Voting Designed for Better Decisions

A comprehensive, logical approach to voting should result in better decisions.

By Roy A. Minet

Although the mechanism of voting is employed to make major, sometimes momentous, decisions, surprisingly little serious attention has been paid to optimizing the process for the "quality" of the decisions that are made. Despite the great importance of voting to the form of government it establishes, the United States Constitution barely mentions voting and does not specifically dictate any method. Majority votes with a quorum in the House and Senate are mentioned; the rest, most notably public elections, is left up to the state legislatures.

The tacit assumption has generally been made, especially within the United States, that the plurality method should be used. The simplicity of plurality (the option receiving the largest number of votes prevails) surely had something to do with its choice as well. When there are only two options, this reduces to a majority vote and works acceptably, except for the problems of weighting and insincere voting (discussed later). However, it has been widely acknowledged(1,3,4) for more than two centuries that the plurality method is seriously flawed. When one of more than two options must be chosen, the possibility then arises that the winning option may not have received a majority of the votes and, in fact, may not be preferred by a majority of the voters. A recent reminder of this was the US Presidential election of 2000 when the top two candidates received nearly equal shares of the popular vote, but neither had a majority. Plurality provides no way to re-allocate the votes that went to other candidates so as to ascertain who would rightfully achieve a majority.

Some 40 or 50 different voting system variants have been proposed(3,4) over the years and attempts have been made to define, analyze, measure and compare various methods on the basis of their "fairness." Fairness has turned out to be an unexpectedly slippery concept. About a dozen fairness criteria have been proposed. Each seems to be a very desirable characteristic for a voting procedure to possess. However, as Kenneth Arrow showed in 1951(2), while voting systems can be designed to meet various subsets of these criteria, no system is possible that would be judged fair by all of them.

Most would readily agree that "fairness" is highly important. However, I submit that it is probably *not* as important as the "quality" of the decisions that are made. Whoa!!! If the concept of "fairness" is slippery, defining and/or measuring "quality" is at least slippery squared. We don't even want to go there. Fortunately, it may be possible to start at the beginning, step through the problem logically and arrive at a persuasive result without ever needing to deal very directly with "quality."

<u>The Beginning</u> – Voting is a method for collecting the desires of a group of individuals as inputs and processing them to produce an output that is a choice between or among two or more alternatives. Voting can be used to select the President of the United States or to designate the "class member most likely to succeed." This discussion addresses those uses of voting where:

- The decision will or can have a significant impact upon all or many of the group members;
- The group is bound legally or by prior agreement to accept and live by the decision.

In any non-frivolous application of voting, it should be apparent that the most important consideration regarding any given voting issue is that the choice made should be the best one for the overall well being of the group. Simply stated, **the primary objective of voting on any issue is to make the best possible choice**. That having been said, there are some other quite important overall characteristics that voting systems should possess:

- Secret Ballot – Voters should have the right to cast their ballots privately, free from force or coercion;
- Auditability – It should be possible to verify counts and results so that confidence is high that each input is being properly taken into account and that any fraud will be detected;
- Transparency – Electors should be able to understand the voting methodology and how the process works.
- Fairness – A large majority of electors should feel that the process is "fair" (by whatever standards they may hold to be critical);
- Openness/participation – The process is easily accessible to a substantial percentage of electors so that decision making power cannot gravitate towards a small group or an individual.

A voting system lacking these characteristics is not likely to be widely accepted and supported, making its long-term viability questionable.

<u>Who Should Vote</u> – The modern socio-political climate seems to be trying hard to misdirect our attention. A prime example is the oft heard old saw, "It doesn't matter how you vote; the important thing is just to vote." Well, *of course* it *does* matter how you vote. If it did not matter, there would be no point in holding an election; choices could be made by some random selection technique such as flipping coins.

Perhaps there is an individual or a small number of individuals considered by everyone to be best qualified to decide a certain issue. It would seem wise for everyone else to abstain and leave the decision to those best qualified to make it. This surely happens to some extent in current practice when some voters skip making selections in races about which they are not sufficiently informed. Clearly, the quality of a decision is not necessarily improved and could be degraded if everyone votes.

