

**Do What Consumers Say Matter?
The Misalignment of Preferences with Unconstrained Ethical
Intentions**

Pat Auger
Melbourne Business School
200 Leicester Street
Carlton, Victoria 3053 AUSTRALIA
+61 (3) 9349-8194 (phone)
+61 (3) 9349-8414 (fax)
P.AUGER@MBS.EDU

Timothy M. Devinney
Australian Graduate School of Management
Sydney NSW 2052 AUSTRALIA
+61 (2) 9931-9382 (phone)
+61 (2) 663-4672 (fax)
T.DEVINNEY@UNSW.EDU.AU

February 2005

Running Title: What Consumers Say

Abstract

Nearly all studies of consumers willingness to engage in ethical or socially responsible purchasing behavior is based on unconstrained survey response methods. In the present paper we ask the question of how well does asking consumers the extent to which they care about a specific social or ethical issue relate to how they would behave in a more constrained environment where there is no socially acceptable response. The results of a comparison between traditional survey questions of “intention to purchase” and estimates of individuals willingness-to-pay for social attributes in products reveal that simple survey questions are too “noisy” to provide operationally meaningful information and overstate intentions to a considerable extent.

Keywords: Ethical Product Features, Willingness-to-pay, Survey Methods

Introduction

Ethical consumerism has become an important topic in both the popular press and academic journals (Crane and Matten, 2004). High visibility issues such as the use of child labor in developing countries or the global environmental impact of production (and consumption) appear to increasingly affect the purchase decisions of consumers around the world (Auger et al., 2003; Creyer and Ross, 1997; Elliott and Freeman, 2001). This apparent change in the behavior of consumers has important implications for managers and could impact decisions as varied as the choice of location for production facilities, the human resources policies of the organization, or the role of ethical business practices within the corporate environment.

Despite this increasing focus on issues related to ethical consumerism the evidence on its importance seems, at best, contradictory. For example, a large scale survey by Market & Opinion Research International (MORI) found that over one third of consumers in the UK were seriously concerned with ethical issues (Mason, 2000). The same survey also suggested that the potential for ethical products in the UK could be as high as 30 percent of consumer markets. A similar survey by Corporate Edge reported even more support for the importance of ethical corporate behavior with 57 percent of consumers saying that they would stop buying a product if they knew that children were being employed to make it (Rogers, 1998).

Conversely, other researchers have suggested that the opinions of consumers do not appear to translate into changes in purchasing behavior (Carrigan and Attala, 2001). That is, there appears to be a gap between what consumers say about the importance of ethical issues and what they do at the checkout counter. This gap has even led some researchers to believe that research on ethical consumerism is inherently unreliable (Ulrich and Sarasin, 1995).

However, part of the gap may be due to the nature of the survey instruments used in the consumer research. In general, surveys on ethical consumerism have used simple ratings scales that may overstate the importance of the ethical issues, since there are clearly more socially-acceptable answers. These issues relate to the use of self-reports and survey instruments and have been well-studied in fields such as social psychology (Schwarz, 2003). Those studies have consistently indicated that the way in which the questions are asked can influence the answer given (Schwarz, 1999).

The primary purpose of this article is to investigate whether common survey instruments may misrepresent the opinions of consumers with respect to ethical issues. To accomplish this, we directly compare the results from a replication of the MORI poll to results from a structured choice experiment. Structured choice experiments force consumers to trade-off product attributes (including ethical attributes) against one another, which leads to more reliable estimates of relative valuation (or utility) (Auger et al., 2003; Louviere et al., 2000) than would be the case where such constraints on choice not imposed. This has importance not just to the veracity of academic research in consumer ethics—where clear and consistent measures of consumer ethical positions must be attained—but to the practical realm as well—where corporate and public policy debates based on erroneous information can have large-scale societal repercussions.

The next few sections briefly review some of the relevant literature on ethical consumerism and survey and experimental methods. The research approach is then described in detail followed by a presentation of the results. The article concludes with a discussion and implications of the research for both academia and corporate and public policy.

Literature Review

Ethical Consumerism

The concept of ethical consumerism has evolved considerably over the last decade from an almost exclusive focus on environmental issues (i.e., green consumerism) to a concept that more broadly incorporates matters of conscience (Carrigan et al., 2004). These matters of conscience can include such varied issues as animal welfare, labor standards, human rights, health-related issues, and, of course, environmental issues (Carrigan et al., 2004; Crane, 2001; Strong, 1996). In its broadest form, ethical consumerism can be defined as “the conscious and deliberate choice to make certain consumption choices due to personal and moral beliefs” (Crane and Matten, 2004). From a business or managerial perspective, ethical consumerism “implies that individual consumers can have a significant role, through their daily purchase decisions, in promoting ethical corporate practices” (Uusitalo and Oksanen, 2004, p. 215). Some of the ways by which consumers can more directly accomplish this is by purchasing (or not purchasing) certain products and/or by paying more for more socially acceptable products.

Most research, both commercial and academic, on ethical consumerism suggests that a growing number of consumers are increasingly taking ethical and social issues into account when purchasing products (Carrigan et al., 2004; Mason, 2000; Rogers, 1998; Uusitalo and Oksanen, 2004). Several factors including the emergence of pressure groups, increasing media interest in social and ethical issues, increasing focus on corporate social responsibility by major corporations, and the availability of better quality “ethical” products have been proposed as factors contributing to the growing popularity of ethical consumerism (Harrison, 2003; Spar and La Mure, 2003; Strong, 1996).

However, estimates of the importance of ethical issues in consumer purchasing decisions vary significantly depending on the survey methodology and/or source of the survey. For example,

the MORI poll mentioned earlier estimated that a third of consumers in the UK were seriously concerned with ethical issues, which could potentially translate into a relatively large market for socially desirable products (Mason, 2000). Another study of Finnish consumers by Uusitalo and Oksanen (2004) found that almost 70 percent of respondents believed that a firm's business ethics had at least "some influence" on their purchasing decisions. However, these and similar studies used survey instruments with simple ratings scales (e.g., Likert-type scales) that may overstate the importance of these ethical issues as there is no incentive to answer the questions truthfully (common problems with survey instruments will be discussed in the next section). These potential problems with survey instruments have led several researchers to propose the existence of an attitude-behavior gap with respect to the impact of social and ethical issues on consumer purchases (Boulstridge and Carrigan, 2000; Carrigan and Attala, 2001; Simon, 1995; Ulrich and Sarasin, 1995). That is, consumers indicate in surveys that ethical and social issues are important, but do not change their purchase behaviors accordingly.

A more rigorous stream of research has attempted to deal with these same issues and quantify the value of ethical product features for specific groups of consumers. The evidence to date suggests that some consumers are willing to pay more for products that are socially acceptable. For example, several studies conducted at Marymount University (1999) found that 75 percent of consumers would avoid shopping in a store if they knew the goods were produced under bad conditions. More importantly, these same consumers indicated that they would pay \$1 more for a \$20 item that was made under good conditions.

Another study by Elliott and Freeman (2001) produced some additional insights into the behavior of consumers. They uncovered relatively high elasticities of demand for products made under bad conditions but low elasticities for products made under good conditions. Hence, companies can potentially lose from having their products identified as being made under bad

conditions but have little to gain from marketing their products as being made under good conditions.

In a more recent study, Auger et al. (2003) used a choice modelling approach to provide more accurate willingness-to-pay estimates on a sample of consumers from Hong Kong and Australia. Choice experiments are especially well suited for this type of research since they force consumers to trade-off ethical features against traditional features. They found that some consumers were willing to pay more for products that possessed certain ethical attributes such as products that were not tested on animals or were manufactured by companies that did not use child labor. However, their results also clearly showed that consumers were not willing to sacrifice product quality or features for more socially acceptable products.

Overall, the evidence suggests that some consumers are willing to consider ethical and social issues when purchasing products. However, inherent weaknesses in the survey methodologies most often used in research on ethical issues has created doubts about the reliability of some of the results. For example, the sensitivity of ethical issues (e.g., use of child labor or animal testing) may motivate consumers to answer questions in a more socially acceptable way. Problems of this sort are not new to social science research and the next section highlights some of the relevant research on problems with survey methods.

Common Problems with Survey Instruments

Problems associated with the use of survey instruments and ratings scales are well-documented and researched (Box-Steffensmeier et al., 2000; Miller and Mitamura, 2003; Rucinski, 1993). Among the most useful research in this area is a series of articles by Schwarz and several of his colleagues (e.g., Schwarz, 1999, 2003; Schwarz et al., 1998; Schwarz and Hippler, 1995) that yielded several important and relevant insights for the present research. First, there are a variety of relatively well-known and obvious issues that may impact the reliability of

survey instruments, such as: 1) the wording of the questions so that respondents are clear about what the researcher is asking, 2) the choice of response alternatives such as open ended, frequency scales, rating scales and so on, and 3) the context of the questions such as the sponsoring organization and the nature of adjacent questions (Schwarz, 1999). As an example of the latter, Schwarz and Hippler (1995) found that in self-administered mail surveys responses to a question can not only be affected by the nature of previous questions but also by subsequent ones, since respondents often look ahead while answering the surveys. This finding is interesting, but adds a level of complexity in survey designs.

