Paper to be presented at the Health Economics Study Group, University of Leeds, January 2003

Work in progress. Please do not quote or cite

Is one reason enough?

Comparing compensatory and non-compensatory decision-making strategies in discrete choice experiments

Kathy Porter*, Mandy Ryan and Mabel Amaya-Amaya

Health Economics Research Unit, University of Aberdeen, UK.

Abstract

Discrete choice theory assumes that when presented with alternative scenarios, individuals adopt a fully compensatory utility maximising decision-making process. That is, they are assumed to weight and "trade-off" the levels of all the attributes and, based on these overall evaluations, choose the highest utility scenario. Limited research has explored whether this assumption can be empirically falsified. Extensive work in psychology supports the idea that the key to understanding choice behaviour lies in comprehending how decision-making strategies are well-matched to particular contexts. This paper explores this hypothesis using responses to three discrete choice experiments whose context differ with respect to complexity of the choice task, familiarity of respondent and incentive compatibility. Consideration is also given to the influence of socio-economic characteristics on decision strategies. For each dataset, predicted choices using the traditional compensatory regression analysis and an alternative lexicographic heuristic based on self-reported attribute importance rankings are compared to actual choices. It is hypothesised that more complex and/or unfamiliar choices and/or incentive-compatible experiments are more prone to trigger non-compensatory behaviour. The results obtained from this study provide suggestive evidence of context effects and indicate a need to further investigate the nature of the decision-making process in completing choice experiments.

JEL Classification: A12, D12, I1

Keywords: Discrete Choice Experiments, compensatory decision-making, non-compensatory rules, regression analysis, lexicographic rule, context effects

---

**Correspondence to*: Kathy Porter, Health Economics Research Unit, Department of Public Health, University Medical buildings, Foresterhill, Aberdeen, AB25 9ZD. Tel: 01224 552764, Fax: 01224 662994, http://www.abdn.ac.uk/heru/ Email: k.porter@abdn.ac.uk

Acknowledgements: Financial support from the University of Aberdeen is acknowledged. HERU is supported by the Chief Scientific Office of the Scottish Executive Health Department (SEHD). The views expressed here are those of the authors and not necessarily those of SEHD.

1. Introduction

Research in the area of health care benefits valuation has seen an increased interest in the use of discrete choice experiments (DCEs)[1] for eliciting preferences as shown by the explosion of literature applying this technique to direct evaluation of different policy relevant attributes of health care interventions (see Ryan and Gerard 2001 for a recent review). A critical assumption underlying discrete choice models is that the decision strategy leading to choice is **fully compensatory**. Decision-makers are assumed to evaluate ***all*** the available alternatives simultaneously and then select the option with the highest utility considering **trade-offs** among ***all*** the attributes. This embodies the notion of ***unlimited substitutability*** meaning that individuals are able and willing to exchange or substitute the bundle of attributes currently held for another and can do so without affecting their welfare level. This property is at the heart of economists' utility theory and their concept of value and it provides tractable models of of consumer choice (McFadden, 1986).

However, limited research has been conducted in Economics to verify the correspondence between choice decision processes and the formation of value trade-offs to reduce multiple attributes to a uni-dimensional utility criterion, as visualized in the standard economic model. Methodologically, it has proven difficult to evaluate the economic assumption of compensatory decision-making and observed "inconsistencies" in choice behavior have been typically taken to be the result of observational or modeling deficiencies on the part of the analyst (Ben-Akiva and Lerman, 1985).

By contrast, more behaviorally oriented research has devoted considerable attention to identifying and characterising the strategies used by human beings and organisations to make decisions (Bettman et al, 1991; Payne et al, 1993). The view of decision-making that emerges from this work is quite different from that which underlies most economic choice models. Instead of a single, context-free decision rule, individuals appear to possess an assortment of contingent decision heuristics. The heuristics employed in choice depend upon the external representation of the choice problem as well as individual characteristics, and vary over time as the structure of the choice set changes (e.g. Einhorn 1971; Liechtenstein and Slovic, 1971; Payne, 1976; Olshavsky, 1979; Bettman, 1979). In particular, consumers have been found to employ simplifying non-compensatory strategies to reduce cognitive requirements (Bettman et al, 1993) and to vary in the accuracy with which they make their choices and provide preference evaluations (Haaijer et al, 2000; Fischer et al, 2000). Therefore it has been proposed that consumers should be modeled as boundedly rational (see Rubinstein 1998 for a review)

[1] Other labels given to DCEs include Stated Preference Discrete Choice Modelling (SPDCE), Discrete Choice Modelling (DCM), Attribute Based Stated Choice Methods (ABSCM) and Discrete Choice Conjoint Analysis (DCCA)

Utility theory, as posed in Economics underlying choice experiments, and in fact, all preference elicitation techniques, is in contradiction with such a perspective. Thus the investigation of non-compensatory choices and rule-based behaviour may be of paramount importance for economic choice theory in general and valuation methods in particular. For example, if non-compensatory decision are indeed used by a significant proportion of the population, the use of Cost Benefit Analysis itself would be questionable to the extent that it is founded upon the concept of compensating for welfare losses, as expressed in the Kaldor-Hicks potential compensation test (Kaldor 1939, Hicks 1940).

This paper describes work in progress looking at the extent to which the assumption of compensatory decision-making underlying discrete choice analysis can be empirically falsified in a health care context. The predictive validity of two different modes of processing subjective information in DCEs is examined. The traditional compensatory random effects probit widely used in health economics to analyse responses to choice experiments is contrasted with a non compensatory, lexicographic heuristic based on self-reported attribute importance rankings to test the alternative theoretical constructs. The remainder of the paper is organised as follows. Section 2 provides background information on the traditional compensatory economic model of choice behaviour and alternative non-compensatory, context-dependent views emerging from empirical research in behavioral sciences such as cognitive psychology. Following this, section 3 describes the datasets used in the empirical work and the methods of statistcla anlysis. The data sets were chosen due to their variability with respect to complexity of the choice task (number of attributes and levels and number of questions per individual), familiarity (previous experience) and incentive compatibility. A number of hypothesis regarding "context effects" on decision-making are advanced. Consideration is also given to influence of the socio-economic characteristics on decision strategies. In Section 4 the results of the application of both models to the different datasets are presented and the prior hypotheses assesed. Section 5 discusses the results and some practical implications, and identifies future research areas.

