

# Procedural Justice in Public Healthcare Resource Allocation

Aki Tsuchiya,<sup>1,2,3</sup> Luis Silva Miguel,<sup>4</sup> Richard Edlin,<sup>1</sup> Allan Wailoo<sup>1</sup> and Paul Dolan<sup>1,3</sup>

1 School of Health and Related Research, University of Sheffield, Sheffield, UK

2 Department of Economics, University of Sheffield, Sheffield, UK

3 Centre for Well-being in Public Policy, University of Sheffield, Sheffield, UK

4 Research Centre on the Portuguese Economy, Lisbon, Portugal

## Abstract

**Introduction:** The legal studies literature on procedural justice identifies six key characteristics of procedural justice: accuracy, consistency, impartiality, reversibility, transparency and voice. However, the relative importance of these in the context of public healthcare resource allocation is unclear, as is whether they are valuable instrumentally (because it contributes to better outcomes) or inherently (for its own sake).

**Methods:** A survey of 80-odd members of the UK public determined the following: the ranking of all the six characteristics; the pairwise comparisons of the characteristics; and whether each characteristic was important for instrumental reasons, for intrinsic reasons or for both.

**Results:** Respondents ranked the procedures in the order of accuracy, consistency, impartiality, reversibility and transparency. Procedural justice was valued for both instrumental and inherent reasons.

**Discussion/conclusion:** A robust ranking of five of the six procedural characteristics was found. The ranking for voice was sensitive to the question format, which has methodological implications. Around a quarter to a third of respondents regarded a procedural characteristic to have entirely intrinsic value.

The UK NHS has two objectives in terms of consequences of health policy: to maximise population health and to reduce avoidable inequalities in health. Should it also matter that resource allocation decisions are made in an appropriate manner, according to certain procedural rules? In other words, criteria for evaluating a decision could include whether or not the processes that led to the decision satisfy the characteristics of procedural justice. An important issue, then, would be whether the value of the procedures is instrumental or intrinsic. Economics has traditionally been a predominantly consequentialist discipline, and the typical treatment of procedural characteristics within economics has been to interpret them as having solely instrumental value – they are valuable only to the extent that they enhance outcomes. Yet this view has been challenged by those who think that the public may appreciate procedural characteristics for what they are, thereby recognising their inherent value.

The existing theoretical and empirical literature on procedural justice has developed mostly in the legal field, but the concepts discussed there are also highly relevant to healthcare resource allocation decisions (for example, see Coast,<sup>[1]</sup> Daniels<sup>[2]</sup> and Daniels and Sabin<sup>[3]</sup>). The purpose of this article is to empirically explore the relevance of procedural concerns in public healthcare resource allocation by eliciting the procedural preferences of the public.

## Background

The existing literature on procedural justice, which mostly relates to legal and organisational contexts such as court cases, arbitration of disputes, pay decisions, etc., includes six key procedural characteristics: accuracy; consistency; impartiality; reversibility; transparency; and voice (for example, see Thibaut and Walker,<sup>[4]</sup> Leventhal<sup>[5]</sup> and Lind and Tyler<sup>[6]</sup>). This section briefly introduces each of these characteristics and explains their rele-

vance to public healthcare resource allocation procedures. In particular, we base our discussion on the technology assessment mechanism of the National Institute for Health and Clinical Excellence (NICE) for England and Wales, but the principles are relevant to any other similar "independent organisation responsible for providing national guidance on the promotion of good health and the prevention and treatment of ill health".<sup>[7]</sup>

*Accuracy* requires that the information on which decisions are based is factually correct and up to date. Organisations such as NICE are a result of the wider move towards evidence-based medicine, and the value of evidence depends crucially on its accuracy. Technology assessment procedures typically grade different types of evidence and aim to give them due weight.

*Consistency* as a procedural characteristic within a given decision-making body requires that the set of rules and criteria used in decisions remains constant through time and across various cases. So the set of rules applied to assess one technology will be applied for the assessment of another technology. Technology assessment carried out by bodies such as NICE follows published guidelines that set out the steps and criteria involved in the procedure.<sup>[8]</sup>

*Impartiality* means that there is absence of vested interests and biases. For instance, members of decision-making committees would be required to declare all possible conflict of interests and/or would not be allowed to sit in sessions where their professional or personal interests may be relevant to a case discussed.

*Reversibility* requires that there be opportunities, under certain conditions, for affected parties to challenge the decision. NICE arranges for this by giving stakeholders the right to appeal on procedural grounds, but not if it is simply due to a disagreement with the decisions made.<sup>[9]</sup> Also, there can be built-in mechanisms for reversibility, such as having review dates set for decisions.