Second, and highly relevant to the present research, is the possibility that “respondents may want to edit their private judgment before they report it to the researcher, due to reasons of social desirability and self-presentation” (Schwarz, 1999, p. 97). This is fundamentally an ‘incentive compatibility’ issue—to what extent does the structure of the survey instrument allow (or force) respondents to reveal their ‘true’ underlying behaviour, preference or attitude. As mentioned previously, this is potentially a serious issue in ethics research given the sensitivity of the issues under investigation, the lack of any penalty for not revealing the ‘truth’, and the obviousness of the intent of the research to tap ethical attitudes and behavior.

Third, the features of rating scales and their informational content can have dramatic impacts on the results. The first issue is fairly obvious. The wording used in rating scales is critical and vague wording can lead to erroneous conclusions (Schwarz et al., 1998). For example, we cited a study in the previous section on ethical consumerism where 70 percent of respondents stated that a firm’s business ethics had “some influence” on their purchase behavior. This relatively vague label begs for more information—how much “influence” is “some influence” and is one person’s “some influence” equivalent to another person’s “some influence”?

The second issue, the informational content of the rating scale, is much less obvious. In short, rating scales (and more generally questionnaires) are often an important source of information for respondents. That is, respondents draw on the information embedded in rating scales (and questionnaires) to arrive at an answer (Schwarz, 1999). Furthermore, research also shows that respondents utilize this information more heavily when the behavior is poorly represented in memory and/or when the behavior is ill defined.

Unfortunately, there are no easy or full-proof solutions to solve these numerous problems. As Schwarz (2003, p. 593) noted, “every design decision involves complex trade-offs, requiring researchers to think through the issues at hand for every particular study”. Within the context of ethical consumerism, two categories of problems (and their associated solutions) are especially relevant and deserve further discussion: 1) the use of rating scales and basic questionnaire design and 2) a lack of incentive compatibility due to the sensitive nature of the issues under investigation.

The first category, the use of rating scales and basic questionnaire design, has been studied in multiple disciplines. Though there is no silver bullet, results from such studies suggest that researchers can take obvious precautions to reduce the impact of these problems such as 1) the use of cognitive pretests, which focus on differences in the interpretation of questions for a small group of respondents, 2) variation in response alternatives to identify potential order effects, 3) the creation of a context for the questionnaire that closely resembles the context in which consumers make decisions, and 4) the replacement of numeric frequency scales with other alternatives such as open-ended frequency reports (Klein et al., 2004; Schwarz and Oyserman, 2001 ; Zwick and Chen, 1999).

The second category, lack of incentive compatibility, is more difficult to deal with. Fortunately, there is a growing literature on researching sensitive issues such as sexual behaviors

and drug use that offers a number of useful guidelines. For example, Tourangeau and Smith (1996) compared three methods of collecting data about sexual behavior: computer-assisted personal interviewing, computer-assisted self-administered interviewing, and audio computer-assisted self-administered interviewing. Their results show that computer-assisted self-administration surveys tend to increase the willingness of respondents to answer questions about sensitive issues. These results are explained primarily by the greater sense of privacy fostered by self-administration and the greater legitimacy (scientific value of the study) when respondents are able to interact directly with a computer. Overall, the researchers propose that three variables mediate the effects of data collection method on the quality of data: the degree of privacy permitted, the level of cognitive burden imposed, and the sense of legitimacy fostered.

The previous discussion was not meant to be a comprehensive review of the many issues associated with survey research. Instead, it was meant to highlight some of the potential problems of survey research, especially those that are more relevant to research on ethical consumerism. It also does not suggest that results from these surveys are completely unreliable. The relevant question, and the primary issue for this research, is how do responses to survey questions about the importance of ethical issues relate to purchase behavior? Research that directly examines this important question is, to the authors' best knowledge, not yet available. However, researchers in other areas have investigated similar issues and research on the concept of trust presents the best example.

Glaeser et al. (2000) offers the best example of research that compares survey and experimental methods in the measurement of a single construct. Specifically, the authors used two separate trust experiments and a comprehensive survey with twelve different attitudinal measures of trust (a total of 137 questions) to examine how well (or poorly) these attitudinal measures predicted trust. The results showed that only two of the twelve measures of trust were

correlated with the actual amount of trust in their experiment. Furthermore, the authors stated that to find two correlated measures is not surprising given the large number of measures under investigation. In other words, there are bound to be some correlated items given the number of measures even if these correlated measures have little predictive power. Hence, the authors concluded that “trust survey questions at best only weakly predict trust” (Glaeser et al., 2000, p. 826). What they did conclude, however, was that relatively simple ‘revealed’ and ‘related’ behavior was related to trust. For example, they discovered that simple past ‘trusting’ acts, such as leaving ones dorm room unlocked when going to the shower or lending CDs to acquaintances, were strongly related to ‘trust’. The implication of this was that past behavior tapping the same underlying psychological tendencies was a better predictor of future behavior than any set of unconstrained questions.

In summary, a significant portion of the existing research on ethical consumerism suffers from weaknesses in the survey methodology used to gather the opinions of consumers. Issues associated with the use of rating scales and the potential for a lack of incentive compatibility (Carson et al., 2000) puts doubts on some of the results of previous research. Hence we make the following overarching hypothesis to imply that similar phenomena may be operating in ethical consumerism research:

H1: Unconstrained ratings questions on ethical preferences are only weakly related to a consumer’s willingness-to-pay to execute those preferences.

The implication of this hypothesis is that unconstrained responses abstract sufficiently from the realistic aspects of the constraints operating when consumers are making operative purchasing decisions to make such information of marginal validity for scientific evaluation and of limited value for managerial decision making.

Research Methods

We tested our hypothesis by designing and implementing a two-stage experiment that involved two survey instruments. The two instruments were an ethical disposition survey (EDS) and a choice experiment. The ethical disposition survey contained four sections. The first section included items taken almost verbatim from the MORI poll survey (see Appendix A for a list of items). The other three sections contained 20 Machiavellianism statements from the MACH IV Machiavellianism scale of Christie and Geis (1970) and 20 statements from Forsyth's (1980) Ethics Position Questionnaire to measure ethical idealism and moral relativism (i.e., 10 statements for each scale). For each of those statements, subjects were asked to rate their extent of agreement based on a 5-point Likert scale ranging from strongly disagree to strongly agree. MACH IV, moral relativism, and ethical idealism scores were calculated for each subject by summing the ratings for the MACHIV scale and averaging the ratings for the moral relativism and ethical idealism scales.

The choice experiment required subjects to: (1) evaluate their most recently purchased brand, (2) decide whether to consider and purchase 32 hypothetical bath soap or athletic shoe products (see Appendix B for an example of a hypothetical profile for athletic shoes), and (3) answer a series of socio-demographic questions. The important components for this paper are (2) and (3), particularly the 32 choice tasks. For each hypothetical product the subject was asked two questions:

1. If the [shoes/soap] described above were available in your local shops now, would you consider trying them/it (Tick ONE box only)? No Yes
2. If the [shoes/soap] described above were available in your local shops now, would you buy them instead of or in addition to your current [shoes/soap] next time you shopped for [these products] (Tick ONE box only)? No Yes

Bath soaps and athletic shoes were used as products because of familiarity and relevance to specific ethical issues, namely environmentalism, labor and animal rights. Subjects were randomly assigned either to bath soap or athletic shoe experiments; in other words, a subject either was making choices amongst only athletic shoes or only bath soaps, not both. The functional and ethical product attributes and the levels used for the product categories are shown in Appendix C. The functional product attributes were pre-tested to ensure their relevance to consumer purchase decisions and price levels were consistent with prices in both markets at the time of data collection. The ethical product attributes were chosen on the basis of their pertinence to each product category and representativeness of ethical concerns expressed by human rights, environmental and animal activist groups and journalists.

Our choice experiment also involved an overarching experiment in which subjects were supplied with a professionally designed news article about functional and ethical factors related to the production of bath soaps or athletic shoes (see Auger et al. 2003 for a description of these articles). Articles were pre-tested for believability and comparability across both product categories and these tests indicated that they were realistic and effectively similar in style and content. All subjects received a core part of the article describing functional features of their product category.

The choice experiment contained eight experimental conditions plus two control conditions. The eight experimental conditions included all attributes in the hypothetical product profiles (i.e., both functional and ethical attributes; see Appendix B) and the control conditions did not include any of the ethical attributes in the profiles (i.e., only the functional attributes were included). Subjects in the experimental conditions were randomly assigned to one of eight news articles for which the presence or absence of the ethical factors mentioned in the articles were systematically varied. Finally, subjects were randomly assigned to one of three conditions whereby two groups

of subjects received both a choice experiment and an EDS survey, and the third group only got the choice experiment. The order of survey administration was randomized so that subjects received a first survey (EDS or choice experiment), and received a different second survey (EDS or choice experiment).

From the standpoint of the analysis here, we only examined those subjects who did both the choice experiment with both functional and ethical attributes present and the EDS survey. This is necessary as all of our comparisons are at the individual level between the responses to the 32 options in the choice experiment and the EDS questions. Hence, the responses used here included all the subjects in the non-control groups.