2. Background

2.1. Compensatory choice models

DCEs present respondents with a series of hypothetical scenarios compound by two or more alternatives, which vary along several characteristics or attributes of interest, and ask them to choose one alternative. It is assumed that individuals derive utility from the underlying attributes rather than the commodity *per se* and that they reveal their preferences through their choices. The results from the experiment are used to model preferences within a Random Utility Maximisation (RUM) framework (McFadden, 1974). The standard tool for analysing responses to DCEs is McFadden's discrete choice model based on the Random Utility Maximisation hypothesis (McFadden, 1974). The idea behind random utility models is that researchers cannot observe all factors affecting preferences (represented by utilities). Therefore, the researcher can only make probabilistic statements about a decision-maker's choice. The probability that a utility maximising decision-maker *n* chooses alternative i is given by

$$
\begin{aligned}
P_{in} &= \Pr(U_{in} > U_{jn}) \\
&= \Pr(V_{in}(X^k_{in}, m_n, Z_n) + e_{in} > V_{jn}(X^k_{jn}, m_n, Z_n) + e_{jn} \\
&= \Pr(e_{jn} - e_{in} < V_{in}(X^k_{in}, m_n, Z_n) - V_{jn}(X^k_{jn}, m_n, Z_n)) \quad \forall j \neq i
\end{aligned}
\tag{1}
$$

Where the latent utilities U_{in} i=1,… (J alternatives) are considered decomposable in two additively separable parts: an explainable component $V_{in}(X^k_{in}, m_n, Z_n)$ specified as a function of the attributes and corresponding levels (k) of the alternatives (X^k_{in}) and the characteristics of the decision-maker n (Z_n), including income (m_n) and unmeasured variation in preferences e_{in}

In words, the probability that an individual n chooses alternative i over all other in the choice set is given by the probability that the error difference is smaller than the difference in the observed utilities. Different discrete choice models are obtained from different assumptions about the distribution of the unobserved portion of utility. Logit and probit models are commonly used to estimate the measurable component of the utility function.

A critical issue in discrete choice models is the specification of the function $V_{in}(X^k_{in}, m_n, Z_n)$ that relates the observed attributes of the alternatives and the characteristics of the decision-maker to the utility that the decision-maker derives from alternative i U_{in} (so-called "representative" utilities). A linear additive model is normally assumed

$$
V_{in} = ASC_i + b_1 x_1^k + \ldots + b_{J_n} x_{J_n}^k + g(m_n * ASC_n) + q_1(z_1 * ASC_1) + \ldots + q_m(z_m * ASC_m) \tag{2}
$$

where there are J_n attributes with generic coefficients across alternatives, and $(m+1)$ individual-specific socio-economic and attitudinal variables interacted with alternative specific- constants (ASC). An additional alternative-specific (ASC_i) captures the mean effect of the unobserved factors in the error terms for each alternative.

The implication of the above specification is that individuals are supposed to use a weighted additive decision rule, which presumes explicitly considering the extent to which one is willing to **trade-**

off *all* relevant attribute values, as reflected by the relative importance's (or weights). That is, the decision strategy leading to choice is assumed a **fully compensatory** one, whereby decision-makers are assumed to evaluate ***all*** the available alternatives simultaneously. For each alternative, an overall evaluation is developed as the result of multiplying the weight times the value for each attribute and summing these weighted attribute values over all attributes. It is then assumed that individuals, considering **trade-offs** among ***all*** the attributes, select the alternative with the highest overall evaluation.

Compensatory models embody the notion of *unlimited substitutability* meaning that individuals are able and willing to exchange or substitute the bundle of attributes currently held for another and can do so without affecting their welfare level. More of one attribute (e.g. shorter waiting-time) can always compensate for less of another (e.g. less income). This assumption can be shown equivalent to the Archimedian (and related continuity) axiom (Borsch, 1968), hence enabling preferences to be represented by a well-behaved utility function and indifference curves to be constructed. Hence this property lies at the heart of economists' concept of value since it allows estimating the trade offs people are prepared to make between attributes (i.e. marginal rate of substitutions) and the welfare effects of changes in individual attributes and/or the good or service in question.

Some evidence of people potentially acting in the denial of the principle of *unlimited substitutability* has been put forward in different economic fields such as transportation or marketing research and more recently environmental and health economics. For relevant reasons, the review that follows deliberately concentrates on this latter[2]. Here, evidence of some type of non-compensatory behaviour has been reported in most applications of choice experiments to health care valuation (Ryan and Gerard, 2002). This has been investigated by directly asking respondents trade-off questions (Propper (1995) or identifying individuals who always choose the 'best' level of a particular attribute or the same alternative (Bryan et al, 1997; Ratcliffe, 1998; Ryan and Hughes, 1997; McIntosh and Ryan, 2002). A more recent study used self-reported ranking of attribute importance to further refine the identification of dominant preferences (Scott, 2002). In the general Economics literature several writers and researchers having investigated the concept of hierarchical choice and limited substitutability in Economics for more than a century (e.g. Menger, 1871; 1950; Little, 1957; Georgescu-Roegen, 1958; Drakopoulos, 1994). Despite all this, the compensatory model still dominates in Economics.

It may be argued that evidence remains limited in that the researcher cannot really know what respondents to DCEs are doing. However preliminary findings from studies applying qualitative research methods, e.g. Verbal Protocol Analysis (Ericsson and Simon, 1993) where the respondent is asked to

[2] See Lee-Gosselin and Pas (1997) for a discussion of non-compensatory choices in transportation; Bettman et al (1998) for evidence in marketing research and Spash (2000) for a review of studies in environmental economics.

think aloud while making a choice, support the idea that respondents tend to focus on one or two attributes to make a decision (Ryan et al, 2002) [3].

2.2. Non compensatory choice models

In contrast to economists, psychologists and other social scientists have conducted a fairly large number of studies with the explicit purpose of identifying the behavioral mechanisms involved in human decision-making. These researchers have developed a number of models, which not only describe behaviour but also attempt to explain the processes by which choices are made (e.g. Kleinmuntz, 1968; Russ, 1972). The unifying goal of this research has been to arrive at "an adequate characterization of a presumed cognitive process by which the identifiable characteristics of objects become synthesized, eventuating in an observable response" (Hoffman, 1960)

The results of these behavioral studies firmly demonstrate the extensive use of non-compensatory choice strategies. The general conclusion is that "subjects find it rather difficult to weight and trade off values in a compensatory manner … information that does not require "in the head' transformation is preferred in the interest of cognitive economy" (Slovic and MacPhillamy, 1974).