*Transparency* is about making visible the process and the results of decision making. However, publicising all the information that went into the decision-making procedure can be problematic, since some of the material submitted may be commercially sensitive. In fact, the House of Commons criticised that "much of the information on which NICE appraisals are based is unpublished, as it is supplied to NICE by manufacturers in confidence".<sup>[10]</sup> Furthermore, simply making huge amounts of information available does not necessarily achieve transparency if the relevant contents are inaccessible. This can happen when the volume of information makes it costly to identify what one is looking for, and/or when the technical nature of the issues addressed is too complex for simple dissemination without detailed technical notes. There is also the concern that no written material

can make the workings of the group dynamic involved in decision making completely transparent.

*Voice* requires that there is opportunity for concerned parties (especially affected individuals) to have their views and interests heard. It may take the form of direct voice, where the affected individuals will themselves be heard, or indirect voice, where representatives will be used. The increased involvement of patients, carers and lay members of the public in healthcare decision making is a trend in line with voice (for example, the Citizens Council set up by NICE in 2002 [<http://www.nice.org.uk/page.aspx?o=113693>]).

Overall, these six procedural characteristics seem relevant and reasonable to expect of mechanisms that make decisions on public healthcare resource allocations. In this article, we aim to explore two further questions that the existing literature does not address: first, whether or not these procedural characteristics are all equally important in publicly funded healthcare; and second, whether each of these procedural characteristics are relevant in the health sector because they are expected to increase the likelihood of better outcomes or because they have some intrinsic value in themselves.

## Methods

In the following sections we explore whether members of the public consider all six characteristics of procedural justice equally relevant and, if not, what weights they would give to each. In order to place the exercise in context, NICE was used as an example. Preceding the survey, participants were given a brief introduction to the scarcity of healthcare resources, the need for organisations such as NICE, and what NICE does (including a brief outline of its activities). It was not expected that many people would be familiar with how NICE operates or what NICE actually does. However, the actual questions did not rely on such knowledge and were entirely normative (how NICE *should* operate as opposed to how well NICE *is* achieving a given target).

## Questionnaire

A questionnaire was designed consisting of three sets of questions (see appendix). The aim of the first two sets (figures A1 and A2) was to ask the respondents to express their preferences on the relative importance of each of the six procedural characteristics. It is conceivable to simply ask direct questions on these, but without a reference point the responses will be meaningless (my preference ranking for apples and oranges will depend on how many apples and oranges I have already consumed). In order to tap into preferences at the margin, the questionnaire asked respondents to

assume that NICE was performing at a medium level on all six characteristics (on a three-point scale: low, medium, high). Then question 1 (figure A1) asked respondents to choose the order in which they would improve the procedural characteristics from medium to high, under the assumptions that there were additional resources available, and that improvements on any one characteristic would cost the same.

The second set of questions (question 2, see figure A2) presented a series of pairwise comparisons of the six characteristics, giving 15 combinations. Respondents were asked to rank three options regarding each pair: one that left both of the characteristics at 'medium', one that had one characteristic at 'high' and the other at 'low', and one where the 'high' and 'low' characteristics were reversed. This implies a built-in consistency test: by selecting three pairwise comparisons across three characteristics, violations of transitivity (if high  $x$  is preferred to high  $y$  and if high  $y$  is preferred to high  $z$ , then high  $x$  is preferred to high  $z$ ) can be picked up. Given the large number of pairwise combinations, and that the characteristics are not completely independent of each other, it will be difficult for a respondent to avoid any violations of this transitivity test.

Finally, in the third set of questions (question 3, see figure A3), subjects were asked if they value the procedural characteristic because of its instrumental nature, its intrinsic nature, or both. The objective was to understand if respondents are concerned about the consequences, about the way in which decisions are taken, or both. Those respondents indicating that a characteristic was important only "because it may improve the decision" were coded as having an instrumental response. If the reason was "because it is a good thing even if it does not improve the decision", then this was regarded as being of intrinsic value. Respondents were allowed to agree with both of these statements.

To examine the effect of the order in which the six characteristics were presented, two versions of the questionnaire were prepared (see table I). In version 1, the six characteristics were presented in alphabetical order throughout (accuracy, consistency, ...), and question 3 presented the instrumental reason followed by the intrinsic reason. Version 2 was arranged in the following way: in question 1, the last three characteristics of version 1 came first, followed by the first three characteristics; in question 2, the six characteristics were presented in reverse alphabetical order; and in question 3 the intrinsic reason preceded the instrumental equivalent.

**Table I.** The order of presentation of the six characteristics of procedural justice by question and questionnaire version

Question number and description	Order of procedure presentation	
	version 1	version 2
Question 1: relative importance of procedures	A-C-I-R-T-V	R-T-V-A-C-I
Question 2: relative importance of procedures	A-C-I-R-T-V	V-T-R-I-C-A
Question 3: instrumental vs intrinsic reasons for procedural importance	A-C-I-R-T-V; instrumental reason presented first	A-C-I-R-T-V; intrinsic reason presented first

A = accuracy; C = consistency; I = impartiality; R = reversibility; T = transparency; V = voice.