Assume for the sake of discussion that every member of the voter population has exactly the same knowledge, experience, intelligence, abilities, wisdom, judgment and other factors that would contribute to making the right choice on a given issue. Were this to be the case, each voter might be expected to make the same choice on that issue. In reality, overwhelmingly large majorities of voters rarely, if ever, all choose the same option regarding any serious issue. There could be two possible explanations:

1. The level of ability to make this choice that all voters possessed was not good enough to enable them to clearly see the correct decision. That is, flipping a coin to decide this issue might have about the same efficacy as voting on it. Flipping a coin could conceivably even be *better* than voting in some cases where there could be a statistical

tendency for the voters (being "unqualified") to be systematically misled into making a bad decision.
2. All voters really do *not* have the same qualifications, wisdom and ability to make this particular decision, so they reach different conclusions about the options.

In the messy real world, a nasty combination of the two explanations is surely in play. Voters actually do have a wide range of abilities to make good decisions. Those with the best judgment on some issues should tend to have the better judgment on all issues, but wide issue-to-issue variations are a virtual certainty. Also, the difficulty of making the best choice can vary widely from one issue to another. Frequently, voters must predict the future and the future behaviors of various candidates (!) in order to make a wise choice. On a small percentage of issues, a large percentage of voters are able to reliably identify the best option. On a large percentage of serious public issues, even the "best qualified" voters cannot make the best choice with a high confidence. In some cases, there could even be a statistical tendency for the "less qualified" voters to be misled into making bad decisions that offset votes for the "right" decision cast by the "more qualified" voters.

Although hard data may not be conveniently available to support this conjecture, it seems quite probable that the quality of voting decisions could be improved if it were possible to select the "better qualified" voters to make the choices. Because even the most qualified voters cannot make the best choices with a high confidence in many situations, it would not be wise to narrow the number of voters too severely; also, the openness characteristic must be considered in this regard. The optimum might be in the neighborhood of selecting the "most qualified" 40% to 50% of voters to vote on each issue. Such a "goal" is useful in thinking about the problem and the process, whether or not a practical and acceptable way can be found to achieve it.

Selecting the most qualified voters is certainly not a new idea. For many decades, voter qualification tests were used, ostensibly in an attempt to accomplish exactly that objective. Unfortunately, these tests were subverted by some to accomplish other purposes such as discrimination against minorities. Because of persistent abuses, voter qualification tests were essentially prohibited by the Voting Rights Act of 1965. However, the fact that abuses occurred does not mean that there is anything wrong with the basic concept of selecting the most qualified voters. What is needed is an unassailably fair way to accomplish this that cannot be corrupted. Obviously, these two requirements will have to take precedence, even if the efficacy of possible selection mechanisms is limited by them.

One seemingly simple step that could be taken is to completely remove the names of candidates from the actual balloting process. That is, all votes would be write-in votes for selections involving candidates. This introduces an elegant voter qualification test that requires only that the voter know for whom s/he intends to vote. It would be hard to argue that this one question is not always both important and highly relevant. There could be no possible bias for or against any class of voter, except those who are not knowledgeable (or at least somewhat informed) about each specific contest. This would cleanly, decisively, fairly and surgically screen out at least one group of voters who clearly should not be voting. It does not require a separately administered test and it even acts specifically to each contest. It would also be virtually impossible to subvert or corrupt this simple and direct mechanism.

Unfortunately, two objections spring to mind rather quickly. First, the burden of entering each candidate's name accurately and completely would slow the voting process significantly. Second, and much more important, error rates would increase and a bias would be introduced against candidates who happen to have long names or difficult spellings. (Consider the prospect of having Congress made up of 535 people all named Smith and Jones.) Both objections can be dispatched by providing an official alphabetized (or randomized) and numbered list of all candidates together. The voter would have the option of entering a name from the list, entering a name not on the list, or simply entering the number of a candidate from the list. Although the list would provide quite a "cheat sheet," the voter would still need to know which candidates are running for which offices; much of the efficacy of the mechanism would be preserved. This would seem worth implementing regardless of whether the benefit is small or large since there is essentially no downside to doing so. It may be very difficult to devise a voter selection technique that would be more effective and still be unassailably fair and inherently un-corruptible. Of course, voters could take information or notes into the booth with them, but this would at least require some minimum level of interest, comprehension and effort.