Non-compensatory choice strategies come from a number of disciplines and have been described using very different kinds of formalisms. Payne et al (1993) provide a detailed review of these. The work herein concentrates on a lexicographic type rule, which processes experimental data sequentially to allow for the evaluation of alternatives in accordance to the most important attribute first. Here individuals will evaluate the alternatives presented in the choice set in terms of their attributes proceeding in a strictly sequential manner, from the attributes that the individual considers the most important to the attribute that the individual considers least important according to the individual's stated importance ranking (Brunswick, 1955). If the most highly ranked attribute discriminates between options, further search is halted and a decision made. If the level for this attribute is the same it is assumed the individual will consider the second best ranked attribute. Again, if the level differs between option a decision will be made based on this attribute; if not, further attributes will be considered until a decision can be arrived at[4]. This sequential consideration principle pervades the human infomation processing theory of Newell and Simon (1972) and the mathematical probabilistic choice theory referred to as "Elimination by Aspects" (Tversky ,1972a). It also appears in a predecessor of the lexicographic semiorder of Tversky (1972b). The

[3] Yet issues are raised about the extent to which it is easy to verbalize thoughts. Empirical evidence has shown that some of the decision rules applied are used subconsciously and cannot easily be communicated (see e.g. Bem (1972); Nisbett and Wilson (1977). This, however, does not invalidate the use of a plurality of decision rules (White (1980))

[4] If it were the case that none of the attributes discriminate between options, the individual would be randomly assigned a choice as it is assumed s/he just guesses. However, experimental design techniques should ensure this situation will not happen.

principle has been supported by empirical findings in a number of brand choice market research studies (Russ, 1971)

Since non-compensatory models do not involve explicit trade-offs, it is not possible to state exactly which improvement in one attribute of a choice alternative just compensates for a given deterioration in another attribute to keep the respective choice alternative equally attractive. As a result, calculation of welfare estimates are complicated and even the use of Cost Benefit Analysis, based on potential compensation tests, itself would be questionable.

Because of this, non-compensatory models may immediately arouse the hostility of a conventional choice theorist and they may appear unreasonable. Yet they are seen as psychologically more plausible than fully compensatory rules embodied in economic choice models as they are more compatible with humans' cognitive limitations (Kahneman, 1973; Newell and Simon, 1972) and flexible use of information. They may better capture how real minds make decisions under constraints of limited time, knowledge and computational resource.

2.3 The importance of context

Extensive work in psychology and decision theory supports the idea that the key to understanding choice behaviour lies in comprehending how decision-making strategies are well-matched to particular contexts. Ford et al (1989) provide a comprehensive analysis of the research using process tracing approaches to study decision behaviour, looking at the evidence of the impact of the task complexity, environmental and person factors on decision strategy selection.

Empirical applications of choice experiments require respondents to understand in general terms the attributes of options, the way those attributes may vary across a number of levels and the way various combinations of attributes at varying levels may result from alternative resource use options under consideration. They also therefore require respondents to make a number of choices between multiple alternatives. This complexity and cognitive burden of stated choice tasks, in conjunction with the limited abilities of respondents, may give rise to the use of simplified non-compensatory decision strategies and any of a number of more subtle effects such status quo bias due to fatigue effects (Samuelson and Zeckhauser, 1988). In fact, the impact of task complexity on behavioral outputs from choice experiments is an emerging research issue (Opaluch and Mazzota, 1995; Swait and Adamovicz, 2001b; Deshazo and Fermo, 2002).

3. Empirical work: a comparison of compensatory and non-compensatory choice strategies.

3.1 Datasets

To make comparisons between the two models of information processing described above three existing DCEs which also collected information on individuals ranking of attributes are used. The choice of datasets was made with consideration given to a variety of factors potentially influencing decision strategy selection. These include the complexity of the choice task, incentive compatibility and familiarity (experience) with the good or service in question. Following evidence from psychology, it was hypothesised that the more complex the task, the more likely individuals will resort to simplified non-compensatory choice strategies to reduce the cognitive burden of arriving at a choice. Herein two dimensions of the task complexity are considered - number of attributes and number of levels.

Incentive compatibility refers to extent to which we can expect respondents to answer truthfully, in particular when cost is included as an attribute. In this latter case, it is common for a disclaimer to be included in the questionnaire to make it clear that the questions are for research purposes only and there is no way the individual would have to pay any amount of money. This may lead respondents to over state their value of the service rather than revealing the truth. Lack of incentives may cause respondents to adopt non-compensatory decision strategies.

Finally it may be easier for individuals to make trade-offs in a compensatory manner if they are familiar with the good being valued, for example, if they had previous experience so that information provided can be easily understood and processed with relatively low cognitive effort.

Three datasets selected for the analysis are (which all include ranking questions)[5]:

1) A study looking at GPs preferences for jobs characteristics (Scott, 2001) (GP hereafter)
2) Responses to a DCE concerned with preferences for cervical screening tests (Ryan and Wordsworth, 2000) (CS)
3) A DCE conducted alongside a randomised control trial to assess the introduction of a new repeat prescription system (Ryan and Ubach, 2002) (RP)

Table 1 shows the attributes and levels for all three studies. Complexity varies across the datasets. The GP study has seven attributes, with levels ranging between two and three; the CS study has six attributes with three to four levels and the RP study has three attributes with levels ranging between two and four. In the CS study complexity may increased by the inclusion of three risk attributes (chance of

[5] In-depth discussion of these data is given in the references provided to which the reader is directed for further details.

abnormality; chance of being recalled and chance of dying from cervical cancer). The psychology literature indicates that individuals find it difficult to handle risk information (Calman, 1996; Lloyd, 2001)