### Analysis

Non-parametric Mann-Whitney U tests were used to test for order effects across the two versions of the questionnaire. Ranking data were described in three ways: by showing how often a given procedural characteristic was ranked first; by calculating Borda scores; and by calculating Thurstone indices. The distributions of respondents across various pairs of characteristics were cross-tabulated. Two-sided z-tests were used to test for equivalence of distributions, and all tests were conducted at a 5% significance level.

Borda scores are derived by giving score 6 to the first-ranked characteristic, score 5 to the second-ranked characteristic, and so on for each respondent, and then calculating the average score across respondents. The Borda score represents aggregate ordinal preferences by assuming that ranking by individuals satisfies interval scale properties.<sup>[11]</sup> Thurstone indices are derived from the number of times a given characteristic is ranked higher than any other characteristic.<sup>[12]</sup> This is interpreted as a probability assumed to be drawn from a normal distribution, so that it can be transformed into a z-value, or the Thurstone index. These indices represent the degree by which a given characteristic is preferred over another in cardinal terms. In the analysis in the Results section, the Thurstone indices are 'standardised' using the same two characteristics as anchors, so that the indices can be compared across questions 1 and 2.

The reason for using these two summary measures of aggregate ordinal preference is that neither is perfect. Borda scores assume interval scale properties on ranking data, which is difficult to justify, and the sample size we have in this study is, strictly speaking, too small to justify the use of Thurstone indices. However, given that these two measures do not have a common theoretic

**Table II.** Background details of sample (n = 88)

Background characteristic	%
<b>Education</b>	
Has 'a degree or equivalent qualification'	34
Has education beyond minimum schooling age	46
<b>Age (y)</b>	
16-44	25
45-64	41
65+	34
<b>Other</b>	
Ethnicity: White	86
In full-time employment	40
Female	54
Has a school-age child	12

cal base, if they both agree on a finding, the agreement can be interpreted as indirect evidence of its robustness.

#### Sample

The questionnaire was handed out in July 2002 to individuals participating in another project on the distribution of health policy outcomes.<sup>(13)</sup> In Sheffield, UK, 2000 individuals randomly selected from the register in two electoral constituencies were invited to participate in that study. Of these, 257 individuals (13%) replied positively, providing age and sex information on their reply slips. Due to limited resources, 192 individuals were selected for the other study, reflecting the local age and sex structure. Of these invitees, 128 (67%) participated in one of the 24 sessions, in groups of five to eight people. At the end of the sessions, a brief presentation on NICE and technology appraisal was given, then

participants were asked to take home the questionnaire and to return it by post, using the freepost envelope provided.

#### Results

The questionnaire was returned by 91 (71%) participants. Three responses were excluded because none of the questions were completed. Additional responses were lost due to missing answers within the questionnaire, leaving 84 valid answers to question 1, 80 to question 2 and 87 to question 3. The distribution of sample background characteristics is given in table II. Relative to the local population in Yorkshire and Humberside, the sample is representative in terms of sex and smoking distribution of respondents, while those <45 years of age, those with children or those employed are under-represented.

The Mann-Whitney U test was applied to analyse the effect of the order of presentation. Regarding question 1, no statistically significant difference in ranking was found. The test for question 2 involved 15 pairwise comparisons with three options each, implying 45 tests, of which only one was statistically significant (low impartiality/high reversibility was ranked worse when it was presented as the lower choice than when it was the upper choice;  $p = 0.016$ ). The test for question 3 resulted in no statistically significant results. Therefore, the null hypothesis that there are no order effects in this questionnaire could only be rejected in one case. The results from the two versions are therefore pooled in the results reported in this section.

Table III summarises the distribution of first choices (for example, 44% of respondents ranked accuracy first in question 1), the Borda scores and the Thurstone index scores, with their respective ranking. As can be seen, the rank ordering of the two summary measures are in complete agreement across both questions. In

**Table III.** Responses to questions 1 and 2: distribution and ranking of the six procedural justice characteristics

Procedural characteristic	Question 1 (n = 84)				Question 2 (n = 80)			
	first choice		scores (ranking)		first choice		scores (ranking)	
	n	%	Borda	Thurstone index <sup>a</sup>	n	%	Borda	Thurstone index <sup>a</sup>
Accuracy	37	44	4.81 (1)	1.00 (1)	36	60	4.78 (1)	1.00 (1)
Consistency	17	20	4.04 (2)	0.67 (2)	7	12	4.11 (2)	0.66 (2)
Impartiality	7	8	3.26 (4)	0.37 (4)	4	7	3.21 (3)	0.21 (3)
Reversibility	4	5	2.95 (5)	0.25 (5)	3	5	3.14 (4)	0.18 (4)
Transparency	2	2	2.32 (6)	0.00 (6)	2	3	2.75 (6)	0.00 (6)
Voice	17	20	3.62 (3)	0.51 (3)	8	13	3.01 (5)	0.13 (5)

a The Thurstone indices are standardised at 0 and 1 for transparency and accuracy, respectively.