<u>An Intractable Problem</u> – Suppose there is a referendum on an issue that 51% of the voters favor and 49% oppose. By majority rule, it obviously will pass. However, further suppose that those voting for the measure really don't care about it very strongly at all, but the 49% who oppose it do so vehemently. Is passing the proposal a good overall decision? Probably it is not. Here we are concerned both about the fairness of the decision and about its quality, the two being entwined.

There have been a number of proposals which attempt to resolve the problem by allowing voters to weight their votes in proportion to the strengths of their convictions, say on a scale of 1 to 100, instead of indicating a simple yes/no vote. If 51 individuals favor the proposal and their average weighting is 10, while 49 people oppose it and their average weighting is 60, the measure would be overwhelmingly defeated (51 x 10 = 510 for and 49 x 60 = 2,940 against). Of course, all such proposals have zero practical value and serve only the purpose of discussion; even the least qualified voters will vote insincerely, realizing that a weight of 100 (or whatever the maximum weighting may be) gives them the most influence over the outcome without any penalty to them.

Is there any way to obtain inputs from voters that are reasonably accurate measures of their passion for each issue? The only conceivable, workable way is to require that they actually give up something of value, which they will only do in proportion to the strength of their convictions. The standard of value is money, so requiring payment for each vote should do the trick. Suppose each of the 51 voters in the previous example were willing to pay an average of $10 each to have the referendum pass, but each of the 49 opposed to it were willing to pay an average of $60 each to prevent its passage. The problem has been solved.

But wait a minute: won't wealthy people be able to control every election? OK, put a limit of say $25 on what can be spent on a single issue or race. If the limit is high enough, the control of the wealthy is not sufficiently curtailed. If the limit is low enough to solve the rich-guy-control problem, we are mostly back to the original situation where a large fraction of voters will vote the maximum amount all or most of the time. Of course, there is an argument that wealthy voters may correlate reasonably well with the "best qualified" voters and *should* be given more influence over elections. Actually, this view was held by some of the framers of the U.S. Constitution, but whether or not it has any validity, it certainly is DOA in the current political climate.

One more try: each voter receives an equal allocation of voting points, say 20 per issue or office to be voted upon; up to a limit of 100 points may be applied to any given issue or race. The points take on a certain value as a limited (scarce) resource and allow voters to "spend" more points on the races they consider most important at the expense of having fewer points to apply to decisions they think are less critical. Additionally, points might be given a value by allowing unused points to be turned in for money, say $1 per point. This could provide a "social safety net" where those taking the money automatically forfeit their vote – a very important thing to preserving long-term stability. Perhaps some sort of point system merits further thought, but administrative complexity would be high as would resistance to such a change. All of the forgoing serves to amply illustrate the intractability of engineering a way for voters to meaningfully weight their votes.

Voting Methodology – Returning now to the choice of voting procedures, I would like to propose yet another; it is most similar to Baldwin's method[(3)]. Call it Minet Ranked Choice Voting (MRCV). A fundamental problem with the plurality method is that it fails to collect enough information from voters. Without the data, it is impossible to obtain better results. Most proposed alternative methods ask the voter to provide more information and most depend upon obtaining a complete ranking of all the options from each voter for each race. (Substantially all of the mathematics of fairness has also been based upon the lighthearted assumption that complete, valid and meaningful rankings can be gathered from voters.)

In the pesky real world, collecting a complete ranking would be cumbersome and burdensome to voters when there are more than, say, three or four candidates; many may decline to provide this information. Fortunately, the importance of the data and the probability that it will even be needed decreases rapidly after the first choice. Obviously, knowing the first choice is essential (plurality depends upon this *only*). Knowing the second ranked choice adds key information that becomes essential whenever no single option receives a majority of the first choices. On the other hand, deciding which option to rank as fifth, sixth or beyond has a vanishingly small probability of being important; few voters would devote much, if any, time or thought to making those decisions and likely would not know enough about some of the minor candidates to make intelligent ranking judgments. Also, bear in mind that write-in candidates must be accommodated!

