SOCIAL SCIENCE RESEARCH 14, 195–204 (1985)

The Effect of Restricted and Unrestricted Verdict Options on Juror Choice

BERNARD GROFMAN

School of Social Sciences, University of California, Irvine

Data on the effect of limiting the number of verdict options open to jurors on the probability of acquittal are reanalyzed. Strong support is found for a model which postulates that jurors' preferences are single peaked with respect to an underlying verdict severity continuum. Limited support is found for an anchoring effect in which the addition of new verdict options affects the perceived relative fairness of other verdict options. The implication of the single-peakedness model is that some jurors will refuse to vote for conviction if the verdict (punishment) is seen as too harsh even though the defendant is perceived to be guilty of committing a crime. © 1985 Academic Press, Inc.

CHOICE AMONG MORE THAN TWO VERDICT OPTIONS: EXPERIMENTAL EVIDENCE

An important question for jury trials is how the nature of the alternatives open to the jury affects their decisions. Vidmar's (1972) experiment consisted of presenting an abridged transcript of a murder trial to students in an introductory psychology class who were instructed to act as jurors. In Vidmar's experiment there were four verdict options in the unrestricted case, which we shall denote F = first-degree murder, S = second-degree murder, M = manslaughter, N = not guilty.

Groups of 24 simulated jurors were run through each of seven verdict option conditions: (1) F or N, (2) S or N, (3) M or N, (4) F or S or N, (5) F or M or N, (6) S or M or N, (7) F or S or M or N. The data collected by Vidmar are presented in Table 1. Vidmar's (1972) data do

The author is indebted to Lee Hamilton and Neil Vidmar for providing access to the data used in this paper and for helpful suggestions in expanding the analysis. The programming assistance of Paula Newhouse was invaluable. This research was supported by National Science Foundation Grants SOC 7514091 and SES 82 18588, Law and Social Sciences Program, and the initial draft was written while the author was a member of the Department of Political Science, State University of New York at Stony Brook. Send requests for reprints to Bernard Grofman, School of Social Sciences, University of California, Irvine, CA 92717.

0049-089X/85 $3.00

TABLE 1
Decision Alternatives and Frequency of Verdicts in Vidmar (1972)

	Condition						
Alternative	1	2	3	4	5	6	7
First-degree murder	11 (46%)	—	—	2 (8%)	7 (29%)	—	2 (8%)
Second-degree murder	—	20 (84%)	—	22 (92%)	—	11 (46%)	15 (63%)
Manslaughter	—	—	22 (92%)	—	16 (67%)	13 (54%)	5 (21%)
Not guilty	13 (54%)	4 (17%)	2 (8%)	0 (0%)	1 (4%)	0 (0%)	2 (8%)

Note. Blank cells indicate that the decision alternative was not allowed for subjects under the condition.

not permit direct conclusions about individual decision rules, since for each individual only one data point is generated—there is no individual overlap over conditions. Thus, any conclusions about the consequences of verdict option restriction must be based on inferential chains of reasoning from the available aggregate data. Vidmar (1972, p. 215) hypothesized that "under conditions of restricted decision alternatives, the more severe the degree of guilt associated with the least severe guilty alternative, the greater were the chances of not obtaining a guilty verdict." In the Vidmar experiment, severity of guilt was implicitly defined in terms of the length of sentence attached to each verdict, rather than in terms of some elusive notion of degree of guilt implied by each. In terms of verdict/sentence severity, the ordering is F most severe, S next most severe, then M, and then, of course, acquittal.

Let us designate p_{xi} as the proportion of cases in which we obtain a verdict of x (=F, S, M, or N) in condition i. Vidmar's hypothesis may be reformulated as the unconditional assertion that

$$p_{N1} > p_{N2} \geqslant p_{N3},\quad p_{N1} > p_{N4},\quad p_{N2} > p_{N5},\quad p_{N2} > p_{N6},\quad p_{N2} > p_{N7},\quad \text{and}\quad p_{N4} > p_{N5}. \tag{1}$$

As we can see from Table 1, all but one of these inequalities hold, and the one minor discrepancy is readily attributable to random error. Thus, Vidmar's hypothesis seems very well supported (for data on the results of tests of significance, see Vidmar, 1972). It is consistent with other literature (e.g., Kerr, 1978) which shows that conviction rates (for individual mock jurors) are inversely related to the severity of the prescribed penalty.