**Table IV.** Distribution of pairwise preferences in question 1 (%)

Row procedure	Column procedure					
	accuracy	consistency	impartiality	reversibility	transparency	voice
Accuracy		0.62	0.81	0.85	0.87	0.67
Consistency	0.38		0.63	0.65	0.81	0.56
Impartiality	0.19	0.37		0.56	0.70	0.44
Reversibility	0.15	0.35	0.44		0.61	0.40
Transparency	0.13	0.19	0.30	0.39		0.31
Voice	0.33	0.44	0.56	0.60	0.69	

question 2, a majority (54%) of respondents exhibited no violations of transitivity at all.

In both questions 1 and 2, accuracy and consistency are the most important procedures, and transparency the least important. However, voice, which was ranked above impartiality and reversibility in question 1, is ranked below both in question 2. To explore this, we also considered the pairwise preference between different procedures. Table IV displays the pairwise preferences in question 1, where the figures in each cell represent the proportion of respondents who preferred the *row* procedure to the *column* procedure. For example, 62% of respondents considered accuracy to be more important than consistency. Table V presents the equivalent data for question 2.

In terms of preference ordering at the aggregate level, the cells that change across questions 1 and 2 from  $>0.5$  to  $<0.5$  (or vice-versa) are of interest, indicating a change in the median preferences between the two procedures. The change in preference between impartiality and voice (from 0.44 to 0.55;  $p = 0.16$ ) and between reversibility and voice (from 0.40 to 0.56;  $p = 0.04$ ) are thus consistent with the changes in the ranks. However, this does not enable us to understand whether the change reflects a lower valuation of voice or a higher valuation of impartiality and reversibility in question 2 relative to question 1. The standardised Thurstone index scores can offer some insight. The results in table III indicate that voice is given a much lower importance in

question 2, probably affecting the differences in the ranks between the two questions. The indices also suggest that the relative importance of accuracy may be higher in question 2.

Mann-Whitney U tests were also performed to compare the rank of each procedure across questions 1 and 2, with only the test for differences in the voice characteristic ( $p = 0.088$ ) approaching significance. Although this voice test is insignificant, that the  $p$ -value is very small relative to those for the other procedures ( $p > 0.40$  in each case) may suggest a difference between the valuation of voice and the other characteristics.

The structure of question 2 also allows us to determine the proportion of respondents that was willing to change the current situation. Interestingly, the greater the difference between the valuations of each pair of procedures measured by the Thurstone index, the more respondents were willing to reallocate resources. This can be interpreted as an indication of data quality since we expect people to assign more resources to the procedures they value most.

Figure 1 illustrates the results of question 3. This question asked respondents why they believed each procedural characteristic was important in terms of the resulting distributions of responses across 'instrumental only', 'both' and 'intrinsic only'. Accuracy was considered the most important procedure in questions 1 and 2, and in question 3 it was the procedure with the highest value in 'instrumental only' and the second lowest in

**Table V.** Distribution of pairwise preferences in question 2 (%)

Row procedure	Column procedure					
	accuracy	consistency	impartiality	reversibility	transparency	voice
Accuracy		0.74	0.79	0.81	0.74	0.70
Consistency	0.26		0.71	0.68	0.76	0.70
Impartiality	0.21	0.29		0.55	0.61	0.55
Reversibility	0.19	0.32	0.45		0.61	0.56
Transparency	0.26	0.24	0.39	0.39		0.48
Voice	0.30	0.30	0.45	0.44	0.52	