There were three groups of subjects: (1) MBA students at an Australian university, (2) undergraduate students at a Hong Kong university (Calder (1982) and Calder et al. (1981) discuss the use of students in experiments), and (3) supporters of the human rights organization Amnesty International (in Australia). As an incentive to participate, the student subjects were given an opportunity to enter a lottery with total prizes of A\$500 (in Australia) and \$HK4,000 (in Hong Kong). The Amnesty International supporters had A\$5 donated to Amnesty International for each set of surveys completed. Hong Kong surveys were translated into Chinese and back-translated to ensure consistency.

We should emphasize that these samples are not meant to be representative of an underlying population but are chosen to provide a range of preferences that is as extreme as possible thereby generating sufficient variance to make meaningful comparisons. As the comparisons to be made in this study are to compare the responses of an individual from a set of survey questions to their responses to a set of choice experiments, the lack of representativeness of the sample to any specific population is not an issue of concern. What is relevant for this study is to ensure that the

sample is broad in terms of the variance of ethical attitudes, demographics, and potential response to the choice experiment.

In summary, the subjects were assigned to the different experimental conditions as follows: (1) each subject was randomly assigned to one of the two products (shoes or soap), (2) each subject was randomly assigned to one of two choice experiments (full profiles or control profiles), (3) subjects assigned to full profiles were randomly assigned to one of eight news articles; subjects assigned to control profiles were randomly assigned to one of two news articles, and (4) each subject was randomly assigned to one of three conditions for the EDS survey; i.e., subjects completed the EDS survey before the choice experiment, after the choice experiment, or did not complete the EDS survey at all.

Results

Simple Sample Results

In total, 1,253 people were surveyed: 396 undergraduate students in Hong Kong, 357 MBA students in Australia, and 500 Amnesty International supporters (also in Australia). 111 surveys were completed and returned from the Hong Kong students (28 percent), 162 from the Australian MBAs (45 percent), and 172 (34 percent) from the Amnesty International supporters. Hong Kong subjects are mostly undergraduates, hence are considerably younger than both the Australian graduate students and the Amnesty International supporters. They are also less likely to have children, be married or have postgraduate degrees. Considerably more women responded from the Amnesty International supporter (67 per cent) and Hong Kong (54 per cent) samples than in the Australian university sample (32 per cent) due largely to a greater proportion of females in these samples (the gender balance in each sample did not differ significantly from the sample characteristics).

As noted earlier, to investigate whether unconstrained and constrained opinions were related required variance in the sample. Overall, we see this in terms of the distribution of the lifestyles as indicated in the demographics as well as in the responses to the EDS items. Table 1 shows that the Amnesty International sample is less Machiavellian and less likely to take a morally relativistic stance than either the HK or Australia sample. The Australian sample was the least ethically ideal. Hence, these statistics indicate that a wide and diverse group of individuals with disparate dispositions and opinions took part in the study.

Simple MORI Poll Results

Table 1 presents the mean responses for all MORI poll questions for our three samples as well as the aggregate means for all respondents. In general, the Amnesty International (AI) sample showed greater concern about social and ethical issues than the Hong Kong (HK) and Australian MBA samples. These results were expected given the nature of Amnesty International and its primary focus on human rights issues. The HK sample showed the least concern for most social and ethical issues while the Australian MBA sample took the middle ground on most issues. As expected, the responses of the Australian MBA respondents tended to be more similar to the HK sample than to the AI sample. These differences are positive from a sample standpoint as they indicate a high degree of variance in peoples' responses to the survey items.

More interesting results can be observed from a correlation matrix of all items in the MORI poll (see Table 2). The correlation analysis shows relatively high correlation coefficients between the ethical items within each section. For example, item 5 (people paid enough to live on) is highly correlated (0.65, 0.72, 0.46, and 0.45) with the other ethical items in section 1 (items 6, 7, 8, and 9). Furthermore, that same item (item 5) is only weakly correlated (-0.08, -0.06, -0.36, -0.18, and -0.11) with the other, traditional features (items 1, 2, 3, 4, and 10). In fact, the only significant correlation is a negative correlation (-0.36) with item 3 (brand). The same

pattern can be seen for the second group of items (items 11 to 20). In all cases, items are highly inter-correlated with the lowest correlation coefficient between two items at 0.30.

Of greater interest are the relatively high correlation coefficients between the ethical items in the two sections. For example, item 12 (did not earn enough wages to live on) is highly correlated (0.50, 0.33, 0.30, 0.36, and 0.28) with all ethical items in section 1. In fact, only one of the fifty pairs of correlation coefficients is below 0.20 (between items 6 and 14). These results can be interpreted in two different ways. On the one hand, the results demonstrate relatively high levels of consistency in the way respondents answered the different questions about ethical issues in the MORI poll. On the other hand, the results also show that respondents do not engage in much differentiation amongst the issues. That is, respondents tended to have (or answer the questions as if they have) similar opinions of the ethical issues irrespective of the nature of the issue. The implication of this is that either individuals do indeed care about “everything” or that such patterns are an artefact of their desire to answer the questions in an “acceptable” manner.

Factor Analysis of Mori Poll Responses

The MORI poll items were factor analysed (using maximum likelihood extraction with Varimax rotation) to reduce the total items to a more manageable number and detect patterns in the data. The analysis yielded a 4-factor solution that explained approximately 60.5 percent of the variance (see Table 3). We labelled the four factors as follows: 1) worker rights, 2) general rights, 3) health and harassment, and 4) utilitarian. Only factor 4 (utilitarian) does not contain any ethical items.

A correlation matrix of the factor scores and the demographics reveals some interesting patterns (see Table 4). First, two of the three ethical scales (i.e., ethical idealism and Machiavellianism) were only weakly correlated with the factor scores. Only the Machiavellianism scale had a significant negative correlation with factor 2 (general rights).

Ethical idealism was not significantly correlated to any of the factors. For its part, moral relativism was negatively correlated to all of the factors including factor 4 (utilitarian), which does not contain any ethical items.

The implications of this is that the general ethical personality disposition measures reveal little about even how people answer unconstrained questions pertaining to their purchasing behaviour. This is confirmatory of the findings of Auger et al. (2003; p 292–293) showing that such dispositional measures have little discriminatory power when examining product consideration or purchasing intentions in a constrained choice experiment, even though they may be useful in identifying distinct personality types (e.g., Al-Khatib et al., 2005). It should be noted that this result is not due to any specific sample effects as revealed by these variables. The three samples have similar correlations between the three EDS constructs with only the correlations between Machiavellianism and moral relativism (0.303, $p < 0.001$) and ethical idealism (0.259, $p < 0.01$) being significant (that between moral relativism and ethical idealism is not significant at 0.016).

Second, respondent demographics (e.g., age, income, gender, etc.) were also very weakly correlated with the factor scores. In fact, income, education, age, and gender had only 4 out of 16 significant correlations with the factor scores. Only the sample variables had a large number of significant correlations with the factor scores, revealing that the different samples did indeed answer the questions in a different manner but this is unrelated to what can be revealed through examining the demographics alone. Of special interest is the very high correlation coefficient (0.616) between the AI sample variable and factor 2, which includes the human rights MORI poll item. Hence, one could indeed conclude that human rights do matter to people who support a human rights organization like Amnesty International.

Cluster Analysis of Mori Poll Responses

The next level of analysis involved cluster analyzing the factor scores to group individual respondents; in essence, segmenting our respondents based on their answers to the MORI poll. Our optimal solution produced six clusters (see Table 5) that focused on different groups of factors. For example, cluster 4 has positive factor scores for factor 3 (health and harassment), but negative scores for the other three factors. For its part, cluster 6 has positive scores for both factors 2 and 4, but negative for the other two factors (i.e., factors 1 and 3). Hence, we can characterize (or describe) the clusters based on the factors as follows: (C1) ethically uninterested, (C2) utilitarians with a (minor) conscience, (C3) social utilitarians, (C4) anti-exploitationists, (C5) anti-utilitarians, and (C6) liberal utilitarians. The ethically uninterested (C1) have few concerns with respect to any issues, utilitarian or ethical, other than some minor interest in basic worker rights. Three groups are basically utilitarian with a twist, C2, C3 and C6. C2, the utilitarians with a (minor) conscience, are strongly utilitarian but believe in some basic labor rights. C3, the social utilitarians, are broadly concerned about human, environmental and animal rights as well as worker rights and protection from harassment. The liberal utilitarians (C6) are harder to characterize as they are defined basically on their positive attitudes toward general rights (human, animal and environmental) but a general negative attitude to worker rights. Two groups are defined based on their disavowal of utilitarianism. C4, the anti-exploitationists, are strongly orientated against exploitative behavior (e.g., sacking due to pregnancy), while the anti-utilitarians (C5) have the strongest disavowal of utilitarian motives and broad support of all social rights and issues.

We attempt to get a better picture of the clusters by examining the ethical scales and basic demographic characteristics to determine if there were any differences between the clusters (see Table 5). Overall, there were very few differences between the clusters for the three ethical

scales. Specifically, only 5 of the possible 45 mean comparisons were significantly different with moral relativism (4 out of 15 significant differences) showing the greatest differences.