Table 1. Attributes and levels for all three studies

Attributes	**Levels of each attribute**	
1. Repeated prescriptions		
Convenience of ordering and collecting your repeat medicines	• You attend the surgery and visit the pharmacy • You visit the pharmacy only	
Cost of ordering and collecting your repeated prescription	• Nil; 25p per month; £1 per month; £2 per month	
Quality of advice when collecting your repeat prescription	• No advice • Spoken directions • Spoken directions with discussion of any medicine problems	
2. Cervical screening	**Questionnaire 1**	**Questionnaire 2**
Time between smears	• 1,3,5 years	• 1,3,5 year
Time for results	• 7, 10, 21, 28 days	• 5, 10, 18, 28 days
Chance of being recalled	• 11, 15, 17, 20 %	• 11,15, 17, 20%
Chance of abnormality	• 3,5, 8, 10%	• 3,5,8,10%
Chance of dying from cervical cancer	• 0.5, 0.8, 1.5, 2 %	• 0.4, 0.8,1.3, 2 %
Cost of each smear	• £2,£8,£20, £35	• £7,£30,£40,£60
3. GPs preferences for job characteristics		
Opportunities to develop special interests	• Yes/ No	
List size per full-time equivalent GP	• 1400, 2000, 2600 patients	
Total daytime hours worked per week	• 35, 40, 45 hours	
Change in income per year (after tax)	• £2500 less, No change ; £2500 more	
Time spent on administration per week	• 7, 10,13 hours	
Out of hours arrangement	• No nights or weekends on call • One night per week and one week in six • Two nights per week and one week in four	
Use of guidelines for certain areas of care	• Yes/No	

Familiarity also varies across the studies. Whilst it may be hypothesized that GPs would be familiar with all attributes presented in the DCE, women may only be familiar with the frequency of screening for cervical cancer and only a patient on the intervention arm of the randomised control trial would be familiar with both methods of repeat prescription.

Finally, regarding incentive compatibility, this is hypothesised high for the GPs study as there are no particular reasons to suspect that GPs should be anything but truthful. However, the CS study may not be incentive compatible due to the inclusion of a cost attribute and the corresponding disclaimer mentioned above. The opposite is true of the RP study. Although cost is included this is *not* the direct cost of a prescription but rather the indirect costs that individuals bare in the process. Therefore incentive compatibility is thought to be good.

Given the characteristics of the data-sets ,the following hypotheses are advanced:

H_1: ***For the RP study, it is expected that the compensatory regression model would predict better than the non-compensatory*** This seems reasonable a prior given not only the level of complexity is low (three attributes) but also the task was both familiar to respondents and incentive compatible.

H_2: ***In the CS study, it would be expected that the non-compensatory model would outperform the compensatory regression model.*** The inclusion of several risk attributes as percentages and the lack of incentive compatibility justify this prediction.

H_3: ***In GPs preferences for job characteristics study either the compensatory or non-compensatory models will perform better.*** A fairly high degree of familiarity and high incentive compatibility suggested the compensatory model would fit better. However, the choice task is complex (seven attributes) and a time constraint could be operating, which may or may not outweigh familiarity and incentive compatibility

Table 2 provides a summary of the previous discussion along with the hypotheses made for each study.

Table 2 Context of studies

STUDIES	No. attributes	No. levels	Familiarity	Incentive Compatibility	**Hypothesis**
1. Repeated prescriptions	7	2-3	Yes	Yes	**Lexi or Comp**
2. Cervical screening	6	3-4	No	No	**Lexi >Comp**
3. GPs preferences for job	3	2-4	Yes	Yes	**Comp> Lexi**

3.2. Statistical analysis

The data sets are first analysed using the standard compensatory regression model. Using the random effects probit model, equation (2) is estimated. From this model predicted choices are estimated -– here it is assumed that a predicted choice probability greater than 0.5 for an alternative would lead the individual to choose it. Following this, Cohen's κ coefficient of gross agreement corrected for chance (Cohen, 1968) is used to assess the correlation between actual choices and those predicted by the estimated regression model This summary measure ranges between -1 and +1, with 0 being no agreement beyond that expected by chance, 1 being complete agreement, and -1 being contrary to agreement.[6]

Subsequently datasets are "re-analysed" assuming that the lexicographic rule was employed. For each individual and each choice a predicted choice is worked out from their ranking responses. Three steps were followed here:

STEP 1: Choose the attribute with the highest ranking that has not yet been looked at. Look up attribute values of the two scenarios.

STEP 2: If one scenario has a higher value for given attribute then stop and go to Step 3. Otherwise go back to Step 1 and look at next highest ranked attribute.

STEP 3: If scenario with the highest value of said attribute is also respondent's actual choice of scenario record a 'hit', if not, then a 'miss' is recorded.

As an example, assume the ranking obtained from an individual in the RP study indicates that the most important attribute is convenience, followed by the quality of advice and finally the cost. Then, when the first ranked attribute is the same across choices the lexicographic rule cannot differentiate. The rule would then move to the second ranked attribute to see if they differentiate. The search is halted if there is a difference, and choice can be predicted. A record is kept of the number of attributes employed to arrive at a choice (so-called frugality). Only individuals who correctly ranked all the attributes in the questionnaire are used for the analysis. As with the regression analysis, actual choices are compared with the predictions from the non-compensatory model using Cohen's κ coefficient (Cohen, 1968).

Following this, measures of agreement for the compensatory regression model and non-compensatory lexicographic models are compared. Goodness of fit in terms of proportion of correctly predicted choices by the two decision-making strategies are also compared. Consideration is finally given to influence of the socio-economic characteristics (age and income) and particular features of studies (time of completion and ease of completion) on decision strategies.

[6] Usually it is considered that a value of κ less than zero implies worse than chance agreement, and <0.2 implies poor agreement, 0.21 – 0.4 fair agreement; 0.41 – 0.60 moderate agreement; 0.61 – 0.80 good agreement and 0.81 to 1.00 very good agreement.

4. Results

Table 3 presents the estimation results from the random effects probit model for all three studies. The signs on the coefficients are all as expected. Assuming compensatory decision-making, the ratio of the coefficients indicates the marginal rate of substitution between attributes (Hensher and Johnson, 1981). Thus, it is possible to estimate a monetary value for all attributes. Such values are useful at the policy level. For example, other things equal, individuals in the cervical screening study were prepared to pay £51.07 for a 1% reduction in the chance of dying and respondents in the GP study would give up £2,386 to have opportunities to develop special interests, and would require a compensation of £9.49 per extra hour worked per year. All this would lead to conclude that the results are valid, to the extent that they seem consistent with economic theory.

Table 4 shows the predictive ability (or "goodness of fit") for both the compensatory regression model and the lexicographic rule in all three studies. Starting with the compensatory model, the random effects probit correctly predicts 78% of choices in the cervical screening study and 79% in the GPs study. In both cases the associated kappa coefficient indicates a moderate agreement between predicted and actual choices ($0.41<\kappa<0.6$). By contrast, in the repeat prescription study the regression model only predicted 67% of scenarios actually chosen, which is shown as a "fair agreement" ($\kappa=0.169$).