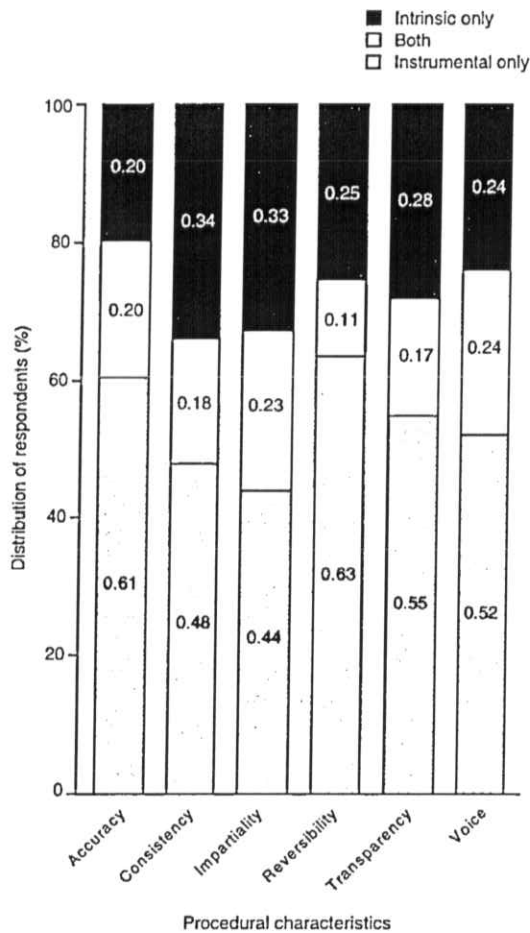


Fig. 1. Respondents' reasons for each procedural characteristic being important.

'intrinsic only'. These results lead us to speculate that some relationship might exist between the relative importance of each procedure and the answer to question 3. Chi-square tests were performed to address this. However, only transparency resulted in a significant result ( $p = 0.047$ ), and the results for the other characteristics were insignificant ( $p = 0.09$  for reversibility;  $p > 0.20$  for the rest).

## Discussion and Conclusions

Several characteristics have been demonstrated to be key determinants of procedural justice in the legal and organisational literature. Some of these are also relevant to bodies set up to make resource allocation decisions in public healthcare, such as NICE. This exploratory study has looked at the way in which members of the public view these characteristics.

The study sample ranked the procedural characteristics in the order of accuracy, consistency, impartiality, reversibility and

transparency. The importance placed on voice is unclear, as its relative value changed between questions 1 and 2. This is in contrast to most other empirical studies where voice is consistently highly valued. It was the only characteristic in our study whose relative ranking changed significantly, which implies two things: first, that the results regarding the other characteristics are robust; and second, a need to explore why the ranking for voice alone changed.

A plausible explanation may lie in the distinction between (a) the impact of improving one procedural characteristic that may affect the performance of other procedural characteristics and (b) the relative importance of the different procedural characteristics on their own. For instance, if involving members of the general public is thought to reduce the possibility of bias in decision making, then improving voice automatically improves impartiality too. In this case, the impact of improving voice is related to the importance of voice, impartiality and also possibly other procedural factors. Even if voice is not that important in itself (as in question 2), it may be selected with greater priority in pairwise choices (as in question 1) because of a belief that it has a broad impact on a range of other procedural characteristics. If so, then it should be noted that, although voice is the only characteristic where we observe differences between questions 1 and 2, this should not be interpreted to mean that voice is the only procedural characteristic that has this knock-on effect. Similar effects may be present for other characteristics as well, such as transparency promoting impartiality, but are simply not large enough to affect the relative ranking between the two questions.

Were such knock-on effects to exist, this would have important methodological implications for ordinal evaluation of multiple items where it cannot be established that all the items are perceived to be independent of each other. This is potentially relevant beyond the issue of procedural justice, since it is often assumed in empirical research that the preference ranking of items is independent of the preference elicitation method.

More generally on the issue of the reasons for procedural justice, responses were observed across all three categories (instrumental only, intrinsic only, or both). While around half the respondents saw only instrumental value to a procedural characteristic, around a quarter to a third of respondents regarded a characteristic to have entirely intrinsic value. In other words, a significant proportion of the present study sample seems to hold the view that decisions should not merely aim to meet consequentialist objectives of the healthcare system, but should also be made according to a set of procedures.

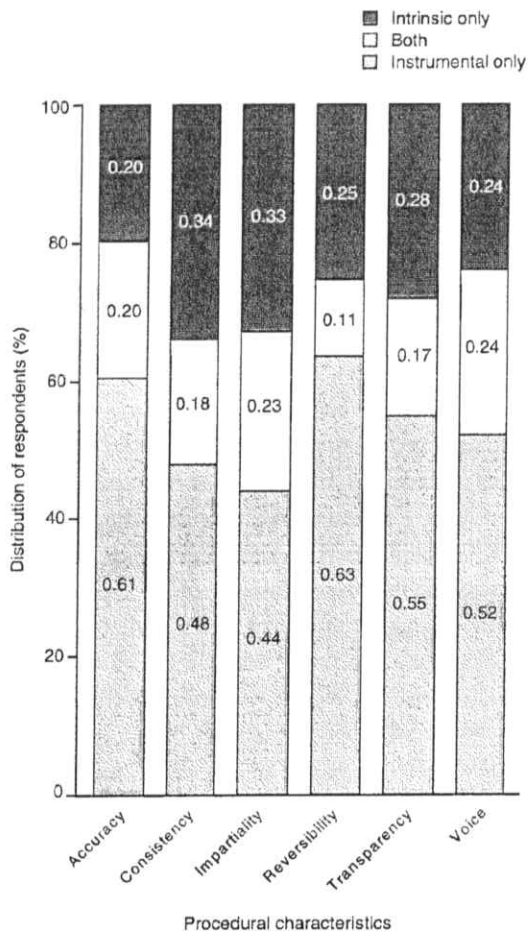


Fig. 1. Respondents' reasons for each procedural characteristic being important.