Machiavellianism and ethical idealism had only one significant difference in cluster means from a pool of 30 comparisons. These results indicate that the three ethical scales do not do an effective job at discriminating between the different clusters, a finding consistent with Auger et al. (2003).

The same basic results apply to the demographics with the exception of the sample variables. For example, only 2 cluster means for age and income (a total of 30 comparisons) were significantly different and the chi-square statistics for gender and education were not significant. This indicates that the clusters were relatively homogeneous in terms of age, gender, income, and education. The major differences in the clusters could be attributed to the samples. That is, the distribution of respondents from specific samples (i.e., HK, Australian MBA, and AI) tended to be concentrated within specific clusters. For example, over 85 percent of HK respondents were either in cluster 1 (the ethically uninterested) or cluster 4 (64% and 23%, respectively). Similarly, over 60 percent of the Australian MBA sample was in two clusters (clusters 2 and 4, utilitarians with a (minor) conscious and anti-exploitationists). The AI sample showed greater variation, but still a relative high level of concentration. For example, AI supporters were the dominant members of clusters 5 (anti-utilitarian) and 3 (social utilitarians) and had minor membership in clusters 1 and 2. In fact, only two clusters, clusters 4 and 6 (liberal utilitarians), had “relatively” balanced membership from all three samples.

Relationships Between MORI Poll Responses and Willingness-to-Pay Estimates

Our core analyses focused on the relationships between the MORI poll responses and the willingness-to-pay-estimates from the choice experiments. One of the most useful features of the choice experiment methodology is its ability to convert the probability of consideration and

purchase directly into conditional dollar equivalents (i.e., willingness-to-pay estimates). By comparing the dollar value of specific bundles of product features it is then possible to determine the dollar equivalent of the utility that a consumer gets from the addition or absence of a specific feature. The details of the procedure are presented in Henscher, Louviere and Swait (2000) and an example is presented in Auger et al. (2003). In short, we calculated willingness-to-pay estimates (WTP) for all respondents and features of both shoes and bath soap based on the responses to the 32 choice options, treating the choice to consider or buy separately (giving us 64 separate choices per individual). We then used these individual WTP estimates for the remaining analyses.

For simplicity in the discussion that follows we consider each ethical attribute separately but pool all of the functional product attributes together (a detailed breakdown is available from the authors).

Willingness to Pay: Ethical versus Functional Product Features

Our first group of analyses examined the differences in the importance of the ethical attributes for the different cluster groupings (see Table 6). We did this by comparing the mean ratios of the WTP for ethical features versus the WTP for functional features. Hence, a higher ratio for a particular cluster would signify that respondents from that cluster were willing to pay a greater portion of the purchase price for ethical features relative to functional features than respondents from other clusters.

Comparisons of the means for shoes found very few differences between the cluster groupings. In fact, two of the ethical features, the use of child labor and whether workers are paid above minimum wages, showed no significant differences between the clusters. The ethical feature with the most differences was “acceptable living conditions” with only 4 out of 15 differences significant. The results for bath soap are even more stunning since none of the

possible 45 mean comparisons are significantly different. Overall, only 10 percent of the differences in means were significant for shoes and none were significant for bath soap. As such, these analyses strongly suggest that respondents from the different clusters valued ethical attributes relative to functional attributes in a similar fashion.

Willingness to Pay: Individual Ethical Features and MORI Poll Responses

Our next group of analyses focused on the relationships between the individual items in the MORI poll and the WTP estimates for the ethical features. For these analyses, we used correlation analysis since regression analysis was not possible due to high levels of multicollinearity in the MORI poll responses (see Table 2). Table 7 presents the results for shoes while Table 8 shows the results for bath soap.

Both correlation matrices show relatively complex patterns of association between the MORI poll items and the WTP estimates for the ethical features. The analysis for shoes (Table 7) reveals strong correlations between most of the MORI poll items and the use of child labor, working conditions, and living conditions, but weak correlations with the minimum wage WTP estimates. For example, 10 of the 15 and 12 of the 15 ethical MORI poll items have correlation coefficients above 0.20 with child labor and living conditions, respectively. At first glance, this seems to support strong relationships between the way respondents answered the MORI poll survey and the WTP estimates from the choice experiments.

However, a closer examination of the correlations paints a much more complex and less supportive picture. For example, the MORI poll item corresponding to the use of child labor (i.e., were under the legal minimum age to work) is correlated with the WTP estimate for child labor (0.34), but the correlation is weaker than for several other MORI poll items such as: people paid enough to live on (0.45), human rights record of the country (0.47), had right to sick pay (0.39), and quality of the product (0.37), among others. The latter is especially troubling given

that it is completely unrelated to ethical consumerism. The result for the relationships between the MORI poll items and the WTP estimates for minimum wage are even more stunning. First, the two MORI poll items dealing with wage issues (i.e., people paid enough to live on and did not earn enough wages to live off) are not correlated to the WTP estimates for minimum wage (0.08 and 0.11, respectively). Even more puzzling is that environment related items such as little damage to the environment and work environment healthy (both 0.25) are correlated with the WTP estimates for minimum wage. In fact, the MORI poll items related to wage issues are highly correlated to all WTP estimates for ethical features except for the minimum wage WTP estimates. Hence, the results of the correlation analysis for shoes support our hypothesis.

The same basic results emerged from the bath soap analyses (see Table 8). In those cases, almost all of the MORI poll items were strongly correlated with animal testing (13 out of 15) including a number that are extremely difficult to explain such as the ability to join a union (-0.47) and minimum wages (-0.44). Nonetheless, the animal testing item in the MORI poll had the highest correlation coefficient (-0.49) of all of the items (though only marginally so). The MORI poll items related to the environment (little damage to the environment) was related to the WTP estimates for biodegradability (0.22), but the latter had stronger correlations with other, unrelated, MORI poll items such as had no right to sick pay (0.29) and the human rights record of the country (0.28). Finally, there were no significant relationships between the ethical MORI poll items and the WTP estimates for the use of animal by-products. In fact, those WTP estimates were more highly correlated with standard product features such as availability (0.20) and quality (0.21). These results once again lend strong support to our hypothesis.

GLM Models: MORI Poll Factors and WTP for Ethical Features

Our final group of analyses investigated the relationships between the factor scores from the MORI poll items (see Tables 3 and 4) and the WTP estimates for the ethical features. We used

regression analysis and obtained several General Linear Models (GLM) with the WTP estimates as the dependent variables and the factor scores as the independent variables. We grouped the WTP estimates for the functional features into a single variable to reduce the number of analyses and simplify the presentation of the results. The analyses are presented in Table 9

Overall, the results for shoes are more supportive of relationships between the MORI poll items and the WTP estimates than the correlation analyses. For example, the health & harassment factor (factor 3) is strongly associated with the WTP estimates for child labor and work conditions as would be expected. However, the worker rights factor (factor 1) is not associated with either the minimum wage or the work conditions estimates. In fact, the only significant association of the worker rights factor is with the estimates for living conditions, which should not be associated at all. These results along with the very weak association between the general rights factor (factor 2) and the estimates for minimum wage (which should also be strongly associated) leads us to conclude that the MORI poll items are very weak predictors of the WTP estimates for the ethical features of shoes.

The results for bath soap are even more supportive of this conclusion. The general rights factor (factor 2) should have the strongest association with the WTP estimates of the ethical features, since that factor contains the environmental items of the MORI poll. In fact, it is only weakly associated with animal testing and not significantly associated with the other dependent variables (biodegradability and animal by-products). The most strongly associated factor is the worker rights factor (factor 1), which should not be associated at all. In fact, it is very strongly associated with animal testing ($p=0.000$). Finally, the only factor associated with animal byproducts is the utilitarian factor (factor 4), which does not contain any ethical items. Hence, it is clear that the MORI poll items are extremely poor predictors of the WTP estimates of ethical features for bath soap, which supports our hypothesis.

Discussion

The primary purpose of this research was to investigate the relationships between traditional survey methods and experimental methods for the evaluation of the importance of ethical issues in consumer purchase decisions. Our overarching hypothesis predicted that the two types of methods would produce weakly correlated results based on the salience of problems with survey instruments (e.g., Schwarz, 1999; Schwarz et al., 1998; Schwarz and Hippler, 1995), the lack of incentive compatibility (Carson et al., 2000), and the existing empirical evidence from research on stated preferences (i.e., Glaeser et al., 2000) study on trust). At the core of this research is the issue of incentive compatibility or more precisely, the lack of incentive compatibility in much of the existing research on ethical consumerism. The issue we have sought to highlight is that unconstrained survey instruments, the modus operandi of much of the empirical research on ethical consumerism, do not force consumers to reveal their true attitudes or intentions due to inherent weaknesses in survey design and the sensitivity of the issues under investigation.

The Main Findings

Overall, our results strongly support our hypothesis. Specifically, the correlation analyses (see Tables 7 and 8) and the General Linear Models (Table 9) show very inconsistent patterns of associations between the MORI poll items and the WTP estimates for the ethical features of bath soap and shoes. For the correlation analysis for shoes, most of the MORI poll items are strongly correlated with the WTP estimates for the use of child labor even if most of these items should not be correlated. Especially difficult to explain are the MORI poll items relationship to traditional products features, such as quality and appearance/style. Furthermore, and equally difficult to explain, is the relatively low correlation of the MORI poll item related to child labor and the WTP estimate for the use of child labor. In fact, the MORI poll item related to child labor has only the sixth largest correlation coefficient with the WTP estimate for child labor.