TABLE 2
Feasible Preference Orderings in the Four-Alternative Case

1	FSMN	7	SFMN	13	MSFN	19	NSMF
2	FSNM	8	SFNM	14	MSNF	20	NSFM
3	FMSN	9	SMFN	15	MFSN	21	NMSF
4	FMNS	10	SMNF	16	MFNS	22	NMFS
5	FNSM	11	SNFM	17	MNSF	23	NFSM
6	FNMS	12	SNMF	18	MNFS	24	NFMS

However, Vidmar's hypothesis on the relationship between severity of the least severe guilty verdict and the probability of acquittal seems remarkably inconclusive in that it predicts only 11 of the 21 possible paired comparisons among p_{Ni} values and predicts nothing about the results of paired comparisons between p_F, p_S, and p_M values (i.e., within-rows cross-column comparisons). Moreover, it does not enable us to make any within-column predictions.

Let us consider some alternative ways of dealing with Vidmar's (1972) data. One approach is to postulate that all jurors' preferences are strongly ordered. There will be 4! (24) such orderings in the four-alternative case. These are specified in Table 2.

If each group of 24 jurors in the Vidmar experiment is assumed to be drawn from the same population, then, if we denote the proportion of that population holding preference ordering i as x_i, we may solve the set of independent simultaneous equations given in Eq. (2) to estimate the percentage of members of that population holding any given preference ordering. Equations (2) are derived from Table 1; numbers are expressed in percentages.

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{13} + x_{15} + x_{16} = 46 \tag{2a}$$

$$x_1 + x_2 + x_3 + x_7 + x_8 + x_9 + x_{10} + x_{11} + x_{12} + x_{13} + x_{14} + x_{15} = 84 \tag{2b}$$

$$x_1 + x_3 + x_4 + x_7 + x_9 + x_{10} + x_{13} + x_{14} + x_{15} + x_{16} + x_{17} + x_{18} = 92 \tag{2c}$$

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_{15} + x_{16} = 8 \tag{2d}$$

$$x_7 + x_8 + x_9 + x_{10} + x_{11} + x_{12} + x_{13} + x_{14} = 92 \tag{2e}$$

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 29 \tag{2f}$$

$$x_9 + x_{10} + x_{13} + x_{14} + x_{15} + x_{16} + x_{17} + x_{18} = 67 \tag{2g}$$

$$x_1 + x_2 + x_7 + x_8 + x_9 + x_{10} + x_{11} + x_{12} = 46 \tag{2h}$$

$$x_3 + x_4 + x_{13} + x_{14} + x_{15} + x_{16} + x_{17} + x_{18} = 54 \tag{2i}$$

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 8 \qquad (2j)$$

$$x_7 + x_8 + x_9 + x_{10} + x_{11} + x_{12} = 63 \qquad (2k)$$

$$x_{13} + x_{14} + x_{15} + x_{16} + x_{17} + x_{18} = 21 \qquad (2l)$$

$$\sum_{i=1}^{24} x_i = 100 \qquad (2m)$$

$$x_i \geq 0 \quad \text{for all } i. \qquad (2n)$$

This approach does not appear promising since we have only 13 equations in 24 unknowns. Even if we design an optimal experiment, this approach remains unpromising. For $n = 4$, there are 11 possible conditions, of which Vidmar (1972) used only 7; but even the full 11 conditions give rise to only 18 independent equations, only 6 more than the 13 we obtained from Vidmar's data and still not enough to solve for 24 unknowns. More generally, for any N, the ideal verdict option experiment generates $\Sigma_{K=2}^{N} \binom{N}{K}$ conditions and $\Sigma_{K=2}^{N} (K-1)\binom{N}{K} + 1$ independent equations, and as N increases the gap between $\Sigma_{K=2}^{N} (K-1)\binom{N}{K} + 1$ and $N!$ increases. If we require that all conditions include acquittal as one of the verdict options, we will be even further from generating enough independent equations to solve for the $N!$ unknowns.