Moving onto the results from the hierarchical model, Figures 1 to 3 show the percentage of times an attribute was ranked first by individuals. As seen in Figure 1, over three quarters of individuals (79%) considered convenience of ordering and collecting repeat medicines as the most important attribute in a repeat prescription system; followed by the quality of advice at the collection point. "Cost " was ranked first only by 10% of respondents. In the CS study (Figure 2) the chance of dying from cervical cancer was seen as the most important attribute by more than 60% of individuals; followed by the time between smears, which 23% of women ranked first. Finally, Figure 3 for the GPs study. There are two prominent attributes, "out of hours arrangements", with 43% of GPs ranking this as the most important characteristic of their job and 'opportunity to develop special interests', which 30% of GPs felt the most important.

Table 3 Regression results

Explanatory variables[7]	Coefficient β_i	t-stat	MRS[a] Price proxy	First ranked (% people)
1. Repeated prescriptions				
Convenience of ordering and collecting your repeat medicines	0.177	2.78	21p for preferred method	79%
Cost of ordering and collecting repeat prescriptions	-.008	-3.84	-	10%
Quality of advice received when collecting your repeat prescription	0.307	1.93	37p for marginal improvement in advice	19%
Log likelihood (full) - LL(full) Log likelihood (cnst only) – LL(cnst) Pseudo R^2=(LL(cnst)-LL(full))/LL(cnst) 31 individuals sand 118 responses	-69.148 -79.575 0.131			
2. Cervical screening				
Time between smears (years)	-0.239	-9.81	£17.07 per year reduction	23%
Time for results (days)	-0.017	-4.77	£1.21 per day reduction	4%
Chance of being recalled (%)	-0.066	-7.8	£4.71 per 1% reduction	1%
Chance of abnormality (%)	-0.41	-6.25	£29.69 per 1% reduction	7%
Chance of dying from cervical cancer (%)	-0.715	-9.99	£51.07 per 1% reduction	62%
Cost of each smear(£)	-0.014	-8.89	-	3%
Log likelihood (full) - LL(full) Log likelihood (cnst only) – LL(cnst) Pseudo R^2=(LL(cnst)-LL(full))/LL(cnst) 509 individuals and 2528 responses	-1266 -1750.674 0.277			
3. GPs preferences for job charact.				
Constant[b]	0.167	2.78	-	-
Opportunities to develop special interests	0.231	2.28	£2386 for opportunity	30%
List size per full time equivalent GP	-0.0009	-8	-£9.40 per extra patient	15%
Total daytime hours worked per week	-.04	-2.98	-£9.49 per extra hour per year	5%
Change in income per year (after tax)	0.0001	3.26	-	6%
Time spent on administration per week	-0.024	-1.07	-£5.23 per extra hour per year	-
Out of hours arrangements	-1.185	-15.74	-£11850	43%
Use of guidelines for certain areas of care	0.360	3.3	£3717 to use guidelines	1%
Log likelihood (full) - LL(full) Log likelihood (cnst only) – LL(cnst) Pseudo R^2=(LL(cnst)-LL(full))/LL(cnst) 200 individuals and 835 responses	-370 -578.514 0.551			

a. Marginal rate of substitution β_i/β_{cost}. . *Ceteris paribus,* this is an estimate of the welfare impact of a change in the specific attribute when the choice set includes only a single before and after change option (i.e. it is assumed the changed scenario would be chosen with certainty)

b. Constant was kept only when significant

[7] The constant term is excluded when it is not significant. This follows the usual practice with non-significant variables. However it is fair to recognise this may be affecting the properties on the disturbances in the model, for example, zero mean cannot be guaranteed.

Table 4 Number of correct predictions ("*hits*") in sample

Study	No. of cases	Correct A (%)	Correct B (%)	Total (%)	κ value (asy. t-stat)
1. Repeated prescriptions					
Random effects probit (Prob>0.5)	118	36 (31%)	42 (36%)	78 (67%)	0.322 (3.52%)
Lexicographic rule		43 (36%)	53 (45%)	96 (81%)	0.627 (6.91%)
2. Cervical screening					
Random effects probit (Prob>0.5)	2528	875 (35%)	1101 (44%)	1976 (78%)	0.561 (28.38%)
Lexicographic rule		964 (38%)	951 (38%)	1915 (76%)	0.516 (26.02%)
3. GPs preferences for job charact.					
Random effects probit (Prob>0.5)	835	337 (40%)	324 (39%)	661 (79%)	0.583 (16.86%)
Lexicographic rule		330 (39%)	327 (39%)	657 (79%)	0.574 (16.60%)

In Figures 4-6 the number of attributes the lexicographic rule had to look at before making a 'hit', the so-called "frugality" of the model, is shown. In more than 60% of cases one attribute was sufficeint to halt search and make a choice and in 30% of cases two attributes sufficed. Whilst this clearly relates to the design of the experiment, to the extent that it ensures minimum overlap, it should be expected that a decision can be reached with one or two attributes. Nevertheless, this may cast doubt about the amount of information people actually use when completing discrete choice experiments.

The results from the non-compensatory model are as follows. As with the compensatory model, for the GP and CS studies there is a moderate agreement ($0.41<\kappa<0.6$) between choices predicted by the lexicographic decision rule and actual choices, with the non-compensatory model correctly predicting 79% and 76% of chosen scenarios respectively. However in this case the kappa coefficient ($\kappa=0.627$) indicates a good, rather than fair, agreement for the repeat prescriptions, in which the hierarchical model correctly predicts 81% of actual choices.

Table 5 provides a summary of the results "in context" so that the hypotheses postulated for each study (see Table 2) can be analysed. Beginning with RP study, the hypothesis (H_1) that the compensatory regression model would predict better than the non-compensatory was not found. Notably the hierarchical choice model had a much better "fit", predicting 81% of actual choices ($\kappa=0.627$) as compared with only 67% correct predictions achieved by the regression model ($\kappa=0.322$). This may be explained by the fact that in this study all ranking questions were asked at the beginning of the questionnaire. Respondents may be trying to be consistent when making their choices.