The results for the WTP estimate for minimum wage are even more difficult to explain. In those cases, the two MORI poll items related to minimum wage issues are not significantly correlated to the WTP estimate for minimum wage while several other items such as the impact on the environment and the human rights record of the country are. The results of the correlation analyses for bath soap show almost the same pattern with the majority of MORI poll items highly correlated with one of the WTP estimates (i.e., animal testing) and weakly correlated (as well as strangely correlated) with the WTP estimates for the other two ethical features.

Implications for Ethics Researchers

These results, along with similar results from the other analyses (e.g., correlation analysis of MORI poll items and the GLM analyses), have important implications for research on ethical consumerism. First, they support the notion that traditional survey methods that make use of simple rating scales may be overstating the importance of ethical issues to the purchase behavior of consumers. Specifically, the correlation analysis of MORI poll items (see Table 2) strongly implies that respondents answered our replication of the MORI poll without discriminating between the different ethical issues. That is, the results are consistent with respondents selecting all (or very few) of the ethical items as important irrespective of the nature of the issues. This pattern of response is glaringly apparent in the correlation analyses (Tables 7 and 8) and GLM (Table 9) where a number of items are highly correlated (or associated) with the WTP estimates even when they should not, or vice versa in the case of minimum wage. What is even more important is that this tendency is as likely to be true of those who reveal themselves to be supportive of social causes (as in the case of Amnesty International supporters) as it is for those that do not. Hence, although we find that Amnesty International supporters are revealed to be more socially conscious when looking at the aggregate MORI poll responses (Table 1), this fact does not carry over when one examines specific ethical items and distinct product features.

We are not proposing that all the results from previous research are inherently unreliable or erroneous. What we are arguing is that traditional survey methods are very likely to not only overstate the importance of ethical issues but that they will add unwanted variance into the measurement process by swamping the true preferences with spurious information. This has implications for research, management, and public policy. For researchers and managers, our results highlight the importance of rigorous methodology and the potential problems associated with traditional survey methods. More importantly, it demonstrates the importance of incentive compatibility within the field of ethical consumerism and the tendency of respondents to answer questions about ethical issues in a socially acceptable way. Hence, we would urge researchers to take great care when designing studies to investigate issues associated with ethical consumerism as the incentive compatibility issue is one that is even more salient in situations where one is concerned with issues loaded with societal implications.

Implications for Policy Makers

For policy makers, our results would suggest a more cautious approach to drawing conclusions about the impact of ethical issues on consumer markets. The stakes for governments and public policy agencies are clearly high, since decisions can have serious repercussions for all parties involved. This is not to say that ethical issues are not important and should not be taken into consideration by policy makers. What we suggest is that policy makers should be more careful about jumping to conclusions based solely on evidence from traditional survey instruments. For example, in related work we show that one of the issues least likely to be supported by average individuals is the banning of genetically modified foods (citation withheld), even in countries (e.g., Germany) where this issue is perennially in the press.

Second, our results also strongly suggest that traditional survey instruments may not be specific enough about the ethical issues and the context under which those ethical issues

influence purchase decisions. Our results, along with evidence from prior research (e.g., Auger et al., 2003; Glaeser et al., 2000), imply that broadly worded questions may be simply too general to have predictive power. Glaeser et al. (2000, p. 827) came to similar conclusions when they discovered that the attitudinal measures of trust that were more strongly associated with actual levels of trust tended to be more specific. They stated that this “high degree of specificity might make these variables genuinely more predictive than those of a more general nature”. We propose that some consumers may be influenced by ethical issues, but that the level of influence may vary depending on the specific situation and context (e.g., the type of product, the specific type of issue, the nature of the purchase situation, etc.); something that is rarely considered when general attitudinal scales are being utilized. Hence, we believe that an experimental methodology that more closely mimics a real purchase situation may be more appropriate for this type of research. At the very least, it would be prudent to utilize a combination of methods instead of relying exclusively on a single method.

Third, our results also show that traditional purchase intention survey instruments may not be suitable for classifying consumers (i.e., for segmentation purposes). We found that the cluster groupings that we generated from our replication of the MORI poll survey were fairly homogeneous about their valuation of ethical features, their responses to ethical disposition measures (e.g., Machiavellianism), and their basic demographic characteristics. Specifically, we found very few differences between the clusters on the relative importance of ethical features versus the traditional features. In fact, only ten percent of the mean ratios of the WTP estimates for ethical features versus traditional features for shoes and none for bath soap were significantly different between the clusters (see Table 8).

There are a number of possible explanations for these results, but two appear most valid. First, the MORI poll items may be extremely poor at characterizing individuals based on their

attitudes towards ethical issues; something at which more tried and tested scales may do better (e.g., Al-Khatib et al., 2005). This is troubling given that this is essentially the purpose of the survey instrument and it is fairly direct in its question about intentions and attitudes. However, it is also consistent with the results that we discussed earlier in this section and in support of our hypothesis. In essence what differences exist are relatively minor differentiators on the variables that really matter. Second, the rating scales used in the MORI poll may be picking up differences in response styles as opposed to differences in attitudes. Response styles are well-researched in the social sciences and describe the tendencies to respond systematically to questionnaire items on some basis other than what the items were specifically designed to measure (Paulhus, 1991). This phenomenon is especially frequent in cross-country studies, since response styles have been shown to be culturally-based (Chen et al., 1995; Steenkamp and Ter Hofstede, 2002). As such, it is possible that our clustering is actually grouping individuals into segments with different response styles as opposed to segments with different attitudes towards ethical issues.

These results have important implications for managers given that firm behavior with respect to ethical issues could be used as a way to differentiate a product and/or organization in the marketplace (e.g., the Body Shop, Reebok, etc.). However, effective differentiation also requires managers to properly segment the market and understand the needs of consumers within those segments (Kotler, 1997). Failure to properly segment the market could lead managers to waste precious resources on communicating and selling to the wrong audience and/or focusing on issues that have no relevance or salience for their target market. Given that the traditional methods used by firms are little more than basic survey response and focus group research, it is unlikely the information they are getting on the responsiveness of their customers to social or ethical positioning is very accurate (indeed anecdotal information supports management frustration on this dimension).

In conclusion, our results are supportive of our hypothesis on the lack of relationships between results obtained from traditional survey instruments and results obtained from experimental methods. To some, our discussion and our conclusion may seem overly critical of prior research based exclusively on traditional survey instruments. This is not our intent. Research is fundamentally a process that builds and tries to improve upon previous work. Previous research on ethical consumerism has played a critical role in highlighting the potential impact of ethical issues on consumer decision making and motivating researchers to study the phenomenon more closely. Our intention is to put out a call for new approaches that can address an inherent issue in a field that is studying issues with inherent methodological complexities that make understanding human behavior even more difficult than normal. We hope that this research will spur and motivate additional research into this important topic.

Table 1: Mean Mori Poll and EDS Survey Responses

	Sample			Total
	Australia	Hong Kong	Amnesty International	
<i>Which, if any, of the following would you take into consideration when you were buying it? (Percent indicating that it would matter)</i>				
Appearance/style	79.8	82.5	67.7	74.0
Availability	40.3	5.3	35.9	32.8
Brand	51.2	45.6	16.7	32.6
Quality	97.7	12.3	86.9	79.4
People paid enough to live on	24.8	3.5	77.3	48.7
Little damage to environment	31.0	28.1	76.3	53.9
Work environment healthy	24.8	12.3	74.2	48.4
No animal testing	15.5	21.1	61.6	40.1
Human rights record of country	19.4	0.0	58.1	36.5
Need for the product	87.6	93.0	81.3	85.2
<i>Which, if any, of the following things about the people who made the product would affect your decision to buy it?*</i>				
Forced to work overtime	1.74	2.43	2.87	2.56
Did not earn enough wages to live off	2.18	2.48	3.62	3.19
Had no job security	2.81	2.68	2.84	2.56
Could be sacked if they became pregnant	2.13	2.86	3.81	3.53
Found their health to be in danger	3.38	2.88	3.90	3.58
Had no holidays or days off	3.37	2.69	3.65	3.23
Were not allowed to join a union	2.82	2.67	3.45	2.94
Had no right to sick pay	2.32	2.56	3.52	3.08
Were subject to discrimination or harassment.	2.66	3.28	3.82	3.54
Were under the legal minimum age to work	3.21	2.87	3.74	3.49
Machiavellianism	98.08	97.05	90.90	95.20
Moral Relativism (5-point scale)	3.18	3.35	2.64	2.05
Ethical Idealism (5-point scale)	3.19	3.81	3.69	3.55
N =	71	77	87	235

* Mean on 4 point scale from (1) Would Still Buy It, (2) Would Still Consider Buying It, (3) Would Make No Difference, (4) Would Definitely Not

Table 2: Correlation Matrix of Mori Poll Responses (All Respondents)