The proposed system encourages the voter to provide his or her first, second and third choices only. Each voter may optionally rank three, two, one or no candidates. In the case where there are no more than four options *and* all voters actually do rank three of them *and* there are no write-ins, a complete data set would be available. However, complete information (i.e., the ability to fully compare every pairing of options) is *not* normally available and is intentionally not collected when there are more than four options. There is little hope of coaxing all voters into consistently providing complete data and its validity/quality would be highly suspect even if it could be collected on a forced basis. No practical system should depend for its success upon a full ranking.

After gathering the top three (different) choices from as many voters as will provide them, the following (sometimes) iterative procedure is applied until Step 1 identifies the winner:

1. Total the first choices. If any option/candidate has a majority of the first choices, it is the winner. If only one option remains, it is the winner.
2. Eliminate the "weakest" option employing a weighted point system vaguely reminiscent of Borda[3,1]. Points are assigned for each ballot so that each choice level has twice the points of the next lower level. Thus, when a voter has provided all three choices, weighting would be: first choice 4; second choice 2; and third choice 1. The choice level weightings are determined on the first iteration and are maintained for any subsequent iteration(s), even if some choices have by then been eliminated (i.e., a first choice always counts 4). Total the points for each option across all ballots. Eliminate the option with the lowest point total. If there should be a tie for the lowest point total, eliminate the tied option having the smallest first choice total; if still a tie, eliminate the tied option having the smallest second choice total; if a tie persists, choose a tied option at random to be eliminated.
3. For any ballot whose first choice was eliminated, promote the second choice (if any) to first. For any ballot whose second choice has been eliminated or promoted, promote the third choice (if any) to second. Any eliminated or promoted choice which has no lower choice will cease to exist. Go back to step 1.

As should be evident, this system will always have the same outcome as the plurality method in cases where a specific option initially receives a majority of first choice votes. The weighted point elimination system is employed only if/when no option receives a majority of first choice votes. See Appendix A for a more in-depth discussion of the point weighting system and some comparisons with some of the other methods. Appendix B illustrates how the proposed system deals with each of the examples used in the referenced 2004 Dasgupta and Maskin *Scientific American* article; all cases, including those cited as problematic, are "properly" resolved. The "magic" is successive elimination of the weakest option combined with "promotion" of voters' lower choices to replace eliminated options, making sure that the voters' ranking orders are always preserved.

Any system involving ranking opens the door to possible manipulation through insincere voting techniques. However, susceptibility to such attempts is reduced somewhat by the shallow ranking depth (three levels maximum) and the promotion of lower choices. On the other hand, there is a very important class of voter whose sincerity will actually be improved. Consider an election with two strong candidates, S1 and S2, who are perceived as nearly equally likely to win. There is also at least one weaker candidate, W1. Under plurality, voters who strongly prefer W1 will frequently vote insincerely for either S1 or S2 (the lesser of two evils), because they would rather influence the choice between S1 and S2 than "waste" their votes on W1. Only when allowed to rank their choices will they be free to indicate their true preferences, without penalty, by ranking W1 first and either S1 or S2 second. If W1 actually is eliminated as feared, the second choice is promoted and counted (and thereafter weighted) exactly as though it had been ranked first. Only then can a true picture of the support for such weaker candidates emerge. This is a very critical consideration that strongly impacts decision quality.

<u>Implementation</u> – Increasing the complexity of voting, as any improved system must, begs for more automation. In view of the current state of information technology, this should be embraced as something already overdue. A good implementation can enhance the important transparency and auditability characteristics. While the secrecy of each voter's ballot must be fiercely guarded, everything else needs to be as out in the open and visible as possible.