Since merely assuming strong orderings is of limited value, let us consider a model which puts restrictions on the possible strong orderings so as to reduce the number of unknowns sufficiently to guarantee at least as many equations as unknowns.

Consider the continuum of alternatives *FSMN*, ordered with respect to severity of punishment. Consider an individual whose most preferred outcome is at some point on the F–N continuum. Let us posit that in any choice among alternatives, whether pairwise or not, that alternative is preferred which is closest (in utility) to the individual's most preferred outcome. If all individuals in the population judge alternatives vis-à-vis relative severity and if each sees a first-degree verdict as more severe than a second-degree verdict as more severe than a manslaughter verdict as more severe than a verdict of acquittal, then each individual will have a preference ordering which is one of the eight shown in Fig. 1. The actual utility assignments (i.e., the desirabilities to the juror of each of the four verdicts) are irrelevant for present purposes; all that matters is the preference ordering.

The preferences in Fig. 1 are single peaked with respect to the F–N continuum. A graph is single peaked if it changes its slope at most once, from up to down. A set of preference orderings is said to be single peaked if the preferences of all individuals in the group can be graphed as single peaked curves with respect to some underlying continuum (Black, 1958; Arrow, 1962; Grofman, 1969). A sufficient condition for single peakedness to occur is the existence of a continuum along which alternatives are perceived as being ordered, such that each individual chooses that available option which is closest to the point on the continuum that represents what Coombs (1964) refers to as his "ideal" point. Single-peaked orderings may be understood as the union of Coombsian *I* scales for some underlying qualitative *J* scale (Coombs, 1964).

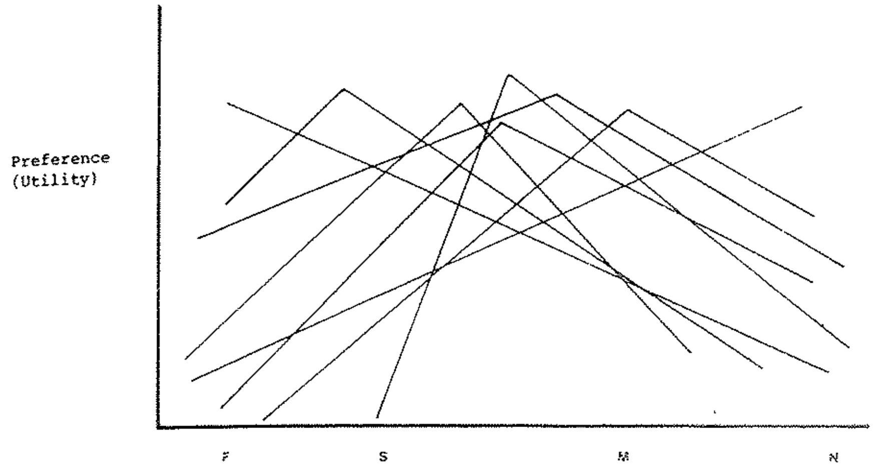

FIG. 1. The eight possible single-peaked preferences schedules in the four-alternative case.

On the assumption that all preference orderings are single peaked with respect to the continuum *FSMN* (or equivalently with respect to *NMSF*), we may use Vidmar's (1972) data to solve for the best fitting x_i values to predict cell frequencies in each of the seven conditions. Table 3 gives