Figure 1 Repeat Prescriptions First Ranked Attribute

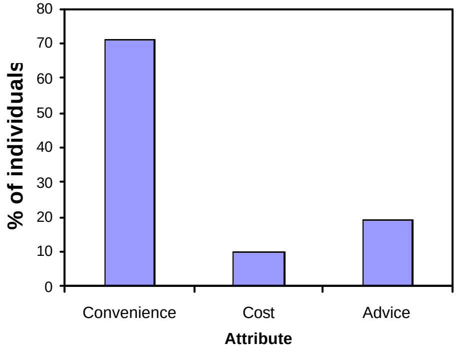


Figure 2 Cervical Screening First Ranked Attributes

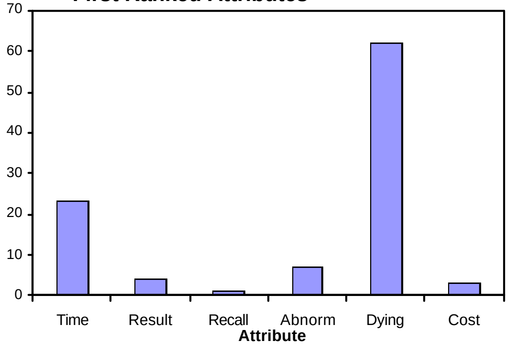


Figure 3 GPs' Job Preferences First Ranked Attribute

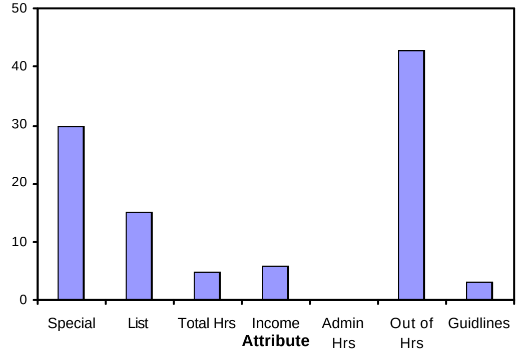


Figure 4 Repeat Prescriptions Frugality

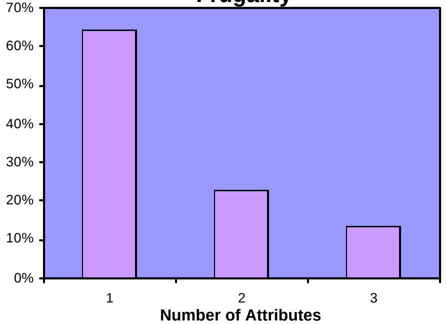


Figure 5 Cervical Screening Frugality

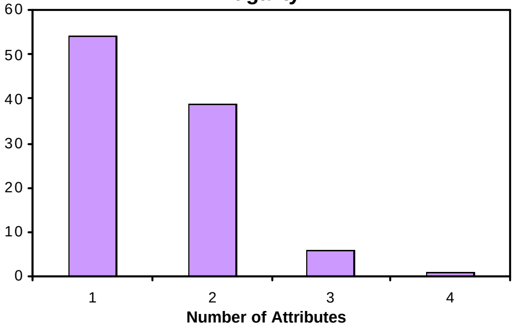
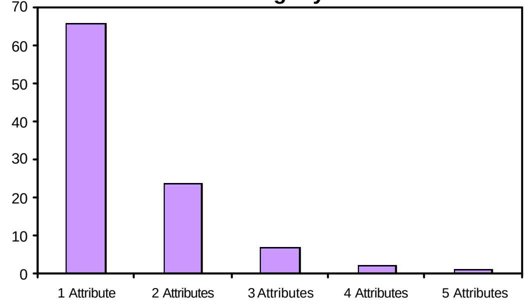


Figure 6 GPs' Job Preferences Frugality



Further, this study followed a randomised control trial, which may point to individuals having a dominant preference for the system they are accustomed to. Segmentation analysis carried out later may reveal some other reasons (e.g. age).

Table 5 Comparison of results "*in context*"

STUDY	No of attrib.	No of levels	Familiar.	Incentive compatib	Total correct predictions	κ value (asy. t-stat)
1. Repeated prescriptions	3	2-4	Yes	Yes	(REP) 67%	0.322 (3.52)
					(LEXI) 81%	0.627 (6.912)
2. Cervical screening	6	3-4	No	No	(REP) 78%	0.561 (28.38)
					(LEXI) 76%	0.516 (26.02)
3. GPs preferences for job chara	7	2-3	Yes	Yes	(REP) 79%	0.583 (16.86)
					(LEXI) 79%	0.574 (16.60)

Turning now to the CS data, whilst it was expected that the lexicographic would outperform a regression model in predicting women's preferences for screening for cervical cancer H_2), performance was very similar in terms of agreement. In both cases the kappa coefficient suggest a moderate agreement ($0.41<\kappa<0.60$). The difference in "goodness of fit", with the regression model correctly predicting the actual choices 78% of times as compared with 76% for the lexicographic rules, does not offer any clear indication on the type of decision strategy that respondents were most likely adopting.

Finally, as stated in H_3, the predictive ability of both compensatory and non-compensatory models turned out to be almost the same in the GPs study. The proportion of correctly predicted choices is 79% for both models and the associated kappa coefficient for both models indicates moderate agreement.

Table 6 presents the results from the segmentation analysis, testing for the influence of socio-economic characteristics and particular features of the studies on decision strategies. The first variable considered is response time - less than 10 minutes and 10 minutes or more. Results suggest that, as expected, the more time it takes a person to complete a questionnaire the more likely a compensatory rule is employed and hence the regression model outperforms the lexicographic rule in predicting actual choices. Closely related to the latter is self-reported ease of completion. The reasonable expectation that a compensatory model will fit better for those individuals who found the choice tasks difficult or very difficult seems to be confirmed in our data. Yet this does not generalise to individuals who found the exercise easy for whom evidence is mixed. We may be tempted to conclude that where the compensatory model predicts better individuals are able to cope with trade-offs while where the lexicographic rule performs better individuals minimising cognitive burden by resorting to simplifying non-compensatory decision strategies.

Table 6 Segmentation analysis. Compared performance

Characteristics	**Segment**	**Performance** (aggregated)
1. Time of completion	<10 minutes	LEX≥COMP
	>10 minutes	COMP > LEX
2. Ease of completion	Easy	COMP $\gtrless$ LEX
	Difficult	COMP >LEX
3. Age	Under 50	COMP>LEX
	Over 50	LEX>COMP
4. Income	Low	COMP~ LEX
	High	COMP>LEX

The results when segmenting according to age are as expected. Whilst respondents under 50 may be able to process information in a compensatory manner, an “information overload” effect leading to non-compensatory decisions is likely to occur for older subjects, as shown by the better performance of the hierarchical model in predicting choices for this group.