Appendix C outlines a proposal whereby each voter is automatically assigned a random seven-digit number. (Optionally, the voter might see this number.) S/he would select candidates' names from the displayed official list for the desired first, second and third choices in each race. A name not on the list could always be entered for any choice as a "write-in." At the end of the day, each polling place would generate a human-readable data file (a text file) of all voter inputs sorted by the random voter numbers. Each file would also contain its identification by country, state, county and polling place. As an ultimate backup, voter verifiable paper ballots are automatically printed and then dropped into a ballot box.

The files from all polling places would form the input to a computer that executes the voting algorithm as described above. Both the initial first choice totals, the trail of eliminations (if any) and the final winner of each race would be published. In addition, the text files from all polling places would be readily available (read only) to anyone (e.g., via the Internet). If voters were shown and took note of their assigned random number, they could easily find their inputs and verify that their wishes were properly recorded. Anyone could independently verify the published results by running the algorithm on other computers or by any other means.

Conclusion – No claim is made that these proposals lay out the ultimate and perfect voting system. Further discussion is invited, encouraged and certainly will happen. However, the methodology described does appear to possess better overall characteristics and to exhibit better behavior over a wider range of circumstances than other systems; it is designed to be practical and to work with whatever data voters actually provide in real-world elections and to still be understandable to substantially all voters.

Much good theoretical work has been done that supports understanding and thinking about voting. But there is no substitute for a comprehensive logical approach that also comprehends practical realities and limitations. One inescapable conclusion is that worthwhile improvements over plurality, both in fairness and in the quality of the voting decisions made, could be achieved at any time and would not be at all prohibitively costly.

References

1. The Fairest Vote of All. Partha Dasgupta and Eric Maskin. *Scientific American*, March, 2004.
2. Social Choice and Individual Values. Kenneth J. Arrow, Wiley, 1951. [Yale University Press, 1990]
3. Voting System. http://en.wikipedia.org/wiki/Voting_System
4. Alternative Voting Methods. James Green-Armytage. http://fc.antioch.edu/~james_green-armytage/voting.htm

Appendix A – Method Considerations/Details

Condorcet – The Condorcet[3,1] criterion is widely considered the most important fairness test that voting systems should satisfy. The basic concept of electing a candidate that can command a majority vote when separately paired with each of the other candidates seems compelling, but Condorcet can do some things that many might view as quirky. Consider a hypothetical election where there are two strong candidates that split the top ranking votes equally. Say, 500 voters rank six candidates in two ways:

250 A, C, D, E, F, B
250 B, C, D, E, F, A

The accompanying Dodgeson matrix then appears as:

	A	B	C	D	E	F	Min.	Min.%
A		250	250	250	250	250	250	50.0
B	250		250	250	250	250	250	50.0
C	250	250		500	500	500	250	50.0
D	250	250	0		500	500	0	0.0
E	250	250	0	0		500	0	0.0
F	250	250	0	0	0		0	0.0

As expected, there is a tie between A and B. Because all 500 voters ranked C second there actually is a three-way tie between A, B, and C. Since no candidate has a majority in all pairings, there is no winner. If one more voter casts a ballot that matches either the first or the second ranking order, giving it 251 votes, then, not surprisingly, either candidate A or B will win with a 251 vote minimum and 50.1% which now is a majority. But perhaps that one additional voter really favors candidate F and submits a third ranking resulting in this picture:

250 A, C, D, E, F, B
250 B, C, D, E, F, A
1 F, E, D, C, B, A

The Dodgeson matrix now adjusts as:

	A	B	C	D	E	F	Min.	Min.%
A		250	250	250	250	250	250	49.9
B	251		250	250	250	250	250	49.9
C	251	251		500	500	500	251	50.1
D	251	251	1		500	500	1	0.2
E	251	251	1	1		500	1	0.2
F	251	251	1	1	1		1	0.2

On the strength of one additional voter who ranks C down at fourth, Condorcet now awards the victory to C. Many may protest that a candidate with zero first place votes should not win over candidates that are only one vote short of a majority of first place votes. Part of the problem here (arguably) has to do with vote weighting, or rather the lack thereof. The ranking of one option over another at the bottom of the ranking carries the same weight as it would at the top of the ranking. However, obtaining good weighting information of any kind from voters is an intractable problem.