TABLE 3
Equations to be Fitted on Assumption that Preferences Are Single Peaked

$x_1 + x_7 + x_9 + x_{13} = 46$	(2a)
$x_1 + x_7 + x_9 + x_{10} + x_{13} + x_{14} = 84$	(2b)
$x_1 + x_7 + x_9 + x_{10} + x_{13} + x_{14} + x_{17} = 92$	(2c)
$x_1 = 8$	(2d)
$x_7 + x_9 + x_{10} + x_{13} + x_{14} = 92$	(2e)
$x_1 + x_7 = 29$	(2f)
$x_9 + x_{10} + x_{13} + x_{14} + x_{17} = 67$	(2g)
$x_1 + x_7 + x_9 + x_{10} = 46$	(2h)
$x_{13} + x_{14} + x_{17} = 54$	(2i)
$x_1 = 8$	(2j)
$x_7 + x_9 + x_{10} = 63$	(2k)
$x_{13} + x_{14} + x_{17} = 21$	(2l)
$x_1 + x_7 + x_9 + x_{10} + x_{13} + x_{14} + x_{17} + x_{24} = 100$	(2m)

TABLE 4
Best-Fitting Solution to Vidmar Data Given Single-Peakedness Assumption

Preferences	x_1	x_7	x_9	x_{10}	x_{13}	x_{14}	x_{17}	x_{21}
Estimated proportions	8	21	4	38	13	0	8	8

Note. Calculated by a generalized least-square minimization technique (BMD Statistical Analysis Package).

the equation set used to determine the x_i. This equation set is overdetermined, having fewer variables than equations. The best fitting x_i values under the single-peakedness constraint are given in Table 4. These values are used to predict cell frequencies, with results as shown in Table 5.

The assumption of single peakedness fits perfectly or near perfectly for five of the seven conditions. Clearly, if we were able to modify the model to account, in particular, for the derivations from predicted values in Condition 6, we would be able to obtain almost perfect predictive accuracy. Of course, the single-peakedness assumption is being only indirectly tested, since we do not have data on individual preference orderings.

So far, we have been assuming that the introduction of new verdict alternatives does not affect jurors' underlying preferences. For example, under the single-peakedness assumption, if first-degree murder is made available as an option, those with preferences FSMN will choose it over whichever other available option they previously most preferred. However, the introduction of the first-degree murder option is not expected to affect the percentage of individuals in the juror population who hold preference ordering FSMN or to affect the percentage of jurors who hold any other preference ordering, for that matter. This assumption may

TABLE 5
Expected Cell Frequencies Using the Single-Peakedness Model

Alternative	1	2	3	4	5	6	7
First-degree murder	46% 87.0	—	—	8% 1.9	29% 7.0	—	8% 1.9
Second-degree murder	—	84% 21.2	—	76% 18.4	—	71% 17.0	63% 15.1
Manslaughter	—	—	92% 22.1	—	63% 15.1	21% 5.0	21% 5.0
Not guilty	54% 13.0	16% 3.8	8% 1.9	16% 3.8	8% 1.9	8% 1.9	8% 1.9
Total discrepancy from Vidmar data	0%	0%	0%	32%	8%	65%	0%

not be an accurate one. Restrictions on the available set of alternatives may also give rise to an important "anchoring effect" (Parducci, 1963; Sherif and Sherif, 1967) with which we have not dealt—in which the presentation of an extreme alternative shifts preferences among nonextreme alternatives in such a way as to increase preferences for the verdict(s) closest to the extreme.

Thus, if we look at Table 1, we see that when first-degree murder was available as an option (Condition 7), more subjects chose second-degree murder as their verdict than chose manslaughter. In Condition 6, on the other hand, more subjects chose manslaughter than chose second-degree murder. It is impossible to reconcile this reversal with the assumption of no shift in preference orderings between the two cases. We conjecture that such an anchoring effect does not affect the single peakedness of preference orderings but merely results in a shift of preferences between orderings. In the four-alternative cases, we propose that the shift would be from MSFN (x_{13}) and MSNF (x_{14}) to SMFN (x_9) and SMNF (x_{10}). Looking at Table 3, we see that the relevant equations for Column 6 are (2h) $x_1 + x_7 + x_9 + x_{10} = 46$ and (2i) $x_{13} + x_{14} + x_{17} = 54$; while for Column 7, they are (2k) $x_7 + x_9 + x_{10} = 63$ and (2l) $x_{13} + x_{14} + x_{17} = 21$. Thus, Condition 7 does, as conjectured, appear to generate a higher proportion of individuals with preferences SMFN or SMNF. While we find this argument a strong one, additional experimental evidence (preferably direct evidence based on ascertaining individual preference ordering) is highly desirable.