'intrinsic only'. These results lead us to speculate that some relationship might exist between the relative importance of each procedure and the answer to question 3. Chi-square tests were performed to address this. However, only transparency resulted in a significant result ( $p = 0.047$ ), and the results for the other characteristics were insignificant ( $p = 0.09$  for reversibility;  $p > 0.20$  for the rest).

## Discussion and Conclusions

Several characteristics have been demonstrated to be key determinants of procedural justice in the legal and organisational literature. Some of these are also relevant to bodies set up to make resource allocation decisions in public healthcare, such as NICE. This exploratory study has looked at the way in which members of the public view these characteristics.

The study sample ranked the procedural characteristics in the order of accuracy, consistency, impartiality, reversibility and

transparency. The importance placed on voice is unclear, as its relative value changed between questions 1 and 2. This is in contrast to most other empirical studies where voice is consistently highly valued. It was the only characteristic in our study whose relative ranking changed significantly, which implies two things: first, that the results regarding the other characteristics are robust; and second, a need to explore why the ranking for voice alone changed.

A plausible explanation may lie in the distinction between (a) the impact of improving one procedural characteristic that may affect the performance of other procedural characteristics and (b) the relative importance of the different procedural characteristics on their own. For instance, if involving members of the general public is thought to reduce the possibility of bias in decision making, then improving voice automatically improves impartiality too. In this case, the impact of improving voice is related to the importance of voice, impartiality and also possibly other procedural factors. Even if voice is not that important in itself (as in question 2), it may be selected with greater priority in pairwise choices (as in question 1) because of a belief that it has a broad impact on a range of other procedural characteristics. If so, then it should be noted that, although voice is the only characteristic where we observe differences between questions 1 and 2, this should not be interpreted to mean that voice is the only procedural characteristic that has this knock-on effect. Similar effects may be present for other characteristics as well, such as transparency promoting impartiality, but are simply not large enough to affect the relative ranking between the two questions.

Were such knock-on effects to exist, this would have important methodological implications for ordinal evaluation of multiple items where it cannot be established that all the items are perceived to be independent of each other. This is potentially relevant beyond the issue of procedural justice, since it is often assumed in empirical research that the preference ranking of items is independent of the preference elicitation method.

More generally on the issue of the reasons for procedural justice, responses were observed across all three categories (instrumental only, intrinsic only, or both). While around half the respondents saw only instrumental value to a procedural characteristic, around a quarter to a third of respondents regarded a characteristic to have entirely intrinsic value. In other words, a significant proportion of the present study sample seems to hold the view that decisions should not merely aim to meet consequentialist objectives of the healthcare system, but should also be made according to a set of procedures.

This study is exploratory in that its aims are to identify the set procedural characteristics that are relevant to resource allocation decisions in publicly funded healthcare. It should be noted, for instance, that the relative importance of procedural characteristics, and the reasons for this (instrumental vs intrinsic), may depend on specific issues addressed by the decision in question and the context in which the characteristics are set. In this respect, the relatively high support given to consistency is promising, since it implies the use of the same decision rules across different contexts. However, the current study asks only abstract questions and thus does not give specific contextual information to test this support for consistency.

Furthermore, the opportunity to explain and to discuss the objectives and workings of NICE, or the concepts of and the relationships between the procedural characteristics, were fairly limited. However, the questions asked were normative, whereby people can have legitimate preferences without detailed knowledge of how NICE or other similar organisations actually operate. Additionally, the preferences revealed by this sample are based on a simplifying assumption: namely, that the cost of improvement for any given procedural characteristic is the same. In reality, it is probably less expensive to achieve a higher level of transparency than to achieve higher levels of accuracy or consistency. The study has illustrated that people value individual procedural characteristics differently, with a robust ordering for five of the six characteristics. Also, while the majority of responses regarded procedural characteristics as having only instrumental value, a persistent minority acknowledged intrinsic value. Although there is increasing interest in procedural justice in economics (for example, see Hahn,<sup>[14]</sup> Ng,<sup>[15]</sup> Frey and Pommerehne,<sup>[16]</sup> Frey and Oberholzer-Gee<sup>[17]</sup> and Sen<sup>[18]</sup>), currently there is little information on how the different characteristics should be weighted against each other and against consequentialist outcomes. This study, exploring rankings between them, offers a starting point.