	Which, if any, of the following would you take into consideration when you were buying it?										Which, if any, of the following things about the people who made the product would affect your decision to buy it?									
	1. Appearance/style	2. Availability	3. Brand	4. Quality	5. People paid enough to live on	6. Little damage to environment	7. Work environment healthy	8. No animal testing	9. Human rights record of country	10. Need for the product	11. Forced to work overtime	12. Did not earn enough wages to live off	13. Had no job security	14. Could be sacked if they became pregnant	15. Found their health to be in danger	16. Had no holidays or days off	17. Were not allowed to join a union	18. Had no right to sick pay	19. Were subject to discrimination or harassment.	20. Were under the legal minimum age to work
1	1.00																			
2	0.33	1.00																		
3	0.20	0.28	1.00																	
4	0.33	0.26	0.15	1.00																
5	-0.08	-0.06	-0.36	-0.18	1.00															
6	-0.05	0.01	-0.25	-0.14	0.65	1.00														
7	-0.05	0.02	-0.23	-0.06	0.72	0.67	1.00													
8	-0.02	0.04	-0.27	-0.06	0.46	0.52	0.49	1.00												
9	-0.04	0.01	-0.14	-0.03	0.45	0.39	0.49	0.37	1.00											
10	0.18	0.19	0.07	0.15	-0.11	-0.14	-0.11	-0.02	-0.14	1.00										
11	-0.11	0.02	-0.27	-0.12	0.37	0.28	0.32	0.29	0.36	-0.01	1.00									
12	-0.06	0.07	-0.22	-0.12	0.50	0.33	0.30	0.36	0.28	-0.07	0.46	1.00								
13	-0.04	0.02	-0.20	-0.07	0.42	0.22	0.26	0.24	0.33	0.09	0.54	0.53	1.00							
14	-0.06	-0.01	-0.16	-0.12	0.42	0.16	0.35	0.23	0.29	-0.10	0.30	0.49	0.39	1.00						
15	-0.04	0.04	-0.25	-0.09	0.40	0.33	0.34	0.32	0.30	-0.12	0.35	0.54	0.30	0.63	1.00					
16	-0.09	0.05	-0.28	-0.15	0.48	0.34	0.43	0.35	0.27	-0.07	0.48	0.67	0.53	0.57	0.56	1.00				
17	-0.07	-0.02	-0.34	-0.21	0.50	0.33	0.46	0.38	0.34	0.02	0.47	0.55	0.62	0.47	0.44	0.67	1.00			
18	-0.13	0.04	-0.21	-0.14	0.51	0.33	0.44	0.38	0.35	-0.01	0.51	0.61	0.57	0.52	0.57	0.68	0.70	1.00		
19	-0.02	0.07	-0.21	-0.04	0.41	0.29	0.37	0.29	0.31	-0.01	0.37	0.51	0.45	0.60	0.64	0.62	0.51	0.66	1.00	
20	-0.04	-0.01	-0.17	-0.11	0.39	0.28	0.28	0.25	0.31	-0.10	0.38	0.55	0.42	0.59	0.64	0.49	0.45	0.50	0.60	1.00

Note: |Correlations| ≥ 0.20 are shown in **bold**. Those |correlations| ≥ 0.50 are shown in **bold italics**.

Table 3: Factor Analysis (Maximum Likelihood Extraction with Varimax Rotation)

	F1	F2	F3	F4
<i>Which, if any, of the following would you take into consideration when you were buying it?</i>				
Appearance/style	-0.065	-0.090	-0.093	0.343
Availability	0.095	0.094	0.108	0.806
Brand	-0.191	-0.246	-0.168	0.305
Quality	-0.104	0.185	0.236	0.379
People paid enough to live on	0.324	0.733	0.277	-0.073
Little damage to environment	0.130	0.739	0.091	-0.045
Work environment healthy	0.220	0.816	0.204	0.007
No animal testing	0.249	0.551	0.147	0.016
Human rights record of country	0.227	0.522	0.236	0.016
Need for the product	0.125	-0.192	-0.212	0.200
<i>Which, if any, of the following things about the people who made the product would affect your decision to buy it?</i>				
Forced to work overtime	0.560	0.196	0.146	-0.063
Did not earn enough wages to live off	0.568	0.228	0.464	0.037
Had no job security	0.752	0.109	0.096	-0.055
Could be sacked if they became pregnant	0.304	0.223	0.701	-0.006
Found their health to be in danger	0.258	0.236	0.763	0.009
Had no holidays or days off	0.644	0.282	0.425	-0.015
Were not allowed to join a union	0.739	0.293	0.193	-0.106
Had no right to sick pay	0.694	0.276	0.392	-0.013
Were subject to discrimination or harassment.	0.410	0.177	0.545	-0.021
Were under the legal minimum age to work	0.364	0.188	0.569	-0.020
Eigenvalue	7.428	1.715	1.663	1.305
Percent of variance explained	37.138	8.577	8.315	6.526

Descriptors:

- F1: Worker Rights
- F2: General Rights
- F3: Health & Harassment
- F4: Utilitarian

Table 4: Correlation Matrix of Factor Scores with Demographics

	F1	F2	F3	F4	EI	MR	M
F1: Worker Rights	1.000						
F2: General Rights	0.068	1.000					
F3: Health & Harassment	0.136	0.071	1.000				
F4: Utilitarian	0.026	-0.006	0.025	1.000			
Ethical Idealism (EI)	0.166	0.104	0.046	-0.121	1.000		
Moral Relativism (MR)	-0.297	-0.350	-0.245	-0.205	0.016	1.000	
Machiavellianism (M)	-0.050	-0.451	-0.150	-0.097	-0.259	0.303	1.000
Sample: MBA	-0.372	-0.260	0.182	0.326	-0.410	0.128	0.260
Sample: Hong Kong	-0.019	0.368	0.478	0.419	-0.227	-0.301	-0.177
Sample: Amnesty Intl	0.341	0.616	0.269	0.098	0.173	-0.421	-0.426
Gender (Female)	0.244	0.107	-0.203	-0.067	0.339	-0.184	-0.304
Ethnicity: White	0.153	0.547	0.332	0.165	-0.048	-0.464	0.350
Ethnicity: Chinese	-0.014	-0.247	-0.457	-0.361	0.196	0.263	0.111
Income (\$000)	0.030	0.153	0.172	0.156	-0.017	-0.189	-0.173
Age (Years)	0.146	0.424	0.188	0.154	-0.097	-0.506	-0.397
Education (University)	0.110	-0.076	-0.037	-0.244	0.035	-0.178	0.111

Note: |Correlations| > 0.20 are shown in **bold**. Those |correlations| > 0.50 are shown in **bold italics**.

Table 5: Cluster Grouping Comparisons^a

	C1	C2	C3	C4	C5	C6	Mean
Mean Factor Score							
F1: Worker Rights	0.126	-0.220	0.847	-0.246	0.183	-1.032	
F2: General Rights	-0.767	-0.444	0.889	-0.804	0.928	0.300	
F3: Health & Harassment	-1.041	0.503	0.168	0.528	0.430	-1.475	
F4: Utilitarian	-0.485	1.166	0.955	-0.557	-0.752	0.459	
Mean EDS Scores^b							
Ethical Idealism	3.56	3.38	4.03	3.67	3.49	3.36	3.56
Moral Relativism	3.29	2.89	2.34	3.33	2.70	3.34	3.05
Machiavellianism	98	95	92	97	92	93	95
Demographics^c							
Gender (% Female)	77	46	70	38	47	36	53
Chinese (Percent)	83	17	10	46	6	27	39
White (Percent)	7	46	90	33	88	45	43
Age (Years)	26.0	31.4	45.0	30.0	39.9	36.8	32.6
Income (\$000)	35.8	46.8	47.0	39.0	50.3	43.3	43.3
University Degree (Percent)	10	27	40	33	76	45	29
Percent of AI Supporters	8	16	18	16	34	9	
Percent Australian MBAs	11	42	3	31	3	11	
Percent HK Students	64	3	3	23	3	5	
Observations	41	41	31	45	56	21	

^aCluster identifiers: (C1) ethically uninterested, (C2) utilitarians with a (minor) conscience, (C3) social utilitarians, (C4) anti-exploitationists, (C5) anti-utilitarians, and (C6) liberal utilitarians.