When we delete the equations drawn from Condition 6 from our equation set and solve the remaining set of 10 equations in eight unknowns, we obtain the same solution as before; however, the fit is tremendously improved.

The single-peaked model has another attractive feature—it subsumes the Vidmar hypothesis.

Result 1. The assumption of a single-peaked ordering along the severity continuum *FSMN* implies Vidmar's (1972) hypothesis that "under conditions of restricted decision alternatives, the more severe the degree of guilt associated with the least severe guilty alternative, the greater the chances of obtaining a guilty verdict"; i.e., the single-peakedness assumption implies that the inequalities of 7 must hold.

Proof. If $P_{N1} > P_{N2} > P_{N3}$, then we must have $x_{10} + x_{14} + x_{17} + x_{21} > x_{17} + x_{21} > x_{21}$. For nonzero x's this result must always hold. Analogous results are readily obtained for the other inequalities in 7.

Vidmar's hypothesis is directly supported by the data; but, of course, the assumption of single peakedness is considerably stronger than Vidmar's (1972) hypothesis. It may be used to predict the directionality of all paired comparisons among the P_{Ni} and all but one of the remaining within-row between-column pairwise comparisons. Moreover, it can be used to make paired-comparison predictions for conditions not utilized in the Vidmar (1972) experiment, e.g., for decision making in which the alternative of not guilty was not available.

Unlike the assumption of strong orderings, the single-peakedness assumption can be tested indirectly by the simultaneous equation technique described above for all values of $N (N \geqslant 3)$, provided a sufficient number of experimental conditions are run. For $N \geqslant 3$ the number of orderings compatible with the assumption of single-peakedness along a severity continuum is 2^{N-1}, and this number is always less than the number of independent equations we can generate from the data obtained by restricting jurors to all possible choice sets of size k ($k = 2, N$). For example, for $N = 3$, we can generate 6 equations in 4 variables; for $N = 4$, we can generate 18 equations in 8 variables; for $N = 5$, 49 equations in 16 variables; for $N = 6$, 130 equations in 32 variables, etc. Even if, like Vidmar (1972), we confine ourselves to conditions in which not guilty is an available verdict, we will still always be able to generate more equations than we have variables to solve for. Given N verdict options, there will be exactly $\Sigma_{K=1}^{N-1} K \binom{N-1}{K} + 1$ independent equations when we require acquittal to be an available verdict. Thus, for $N = 3$, we can generate 5 equations in 4 unknowns; for $N = 4$ we can generate 13 equations in 8 unknowns (as did Vidmar (1972)); for $N = 5$ we can generate 33 equations in 16 unknowns, etc.

Lee Hamilton (1976, personal communication) has kindly made available to us unpublished data from an experiment similar to Vidmar's (1972), in which there are three verdict options, not guilty (N), unpremeditated murder (U), and premeditated murder (P), and two conditions, NUP and NP. Let Z_1 = NUP, Z_2 = UPN, Z_3 = UNP, Z_4 = PUN, Z_5 = NPU, and Z_6 = PNU. The Hamilton data are shown in Table 6.

Solving the appropriate set of five simultaneous equations in four unknowns, we obtain a *perfect* fit to the Hamilton (1976) data $Z_1 = 11$, $Z_2 = 6$, $Z_3 = 9$, and $Z_4 = 3$.

DISCUSSION

The model which postulates single-peaked preferences along some underlying continuum (e.g., severity) both subsumes and extends the Vidmar (1972) model. Its theoretical superiority to the Vidmar model seems obvious, and its additional predictions fit the data quite well, particularly when the discrepancies in accounting for the data in Condition 7 of Table 1 are accounted for in terms of an anchoring effect. Where alternatives may be compared with respect to some basic underlying dimension (e.g., sentence severity, cost, productivity, position on some ideological continuum, etc.), the single-peakedness model affords theoretical parsimony and intuitive plausibility.