**Acknowledgements**

This study was funded by the Arts and Humanities Research Board UK, as part of their Innovation Awards, on project number 13789 "Concepts of procedural justice: from legal studies to social decision making". LSM was funded by the Portuguese Foundation for Science and Technology - III Community Support Framework (Fundação para a Ciência e a Tecnologia). Special thanks are due to Rebecca Shaw and all the respondents who took part in the survey. The usual disclaimer applies.

There are no potential conflicts of interest directly related to this article.

**Appendix**

Questionnaire (Version 1)

Questions 1, 2 and 3 from the questionnaire are shown in figures A1, A2 and A3, respectively.

*Introduction*

The purpose of this questionnaire is to find what you think about the importance of procedures used by NICE during the decision-making process. There are no wrong or right answers; what is important is that you express your opinion.

NICE is part of the NHS and its role is to provide guidance about treatments for health professionals and for patients and their

**QUESTION 1**

NICE's performance in each procedure can be classified as 'low', 'medium' or 'high'. Let us assume that NICE is at a medium level in all these six procedures:

	low	medium	high
Accuracy		X	
Consistency		X	
Impartiality		X	
Reversibility		X	
Transparency		X	
Voice		X	

Imagine that some extra funds are available to change the performance of NICE in one of those six procedures. Bearing in mind that the costs of doing it are the same for all procedures, which one would you choose? Please place a '1' in the box next to your choice (in the table below).

Accuracy	<input type="checkbox"/>
Consistency	<input type="checkbox"/>
Impartiality	<input type="checkbox"/>
Reversibility	<input type="checkbox"/>
Transparency	<input type="checkbox"/>
Voice	<input type="checkbox"/>

Now imagine that even more resources are available and that you would be given the chance to choose another procedure in which you would like to change NICE's performance. Which one would you choose? Please place a '2' in the box next to it.

Please repeat the process until you have numbered all choices from '1' to '6'.

If you are not happy with the final order please change it.

Fig. A1. Question 1 from the questionnaire.



**QUESTION 2**

Now suppose that there are no extra funds available. However, NICE could change the way it makes decisions. For example, it could have a high level of accurate information used in decision making and a low level of consistency. Please rank the three options in each question below (place a '1' in the option that you think is best, a '2' in the next one and a '3' in the one that you think is worst).

A	high accuracy and low consistency medium accuracy and medium consistency low accuracy and high consistency	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	H	high consistency and low transparency medium consistency and medium transparency low consistency and high transparency	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
B	high accuracy and low impartiality medium accuracy and medium impartiality low accuracy and high impartiality	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	I	high consistency and low voice medium consistency and medium voice low consistency and high voice	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
C	high accuracy and low reversibility medium accuracy and medium reversibility low accuracy and high reversibility	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	J	high impartiality and low reversibility medium impartiality and medium reversibility low impartiality and high reversibility	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
D	high accuracy and low transparency medium accuracy and medium transparency low accuracy and high transparency	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	K	high impartiality and low transparency medium impartiality and medium transparency low impartiality and high transparency	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
E	high accuracy and low voice medium accuracy and medium voice low accuracy and high voice	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	L	high impartiality and low voice medium impartiality and medium voice low impartiality and high voice	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
F	high consistency and low impartiality medium consistency and medium impartiality low consistency and high impartiality	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	M	high reversibility and low transparency medium reversibility and medium transparency low reversibility and high transparency	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
G	high consistency and low reversibility medium consistency and medium reversibility low consistency and high reversibility	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	N	high reversibility and low voice medium reversibility and medium voice low reversibility and high voice	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
			O	high transparency and low voice medium transparency and medium voice low transparency and high voice	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Fig. A2. Question 2 from the questionnaire.

carers. In doing so, NICE aims to give patients the best possible healthcare within the available resources and to standardise the treatments provided by the NHS all over the country.

Before asking you some questions let us explain what we mean by the following expressions:

- *Accuracy* of information – quality of the information on which decisions are based.
- *Consistency* of decisions – making decisions in the same way for all people and for different areas of the country.
- *Impartiality* of decision makers – decision makers are impartial if their decisions are independent of any personal preconceptions or interests (financial or not).
- *Reversibility* of decisions – possibility to dispute/discuss a decision that can be changed if there is a reason to do so.
- *Transparency* – revelation of the information used and the activities of decision makers.
- *Voice* – involvement in the decision-making process of the general public and/or individuals that will be affected by the