^bNumber of possible mean difference comparisons = 15. Number of significant mean differences:

Ethical Idealism	1
Moral Relativism	4
Machiavellianism	0

^cNumber of possible mean difference comparisons = 15. Number of significant mean differences:

χ^2 test for dichotomous variables:

Income	0
Age	2

	χ^2	p
Female (df=5)	12.18	0.32
Chinese (df=5)	42.37	0.00
White (df=5)	40.36	0.00
University Degree (df=5)	10.57	0.63
Sample (df=10)	110.53	0.00

Table 6: WTP for Ethical Features Relative to WTP for Functional Features

	C1	C2	C3	C4	C5	C6	Mean
Athletic Shoes^a							
Child Labor	0.12	0.14	0.16	0.13	0.16	0.13	0.14
Minimum Wages	0.12	0.13	0.14	0.13	0.13	0.13	0.13
Working Conditions	0.12	0.11	0.11	0.11	0.09	0.11	0.11
Living Conditions	0.12	0.13	0.13	0.12	0.14	0.12	0.13
Bath Soaps^b							
Biodegradability	0.12	0.13	0.13	0.12	0.13	0.11	0.13
Animal Testing	0.09	0.10	0.09	0.10	0.08	0.10	0.09
Animal By Products	0.10	0.12	0.11	0.10	0.10	0.11	0.10

^aAverage WTP for Functional Features is \$60.58 including brand name and \$44.74 excluding the brand name.
 Number of possible mean difference comparisons = 15. Number of significant mean differences:

Child Labor	0
Minimum Wages	0
Working Conditions	2
Living Conditions	4

^bAverage WTP for Functional Features is \$4.63 for a packet of four bars of soap.
 Number of possible mean difference comparisons = 15. Number of significant mean differences:

Biodegradability	0
Animal Testing	0
Animal By Products	0

Table 7: Correlation Matrix of Absolute WTP with Mori Poll Responses (Athletic Shoe Sample)

	Functional Attributes	Child Labor	Minimum Wage	Working Conditions	Living Conditions
<i>Willingness to Pay</i>					
Functional Attributes	1.00				
Child Labour	-0.11	1.00			
Minimum Wage	-0.04	0.08	1.00		
Working Conditions	0.31	-0.17	0.07	1.00	
Living Conditions	0.00	0.07	0.23	-0.23	1.00
<i>Which, if any, of the following would you take into consideration when you were buying it?</i>					
Appearance/style	-0.21	-0.30	0.12	0.05	-0.11
Availability	-0.05	0.15	0.07	-0.08	0.14
Brand	0.04	-0.15	0.08	0.10	0.06
Quality	0.12	0.37	0.00	-0.02	0.13
People paid enough to live on	-0.02	0.45	0.08	-0.29	0.39
Little damage to environment	-0.05	0.19	0.25	-0.32	0.18
Work environment healthy	0.00	0.35	0.25	-0.41	0.44
No animal testing	-0.06	0.19	0.06	-0.27	0.28
Human rights record of country	-0.02	0.47	0.24	-0.06	0.22
Need for the product	0.17	-0.15	0.00	0.03	0.04
<i>Which, if any, of the following things about the people who made the product would affect your decision to buy it?</i>					
Forced to work overtime	-0.19	0.16	-0.11	-0.18	0.24
Did not earn enough wages to live off	-0.15	0.33	0.11	-0.22	0.40
Had no job security	-0.07	0.10	0.15	-0.14	0.19
Could be sacked if they became pregnant	-0.11	0.27	0.11	-0.33	0.31
Found their health to be in danger	-0.08	0.32	0.04	-0.41	0.34
Had no holidays or days off	-0.13	0.32	0.10	-0.34	0.38
Were not allowed to join a union	-0.14	0.22	0.04	-0.21	0.26
Had no right to sick pay	-0.18	0.39	0.10	-0.35	0.43
Were subject to discrimination or harassment.	-0.08	0.19	0.13	-0.27	0.31
Were under the legal minimum age to work	0.02	0.34	0.23	-0.14	0.30

Note: |Correlations| ≥ 0.20 are shown in **bold**. Those |correlations| ≥ 0.50 are shown in **bold italics**.

Items shaded indicate areas where the variables should be tapping the same underlying effect.

Table 8: Correlation Matrix of Absolute WTP with Mori Poll Responses (Bath Soap Sample)

	Functional Attributes	Biodegradability	Animal Testing	Animal Byproducts
<i>Willingness to Pay</i>				
Functional Attributes	1.00			
Biodegradability	0.10	1.00		
Animal Testing	-0.18	-0.20	1.00	
Animal Byproducts	-0.17	-0.10	0.23	1.00
<i>Which, if any, of the following would you take into consideration when you were buying it?</i>				
Appearance/style	0.06	-0.02	0.05	0.18
Availability	-0.06	0.05	0.18	0.20
Quality	-0.11	-0.15	0.33	0.21
People paid enough to live on	0.19	0.24	-0.41	-0.11
Little damage to environment	0.27	0.22	-0.36	-0.05
Work environment healthy	0.12	0.15	-0.23	0.03
No animal testing	0.18	0.09	-0.49	-0.15
Human rights record of country	0.04	0.28	-0.21	-0.11
Need for the product	0.09	0.11	0.12	0.04
<i>Which, if any, of the following things about the people who made the product would affect your decision to buy it?</i>				
Forced to work overtime	0.12	0.14	-0.28	-0.06
Did not earn enough wages to live off	-0.08	0.08	-0.44	-0.09
Had no job security	0.19	0.24	-0.36	-0.09
Could be sacked if they became pregnant	-0.06	0.13	-0.19	-0.10
Found their health to be in danger	0.04	0.13	-0.22	-0.09
Had no holidays or days off	0.06	0.23	-0.39	0.02
Were not allowed to join a union	0.14	0.27	-0.47	0.03
Had no right to sick pay	0.13	0.29	-0.30	-0.07
Were subject to discrimination or harassment.	0.12	0.12	-0.27	-0.07
Were under the legal minimum age to work	0.04	0.07	-0.14	0.02

Note: |Correlations| ≥ 0.20 are shown in **bold**. Those |correlations| ≥ 0.50 are shown in **bold italics**.

Items shaded indicate areas where the variables should be tapping the same underlying effect.

Table 9: Significance of GLM Estimates of Absolute WTP Estimates Coefficients versus Factor Scores

Dependent Variables	Factors				R-Squared
	F1: Worker Rights	F2: General Rights	F3: Health & Harassment	F4: Utilitarian	
Athletic Shoes					
Child Labor	0.296	0.003	0.009	0.585	0.182
Minimum Wage	0.761	<i>0.070</i>	0.791	0.504	0.003
Work Conditions	0.297	0.009	0.012	0.800	0.143
Living Conditions	0.025	0.008	0.023	0.453	0.186
Functional Features	0.132	0.781	0.734	0.882	0.017
Bath Soaps					
Biodegradability	0.033	0.145	0.657	0.748	0.068
Animal Testing	0.000	0.037	0.841	<i>0.067</i>	0.309
Animal Byproducts	0.950	0.898	0.911	0.028	0.020
Functional Features	0.248	0.114	0.157	0.168	0.063

Note: Bold implies significance at $p < 0.05$; italics $p < 0.10$

References

- Al-Khatib, J. A., Stanton, A. D., and Rawwas, M. Y. A.: 2005, 'Ethical Segmentation of Consumers in Developing Countries: A Comparative Analysis', *International Marketing Review* 22(2), 225–246.
- Auger, P., Burke, P., Devinney, T. M., and Louviere, J. J.: 2003, 'What will consumers pay for social product features?' *Journal of Business Ethics* 42(3), 281–304.
- Boulstridge, E., and Carrigan, M.: 2000, 'Do Consumers Really Care about Corporate Responsibility? Highlighting the Attitude-Behavior Gap', *Journal of Communication Management* 4(4), 355–368.
- Box-Steffensmeier, J. M., Jacobson, G. C., and Grant, J. T.: 2000, 'Question Wording and the House Vote Choice: Some Experimental Evidence', *Public Opinion Quarterly* 64(3), 257–270.
- Calder, B.: 1982, 'The Concept of External Validity', *Journal of Consumer Research* 9(4), 240–244.
- Calder, B., Phillips, L., and Tybout, A.: 1981, 'Designing Research for Application', *Journal of Consumer Research* 8(3), 197–207.
- Carrigan, M., and Attala, A.: 2001, 'The Myth of the Ethical Consumer-Do Ethics Matter in Purchase Behaviour?' *Journal of Consumer Marketing* 18(7), 560–577.
- Carrigan, M., Szmigin, I., and Wright, J.: 2004, 'Shopping for a Better World? An Interpretive Study of the Potential for Ethical Consumption Within the Older Market', *Journal of Consumer Marketing* 21(6), 401–417.
- Carson, R. T., Groves, T., and Machina, M. J.: 2000, 'Incentive and Informational Properties of Preference Questions', *Unpublished Working Paper*, University of California, San Diego.
- Chen, C., Lee, S.-Y., and Stevenson, H. W.: 1995, 'Response Styles and Cross-Cultural Comparisons of Rating Scales Among East Asian and North American Students', *Psychological Science* 6(3), 170–175.
- Christie, R., and Geis, F. L.: 1970, *Studies in Machiavellianism* (Academic Press, New York).
- Crane, A.: 2001, 'Unpacking the Ethical Product', *Journal of Business Ethics* 30(4), 361–373.
- Crane, A., and Matten, D.: 2004, *Business Ethics: a European Perspective : Managing Corporate Citizenship and Sustainability in the Age of Globalization* (Oxford University Press, Oxford).
- Creyer, E. H., and Ross, W. T.: 1997, 'The Influence of Firm Behavior on Purchase Intention: Do Consumers Really Care About Business Ethics?' *Journal of Consumer Marketing* 14(6), 421–433.