TABLE 6
Hamilton Experiment: Verdict Preferences by Verdict–Option Conditions ($N = 58$)

	N	U	P
Two-option condition	.70 (29)	—	.31 (9)
Three-option condition	.38 (11)	.52 (5)	.10 (3)

The single-peakedness model is also important in that it gives rise to two significant policy implications. First, given single-peakedness preferences over verdict/sentence options, some jurors will refuse to vote for conviction if the punishment option associated with a given verdict is too severe *even if they believe the defendant guilty of a crime*. Thus, restricting verdict options or getting very high mandatory sentences for some crimes may in fact reduce the total number of man years of sentences administered—a boomerang effect! Something quite similar to this appears to have happened with New York's harsh drug laws.

Second, single-peakedness preferences across a profusion of verdict options create a virtual certainty that the jury will not be deadlocked. Black (1958) has shown that when preferences are single peaked, there always exists one alternative which can receive a majority in paired contests against each and every other alternative. In the case of the population dealt with by Vidmar (1972), as we reconstructed their preferences, manslaughter is that alternative. Thus, since available empirical data (see review in Grofman, 1976) suggest that the option which is preferred by a majority will become the unanimous verdict of the jury over 90% of the time, we would expect that juries drawn from Vidmar's subject population would almost always reach manslaughter as their unanimous verdict (at least in the absence of our as yet only conjectured "anchoring effect" whose impact would be such as to shift preferences so as to make second-degree murder the most likely jury verdict). Thus, manslaughter should be expected as a unanimous verdict in actual jury deliberations among subjects drawn from Vidmar's juror pool *far* more often than might appear to be the case from the data on individual jurors presented in Table 1, since with all four options available, manslaughter is the *first choice* of only 21% of the jurors.

Finally, the simultaneous equation technique outlined in this paper may prove to be of some general use as a means of drawing inferences about individual preference parameters from aggregate (experimental) data (cf. Goodman, 1953; Grofman, Migalski, and Noviello, 1985).

REFERENCES

Arrow, K. (1962). Social Choice and Individual Values, 2nd ed., Wiley, New York.
Black, M. (1958). Theory of Committees and Elections, Cambridge Univ. Press, New York/London.

Coombs, C. (1964), A Theory of Data, Wiley, New York.

Goodman, Leo A. (1953), "Ecological Regression and the Behavior of Individuals," American Sociological Review **18**(6) (December), 663–664.

Grofman, B. (1969), "Some notes on voting schemes and the will of the majority," Public Choice **7** (Fall), 64–80.

Grofman, B. (1976), "Not necessarily twelve and not necessarily unanimous," *in* Psychology and the Law: Research Frontiers (G. Bermant, C. Nemeth, and N. Vidmar, Eds.), pp. 149–168, Lexington Heath, Boston.

Grofman, B. N., Migalski, M., and Noviello, N. (1985, in press) "The totality of circumstances test in Section 2 of the Voting Rights Act: A social science perspective." Law and Policy.

Hamilton, L. (1975), The Crime of Disobedience: Jury Simulation of a Military Trial, unpublished dissertation, Harvard University.

Kerr, N. L. (1978), "Severity of prescribed penalty and mock juror's verdicts," Journal of Personality and Social Psychology **36**(12), 1431–1442.

Larntz, K. (1975), "Reanalysis of Vidmar's data on the effects of decision alternatives on verdicts of simulated jurors," Journal of Personality and Social Psychology **31**(January), 123–125.

Luce, R. D. (1959), Individual Choice Behavior, Wiley, New York.

Parducci, A. (1963), "Range frequency compromise in judgment," Psychology Monograph **77**.

Sherif, M., and Sherif, C. W. (1967), "The own categories procedure in attitude research," *in* Readings in Attitude Theory and Measurement (M. Fishbein Ed.), pp. 190–198, Wiley, New York.

Vidmar, N. (1972), "Effects of decision alternatives on the verdicts and social perceptions of simulated jurors," Journal of Personality and Social Psychology **22**, 211–218.