Finally an income effect is found – for low levels of income the performance of both compensatory and non-compensatory models is similar. However, for high levels income the compensatory model outperforms the lexicographic rule.

5. Discussion, conclusions and further research

Almost all empirical discrete choice modelling work in the literature assumes a utility-maximising, full information, indefatigable decision-maker who is able to assign values to alternatives, and choose the alternatives with the highest value, independent of context. Taken at face value the results presented herein should be of concern to proponents of this traditional compensatory decision model. The general conclusion emerging from this work is that there are no more reasons for assuming a compensatory decision process that there are for assuming a non-compensatory one. A comparison between the “goodness of fit” (as described by the number of correctly predicted choices) of a lexicographic decision rule and the traditional regression approach suggested that the non-compensatory model predicts behaviour at least as well as than the compensatory model depending on context.

Respondents' use of non-compensatory decision-making may pose a problem for researchers placing an emphasis on the elicitation of trade-offs for use in health care valuation. What if individuals do adopt non-compensatory decision strategies when completing stated choice experiments? How would this affect the "quality" of the information obtained from analysing responses? How can marginal rates of substitution and other behavioral outputs then be interpreted? The answer is still far from clear.

The empirical results presented in this paper are by no means definite. We recognise that the exercise has its limitations. Two seem rather obvious. Firstly, the psychology literature argues that there are a number of different decision-making strategies that individuals may adopt contingent upon context. This study concentrated on a lexicographic rule to get a preliminary insight. Future work should consider the extent to which other decision rules may be operating when individuals are presented with choices. Secondly, much less restrictive choice models than random effects probit are now available (e.g. the heteroskedastic extreme value (Bhat, 1995) or mixed logits (McFadden and Train, 2000). Future work should investigate the results obtained with these more flexible models.

In closing, as with Schwartz (2000), we believe that science often progresses by combining disciplines that previously seemed ill-matched. Therefore economic analysis of choice behaviour could, and maybe should, progress by integrating psychological models of decision making in a way that enriches both disciplines. Fortunately research seems to be moving towards the enhancement of discrete choice models in such a way, allowing for more behaviourally realistic representations of the choice model process hence a better understanding of behaviour. For example, Adamovicz and Swait (2001b) have proposed and tested a model of choice that incorporates the impact of task complexity and Swait (2001) has developed a formal model that can incorporate a wide range of decision strategies inferred from observed choice. Further research in the health field should investigate these approaches in more detail.

References

Bem, D.J. (1972): Self-Perception Theory", in: L. Berkowitz (ed.), Advances in Experimental Social Psychology, Academic Press: New York, 1-62.

Ben-Akiva, M. and S.R. Lerman (1985), "Discrete Choice Analysis: Theory and Application to Travel Demand,"

Bettman, J. R., E.J. Johnson, M.F. Luce, and J.W. Payne (1993), "Correlation, Conflict and Choice", *Journal of Experimental Psychological: Learning, Memory and Cognition* 19, 931-951.

Bettman, J.R. E.,M.F. Luce and J.W. Payne (1998) Constructive consumer choice processes, *Journal of consumer research* 25: 187-217

Bhat,C.R (1995) "A heteroscedastic extreme value model of intercity travel mode choice" Transportation research 29(6) 471-486

Borsch, K. (1968), The economics of uncertainty, Princeton University Press.

Brunswick (1955), "Representative design and probabilistic theory in a functional psychology", Psychological Review 62, pp.193-217

Bryan S, Buxton M, Sheldon R, Grant A (1998) Magnetic resonance imaging for the investigation of knee injuries: an investigation of preferences. Health Economics 7: 595-603

Calman, K.,(1996) Cancer: Science, society and the communication of risk. *BMJ*,. 313: p.799-802.

Drakopoulos, S.A. (1994) Hierarchical choice in economics. Journal of Economic Surveys 8:133-153.