**QUESTION 3**  
 Please select one option, or both, for each question (put a cross in the box next to it):

a) Do you think that accuracy should be increased ...  
 ... because it may improve the decision and/or  
 ... because it is a good thing even if it does not improve the decision

b) Do you think that consistency should be increased ...  
 ... because it may improve the decision and/or  
 ... because it is a good thing even if it does not improve the decision

c) Do you think that impartiality should be increased ...  
 ... because it may improve the decision and/or  
 ... because it is a good thing even if it does not improve the decision

d) Do you think that reversibility should be increased ...  
 ... because it may improve the decision and/or  
 ... because it is a good thing even if it does not improve the decision

e) Do you think that transparency should be increased ...  
 ... because it may improve the decision and/or  
 ... because it is a good thing even if it does not improve the decision

f) Do you think that voice should be increased ...  
 ... because it may improve the decision and/or  
 ... because it is a good thing even if it does not improve the decision

*This is all. Thank you very much.*

Fig. A3. Question 3 from the questionnaire.

decision (e.g. patients). This may be by means of representatives.

**References**

I. Coast J. Citizens, their agents and health care rationing: an exploratory study using qualitative methods. *Health Econ* 2001; 10: 159-74

2. Daniels N. Justice, fair procedures, and the goals of medicine. *Hastings Cent Rep* 1996; 26: 10-2

3. Daniels N, Sabin J. Limits to health care: fair procedures, democratic deliberation, and the legitimacy problem for insurers. *Philos Public Aff* 1992; 26: 303-50

4. Thibaut J, Walker L. *Procedural justice, a psychological analysis*. Hillsdale (NJ): Lawrence Erlbaum Associates, 1975

5. Leventhal GS. What should be done with equity theory? New approaches to the study of fairness in social relationships. In: Gergen KJ, Greenberg MS, Willis RH, editors. *Social exchange: advances in theory and practice*. New York: Plenum Press, 1980

6. Lind EA, Tyler TR. *The social psychology of procedural justice*. New York: Plenum Press, 1988

7. National Institute for Health and Clinical Excellence [online]. Available from URL: <http://www.nice.org.uk/page.aspx?o=home> [Accessed 2005 Aug 30]

8. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal (reference N0515) [online]. Available from URL: [http://www.nice.org.uk/pdf/TAP\\_Methods.pdf](http://www.nice.org.uk/pdf/TAP_Methods.pdf) [Accessed 2005 Aug 30]

9. National Institute for Health and Clinical Excellence. Appraisal process: guidance for appellants (reference N0520) [online]. Available from URL: [http://www.nice.org.uk/pdf/TAP\\_Guide\\_for\\_Appellants.pdf](http://www.nice.org.uk/pdf/TAP_Guide_for_Appellants.pdf) [Accessed 2005 Aug 30]

10. House of Commons Health Committee. National Institute for Clinical Excellence: second report of session 2001-02. Volume 1: report and proceedings of the Committee [online]. Available from URL: <http://www.publications.parliament.uk/pa/cm200102/cmselect/cmhealth/515/515.pdf> [Accessed 2005 Aug 30]

11. Young HP. An axiomatization of Borda's rule. *J Econ Theory* 1974; 9: 43-52

12. Thurstone LL. *The measurement of values*. Chicago (IL): University of Chicago Press, 1960

13. Dolan P, Tsuchiya A. Health priorities and public preferences: the relative importance of past health experience and future health prospects. *J Health Econ* 2005; 24: 703-14

14. Hahn F. On some difficulties of the utilitarian economist. In: Sen A, Williams B, editors. *Utilitarianism and beyond*. Cambridge: Cambridge University Press, 1982

15. Ng YK. Economic efficiency versus egalitarian rights. *Kyklos* 1988; 41: 215-37

16. Frey BS, Pommerehne WW. On the fairness of pricing: an empirical survey among the general population. *J Econ Behav Organ* 1993; 20: 295-307

17. Frey BS, Oberholzer-Gee F. Fair siting procedures: an empirical analysis of their importance and characteristics. *J Policy Anal Manage* 1996; 15: 353-76

18. Sen A. Why health equity? *Health Econ* 2002; 11: 659-66

Correspondence and offprints: Dr Aki Tsuchiya, School of Health and Related Research, University of Sheffield, 30 Regent Street, Sheffield, S1 4DA, UK.

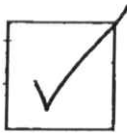
E-mail: a.tsuchiya@sheffield.ac.uk



## DOCUMENT SUPPLY

Boston Spa, Wetherby  
West Yorkshire LS23 7BQ  
www.bl.uk

### Please note the following:



This is the best copy available  
(Some pages may have been copied more  
than once for quality purposes)



This article has a very tight binding



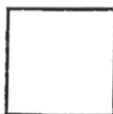
Some pages within the original article are  
advertisements and have therefore not  
been sent

Advertisement pages: .....



Some pages within the original are blank  
and have therefore not been sent

Blank pages: .....



The article you require is on different pages  
to those that you quoted