vol. 1, no. 2, pp. 341-355.

Sen, A. 2005, "Why exactly is commitment important for rationality?", *Economics and Philosophy*, vol. 21, no. 1, pp. 5-13.

Shisana, O., Rehle, T., & Simbayi, L. 2005, *South African National HIV Prevalence, HIV Incidence, Behaviour and Communication Survey* HSRC Press, Cape Town.

Shotland, R. & Straw, M. 1976, "Bystander response to an assault: When a man attacks a woman", *Journal of Personality and Social Psychology*, vol. 34, pp. 990-999.

Silver, M. 1980, *Affluence, altruism and atrophy* New York University Press, New York.

Smallman, C. 1997, "Read all about it - risk trends in the media: a research note", *Disaster Prevention and Management*, vol. 6, no. 3, pp. 160-164.

Solomon, L. & Grotta, P. 1976, "Imitation of a Helpful Model: The effect of level or emergency", *The Journal of Social Psychology*, vol. 99, pp. 29-35.

Stoker, L. 1993, "Judging presidential characteristic: the demise of Gary Hart", *Political Behavior*, vol. 15, pp. 193-223.

Tobin, K., Davey, M., & Latkin, C. 2005, "Calling emergency medical services during drug overdose: an examination of individual, social and setting correlates", *Addiction*, vol. 100, no. 3, pp. 397-404.

Tsuchiya, A., Dolan, P., & Shaw, R. 2003, "Measuring people's preferences regarding ageism in health: some methodological issues and some fresh evidence", *Social Science & Medicine*, vol. 57, no. 4, pp. 687-696.

Ubel, P. 1999, "How stable are people's preferences for giving priority to severely ill patients?", *Social Science & Medicine*, vol. 49, no. 7, pp. 895-903.

Ungar, S. 1979, "The effects of effort and stigma on helping", *The Journal of Social Psychology*, vol. 107, no. 1, pp. 23-28.

van der Pligt, J. & Eiser, J. 1984, "Dimensional salience, judgement and attitudes," in *Attitudinal Judgement*, J. Eiser, ed., Springer-Verlay, New York.

Varese, F. & Yaish, M. 2000, "The importance of being asked: the rescue of Jews in Nazi Europe", *Rationality and Society*, vol. 12, no. 3, pp. 307-334.

Warr, P. 1982, "Pareto optimal redistribution and private charity", *Journal of Public Economics*, vol. 19, no. 1, pp. 131-138.

White, M. 2004, "Can Homo Economicus follow Kant's categorical imperative?", *Journal of Socio-Economics*, vol. 33, no. 1, pp. 89-106.

White, M. 2006, "Multiple utilities and weakness of will: a Kantian perspective", *Review of Social Economy*, vol. 64, no. 1, p. 20.

Wilson, D. & O'Gorman, R. 2003, "Emotions and actions associated with norm-breaking events", *Human Nature*, vol. 14, no. 3, pp. 277-304.

Zaller, J. 1992, *The Nature and Origins of Mass Opinion* Cambridge University Press, New York.