Deshazo, JR and G. Fermo (2002) Designing choice sets for stated preference methods: the effects of complexity on choice consistency. Journal of Environmental Economics and Management,
Einhorn, H.J (1971), "Use of nonlinear, non-compensatory models as a function of Task and Amount of Information", Organizational Behaviour and Human Performance, 6, pp. 1-27
Ericcson, K. A., and H. A. Simon (1993) Protocol Analysis. Cambridge, MA: The MIT Press
Fischer, G. W., M.F. Luce and J. Jia, (2000), "Attribute Conflict and Preference Uncertainty: Effects on Judgment Time and Error", *Management Science* 46, 88-103.
Ford, J.K; N. Schmitt; S.L. Schechtman, Hults, B.M. and Doherty M.L (1989) Process tracing methods: contributions, problems and neglected research questions. Organizational behaviour and human decision process 43, 75-117
Georgescu-Roegen, N.(1958). "Threshold in Choice and the Theory of Demand", Econometrica.
Greene, W, H. 2000, Econometric analysis (Prentice Hall, New Jersey).
Haaijer, R., W. Kamakura and M. Wedel, (2000), "Response Latencies in the Analysis of Conjoint Choice Experiments", *Journal of Marketing Research* 37, 376-382.
Hensher, D. And L.Johnson (1981) *Applied Discrete Choice modelling* Landon-Croom-Helm/New York. John Wiley
Hicks, J. R. (1940). The valuation of social income. *Economica*, 7, 105–124.
Hoffman P.J.(1960) The paramorphic representation of clinical judgement Psychol. Bulletin 57 116 131
Hughes, G.D (1974) Buyer/consumer information processing: an overview of where researchers have been and where they should be going In G.D. Hughes and M.L. Ray (eds.) Buyer/consumer information processing. University of north carolina press, chapel hill, N.C. b) Trends in the development of multi-attribute models of central information processing.
Kahneman, D. (1973). Attention and Effort. Englewood Cliffs, N.J.: Prentice-Hall.
Kaldor, N. (1939). Welfare propositions of economics and interpersonal comparisons of utility. Economic Journal, 49, 549–551
Kleinmuntz, B., Ed. (1968). Formal Representation of Human Judgment. New York: Wiley
Lancaster, K (1972). Operationally relevant characteristics in the theory of consumer behaviour. In M. Peston, & B. Corry (Eds.), Essays in honour of Lord Robbins. London: Weidenfeld and Nicholson, Reprinted in:. In Modern consumer theory. Aldershot: Edward Elgar.
Lee-Gosselin M. and Pas E.I, (1997) The Implications of Emerging Contexts for Travel-Behavior Research. In Stopher P.R. and Lee-Gosselin M. (eds) *Understanding travel behavior in an Era of Change.* Elsevier Science Ltd.
Lichtenstein,A. And P.Slovic (1971), "Reversal of preference between bids and choices in Gambling decisions" *Journal of Experimental Psychology,* 89(1), 46-55.
Little (1957), A critique of welfare economics, Oxford, Claredon Press.
Lloyd, A.J., (2001) The extent of patients' understanding of the risk of treatments. Qual Health Care, 10 Suppl 1: 4-8.
MacCrimmon, K.R. (1968) *Decision-making among multiple alternatives: a survey and consolidated.* Rand Memorandom R.M.,4823- ARPA, Santa Monica.
McFadden, D. (1974), "Conditional Logit Analysis in Analysis of Qualitative Choice Behaviour", P.Zarembka (ed.), Frontiers in Econometrics, New York: Academic Press
- (1986),"The Choice Theory Approach to Market Research", Marketing Science 5, 275-97
McFadden, D. and K. Train (2000), "Mixed Multinomial Logit Models for Discrete Response," Journal of Applied Econometrics Vol. 15, No. 5, pp. 447-470
McIntosh and M. Ryan (2002) Using discrete choice experiments to derive welfare estimates for the provision of elective surgery: Implications of discontinuous preferences, Journal of Economic Psychology, Volume 23, Issue 3 , Pages 367-382
Menger, C. (1950[1871]) *Principles of Economics* Trans. By J. Dingwall and B. Hoselitz, Glencoe, Illinois: The Free Press
Nisbett, R.E., and Wilson T.D. (1977): "Telling More Than We Can Know: Verbal Reports on Mental Processes", Psychological Review 84, 231{259.
Olshavsky, R. W. (1979). Task complexity and contingent processing in decision making: A replication and extension. Organizational Behavior and Human Performance, 24, 300–316.
Mazzotta, M., & Opaluch, J. (1995). Decision making when choices are complex: A test of Heiner'shypothesis. Land Economics, 71(4), 500–515.
Payne, J.W. (1976) "Task complexity and Contingent Processing in Decision Making. An informatio Search and Protocol Analysis", *Organizational Behaviour and Humand Performance,* 16(August) 366-387.

Payne, J.W., Bettman, J.R, and Johnson, E.J. (1988), "Adaptive strategy selection in decision making", *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14, pp. 534-552.-

- (1992) "Behavioural decision research: a constructive processing perspective". *Annual Rev of Psych* 43, 87-131

- (1993), "The adaptive decision maker". Cambridge, England: Cambridge University Press.

Propper C. (1995), The disutility of time spent on UK national health service waiting lists. *Journal of Human Resources* 30.

Ratcliffe J. (1998), Patient's preferences regarding the process and outcomes of high technology medicine: an application of conjoint analysis to liver transplantation. Presented at the Health Economists' Study Group Meeting, University of Sheffield, January.

Rubinstein, A., (1998), *Modeling Bounded Rationality*, MIT Press, Cambridge, Mass.

Russ, F.A. Evaluation process models: objective and subjective comparisons. 19th Int IMS Meeting Houston TX

Ratcliffe J, Buxton M (1999) Patients' preferences regarding the process and outcomes of life-saving technology. International Journal of Health Technology Assessment in Health Care 15)2): 340-351

Ryan M and Hughes J. (1997) Using conjoint analysis to assess women's preferences for miscarriage management. Health Economics 1997, 6, 261-273.

Ryan M. and Gerard K. Using choice experiments to value health care programmes: current practice and future challenges. Paper presented at International Health Economic Association Conference, Univ. of York, July 2001.

Ryan, M, Reeves, C. and Entwistle.V. Listening to respondents: a think-aloud study of a discrete choice experiment. Paper presented at the II European Health Economics Conference, Paris, July, 2002

Ryan M. and Ubach C. Testing for an experience endowment effect within choice experiments *Applied Economic Letters (forthcoming)*

Ryan, M., Wordsworth, S. 2000. Sensitivity of willingness to pay estimates to the level of attributes in discrete choice experiments. Scottish Journal of Political Economy, 47;5 504-524.

Samuelson, W., & Zeckhauser, R. (1988). "Status quo bias in decision making." J. Risk & Uncertainty, 1:1, 7-59.

San Miguel, F., Ryan, R. and McIntosh, E. (2000), 'Applying Conjoint Analysis in Economic Evaluations: An Application to Menorrhagia', Applied Economics 32, 823-33.

Schwartz, H (2000), "Rationality gone awry? Decision Making Inconsistent with Economic and Financial Theory", Praeger.

Scott, A. (2002) Identifying and analysing dominant preferences in discrete choice experiments: An application in health care Journal of Economic Psychology 23, 383–398

Slovic, P. and MacPhimally, D. (1974) Dimensional commensurability and cue utilization in comparative judgement. *Organizational Behaviour and Human performance* 11, 172-194.

Spash, C.L (2000) Ecosystems, contingent valuation and ethics: the case of wetland re-creation Ecological Economics 34 Special issue Social Processes of environmental valuation 195–215

Swait, J. (2001) A non-compensatory choice model incorporating attribute cutoffs. *Transportation Research Part B* 35 (2001) pp 903-928

Swait, J. and Adamovicz, W (2001b). ''Choice Complexity and Decision Strategy Selection,'' Journalof Consumer Research, 28 (June), 135–148.

(2001b) Choice environment, market complexity and consumer behaviour: a theoretical and empirical approach for incorporating decision complexity into models of consumer choice. Organizational Behaviour and Human Decision Processes. 86(2) PP. 141-167.

Tversky, A and D. Kahneman.(1974), "Judgement under uncertainty: heuristics and biases", Science, 185,1124-1131.

Tversky, A. (1972a) "Elimination by aspects: A thoeyr of choice" Psychological Review, 79(July) 281-299(1972b) Choice by elimination *Journal of Mathematical Psychology* 9 341-361

White, P. (1980): Limitations on Verbal Reports of Internal Events: A Refutation of Nisbett & Wilson and of Bem", Psychological Review 87, 105{112.