- Elliott, K. A., and Freeman, R. B.: 2001, 'White Hats or Don Quixotes? Human Rights Vigilantes in the Global Economy', *National Bureau of Economic Research* (8102), Cambridge, MA.
- Forsyth, D. R.: 1980, 'A Taxonomy of Ethical Ideologies', *Journal of Personality and Social Psychology* 39(1), 175–184.
- Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., and Soutter, C. L.: 2000, 'Measuring Trust', *Quarterly Journal of Economics* 115(3), 811–846.
- Harrison, R.: 2003, 'Corporate Social Responsibility and the Consumer Movement', *Consumer Policy Review* 13(4), 127–131.
- Klein, J. G., Smith, N. C., and John, A.: 2004, 'Why We Boycott: Consumer Motivations for Boycott Participation', *Journal of Marketing* 68(3), 92–109.
- Kotler, P.: 1997, *Marketing Management: Analysis, Planning, Implementation, and Control* (Prentice Hall, Inc., Upper Saddle River, NJ).
- Louviere, J. J., Hensher, D. A., and Swait, J. D.: 2000, *Stated Choice Methods: Analysis and Applications* (Cambridge University Press, Cambridge, UK).
- Marymount-University. 1999. The Consumers and Sweatshops, Vol. 1999: Marymount University Center for Ethical Concerns.
- Mason, T.: 2000, 'The Importance of Being Ethical', *Marketing* 26 October, 27.
- Miller, A. S., and Mitamura, T.: 2003, 'Are Surveys on Trust Trustworthy?' *Social Psychology Quarterly* 66(1), 62–70.
- Paulhus, D. L.: 1991, 'Measurement and Control of Response Bias', in J. P. Robinson, P. R. Shaver, and L. S. Wright (Eds.), *Measures of Personality and Social Psychological Attitudes* (Academic Press, San Diego, CA), pp. 17–59.
- Rogers, D.: 1998, 'Ethical Tactics Arouse Public Doubt', *Marketing* (August 6), 12–13.
- Rucinski, D.: 1993, 'Rush to Judgment? Fast Reaction Polls in Anita Hill-Clarence Thomas Controversy', *Public Opinion Quarterly* 57(4), 575–592.
- Schwarz, N.: 1999, 'Self-Reports: How the Questions Shape the Answers', *American Psychologist* 54(2), 93–105.
- Schwarz, N.: 2003, 'Self-Reports in Consumer Research: The Challenge of Comparing Cohorts and Cultures', *Journal of Consumer Research* 29(4), 588–594.
- Schwarz, N., Grayson, C. E., and Knauper, B.: 1998, 'Formal Features of Rating Scales and the Interpretation of Question Meaning', *International Journal of Public Opinion Research* 10(2), 177–183.

- Schwarz, N., and Hippler, H. J.: 1995, 'Subsequent Questions May Influence Answers to Preceding Questions in Mail Surveys', *Public Opinion Quarterly* 59(1), 93–97.
- Schwarz, N., and Oyserman, D.: 2001, 'Asking Questions about Behavior: Cognition, Communication, and Questionnaire Construction', *American Journal of Evaluation* 22(2), 127–160.
- Simon, F. L.: 1995, 'Global Corporate Philanthropy: A Strategic Framework', *International Marketing Review* 12(4), 20–37.
- Spar, D. L., and La Mure, L. T.: 2003, 'The Power of Activism: Assessing the Impact of NGOs on Global Business', *California Management Review* 45(3), 78–101.
- Steenkamp, J.-B. E. M., and Ter Hofstede, F.: 2002, 'International Market Segmentation: Issues and Outlook', *International Journal of Research in Marketing* 19(3), 185–213.
- Strong, C.: 1996, 'Features Contributing to the Growth of Ethical Consumerism: A Preliminary Investigation', *Marketing Intelligence & Planning* 14(5), 5–13.
- Tourangeau, R., and Smith, T. W.: 1996, 'Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context', *Public Opinion Quarterly* 60(2), 275–304.
- Ulrich, P., and Sarasin, C.: 1995, *Facing Public Interest: The Ethical Challenge to Business Policy and Corporate Communications* (Kluwer Academic Publishers, Boston).
- Uusitalo, O., and Oksanen, R.: 2004, 'Ethical Consumerism: A View from Finland', *International Journal of Consumer Studies* 28(3), 214–221.
- Zwick, R., and Chen, X.: 1999, 'What Price Fairness? A Bargaining Study', *Management Science* 45(6), 804–823.

Appendix A: MORI Poll

Please answer the questions given below by following the instructions given with each question.

1. If you were buying a product (for example a pair of shoes, clothes or fruit) that had been produced in a developing country, which, if any, of the following would you take into consideration when you were buying it? (Please, choose as many or as few as you like.)
 - Appearance/fashion/style or trend
 - Availability
 - Brand name
 - Quality
 - That the people who had produced it were paid enough money to live on
 - That it caused as little damage as possible to the environment and that its production processes are environmentally-friendly
 - That the people who produced it worked in an environment that did not affect their health
 - That the product had not been tested on animals (or had not used new ingredients tested on animals)
 - The human rights record of the country of origin
 - Your need (for buying it)
 - None of the above
 - Don't know

2. Do you think companies should have a minimum agreed standard of labour conditions (for example on health, safety, pay and working hours) for their workers in developing countries? (Tick the appropriate box.)
 - Yes
 - No
 - Don't know

3. If you were buying a product that had been made in a developing country, which, if any, of the following things about the people who made the product would affect your decision to buy it? Would you still buy it, still consider buying it but may not, would it make no difference to you either way, or would you definitely not buy it? (Circle the appropriate number, 1 to 5—note this is left off here)

If the people who produced it:

- (a) Were forced to work overtime
- (b) Did not earn enough wages to live off
- (c) Had no job security
- (d) Could be sacked if they became pregnant, and were subjected to forced pregnancy testing
- (e) Found their health to be in danger
- (f) Had no holidays or days off
- (g) Were not allowed to join a union
- (h) Had no right to sick pay
- (i) Were subject to discrimination or harassment
- (j) Were under the legal minimum age to work

Appendix B: Hypothetical Product Profile for Athletic Shoes

Full Profile

In this section, we will present you with thirty-two sets of *hypothetical*, that is, *conceptual or imaginary*, athletic shoes using the same 14 features as in the previous section. For each pair of athletic shoes, we would like you to:

1. indicate (by ticking the yes or no box) if you would **consider trying** the athletic shoes if they were available in your local stores today. That is, we would like you to tell us if the athletic shoes described (in each table) are attractive enough for you to consider them (i.e., include them in your list of possible choices) or if they are simply not worth considering.
2. indicate if you would **buy** the athletic shoes **instead of or in addition** to your current athletic shoes the next time you purchase athletic shoes. That is, we would like you to tell us if you would be willing to buy the athletic shoes described (in each table) to replace or to complement your current athletic shoes.

Features of the Shoes	Features of Shoe #1
Shock absorption/cushioning	Low
Weight	Lighter
Ankle support	Low cut
Sole durability	Short
Breathability/ventilation	High
Fabrication materials	Synthetic (eg nylon)
Reflectivity at night	No
Comfort/fit	High
Is child labour used in making the product?	No
Are workers paid above minimum wage?	Yes
Are workers' working conditions dangerous?	No
Are workers' living conditions at the factory acceptable?	Yes
Brand of Shoe	Reebok
Price	\$100
1. If the shoes described above were available in your local shops now, would you consider trying them (Tick ONE box only)? <input type="checkbox"/> No <input type="checkbox"/> Yes	
2. If the shoes described above were available in your local shops now, would you buy them instead of or in addition to your current shoes next time you shop for shoes (Tick ONE box only)? <input type="checkbox"/> No <input type="checkbox"/> Yes	

Appendix C: Product Features and Ethical Attributes Used in the Experiment

Athletic Shoes	Soap
<i>Basic Product Features:</i>	
Shock absorption/cushioning (LOW or HIGH)	Shape (ROUNDED or SQUARE)
Weight (LIGHTER or HEAVIER)	Natural ingredients (NO or YES)
Ankle support (LOW CUT or HIGH CUT)	Scented (NO or YES)
Sole durability (SHORT or LONG)	Artificial colors (NO or YES)
Breathability/ventilation (LOW or HIGH)	Moisturizer (NO or YES)
Fabrication Materials (SYNTHETIC or LEATHER)	Anti-bacterial protection (NO or YES)
Reflectivity at night (NO or YES)	Will it clog your pores? (NO or YES)
Comfort/fit (LOW or HIGH)	Will it worsen your acne? (NO or YES)
Brand of shoe (Nike, Adidas, Reebok, New Balance, Converse, Brooks, Fila, Puma, Etonic, Asics, Saucony)	Brand name (MAJOR MULTI-NATIONAL or LOCAL BRAND)
Price (\$40, \$70, \$100, \$130)—in Australia	Price (\$2.25, \$1.65, \$1.05, \$0.45)—in Australia
Price (\$300, \$550, \$800, \$1,050)—in HK	Price (\$6, \$9, \$12, \$15)—in HK

Ethical Features:

Is child labour used in making the product? (NO or YES)	Biodegradable formulation? (NO or YES)
Are workers paid above minimum wage? (NO or YES)	Tested on animals? (NO or YES)
Are workers' working conditions dangerous? (NO or YES)	Animal by-products used as ingredients? (NO or YES)
Are workers' living conditions at the factory acceptable? (NO or YES)	

Two level items coded as -1 and +1. The first item (e.g., NO) is coded -1 and the second (e.g., YES) is coded +1