For reasons discussed in the main article, in an election of any size and complexity, a complete data set just will not be available. Write-in candidates, by definition, do not appear on the ballot and so cannot be ranked by all voters. One could assume that a write-in candidate appearing on some ballots should be ranked last on ballots where it does not appear, and this might be reasonably correct for most voters. However, even this fails with two or more write-in candidates. Thus, the Condorcet method is an interesting logic construct primarily useful in thinking about voting systems. Also, there very well may be some circumstances where it's OK or even good to violate the Condorcet criterion.

Borda – Whether or not they think about it in numeric terms, voters always have weightings of the various options in mind when voting. Ranking the options provides the sort order of the option weightings but no actual values for the weights. That additional information is quite important and would aid greatly in arriving at the best decisions. Just because nobody has come up with a way to collect the data in a meaningful and reasonably accurate way does not mean that its existence should be completely ignored.

The Borda[3,1] method was the first (or at least the best known) attempt to assign weightings and utilize their values to determine the winner. Rankings are simply numbered in reverse order and these weighting "points" are then totaled for each option. The option having the highest point total is declared the winner. Using the same example as used above for Condorcet, candidates would receive points as shown below:

C 2,500 (250 x 5 + 250 x 5)
D 2,000 (250 x 4 + 250 x 4)
A 1,750 (250 x 6 + 250 x 1)
B 1,750 (250 x 1 + 250 x 6)
E 1,500 (250 x 3 + 250 x 3)
F 1,000 (250 x 2 + 250 x 2)

Notice that A and B are still tied, but for third place! C is the overwhelming winner and D comes in second (with or without the additional voter). Borda has little regard for the Condorcet criterion and has a very strong "broad consensus" proclivity. First choice votes don't get the "respect" most people would think they deserve. Some other ways of assigning weighting points have been proposed (e.g., Nauru) in order to tweak weighting characteristics.

Baldwin Method – A few "successive elimination" systems have been proposed and Baldwin's is one of these. Baldwin[3] obtains a full ranking, then the Borda weighting point system is used iteratively to eliminate the weakest (option having the smallest total points). The process is repeated if/as necessary to obtain a winner (which is the last candidate standing). This general approach is very good and arguably produces the best results of the systems discussed so far. Again using the same election example, candidates F and then E are pretty obviously and easily eliminated first. Remaining are:

Rankings		Points 500 Votes		Points 501 Votes	
250	A, C, D, B	C	1,500	C	1,503
250	B, C, D, A	B	1,250	B	1,252
1	D, C, B, A	A	1,250	A	1,251
		D	1,000	D	1,004

Candidate C still garners the most points. But, instead of being a strong second, D now is the next to be eliminated, leaving:

Rankings		Points 500 Votes		Points 501 Votes	
250	A, C, B	A	1,000	C	1,003
250	B, C, A	B	1,000	B	1,002
1	C, B, A	C	1,000	A	1,001

Without the 501st vote, an A/B/C tie persists which would have to be resolved by random choice of a candidate to be eliminated. The additional vote would require that A be eliminated leaving:

Rankings		Points 500 Votes		Points 501 Votes	
250	C, B	C	750	C	752
250	B, C	B	750	B	751
1	C, B				

In the 500-vote case, either B or C would have to be randomly chosen for elimination. In the 501-vote case, B must be eliminated; candidate C is declared the winner, the same result as with Condorcet.

Minet RCV Method – This approach is designed to work well in real-world elections and probably is most similar to Baldwin. Voters are encouraged to rank their top three choices. Any choice may be a write-in candidate not on the ballot. No choice may occur more than once. Voters are not forced to rank all three, but are limited to a maximum of three because:

- It is virtually impossible to obtain full rankings from all voters in an election of any size and/or complexity;
- If collected, rankings after the first three or four are not likely to be high quality (have much validity);
- Rankings below the first two or three are extremely unlikely to be of importance in making the decision, even if they could be collected and relied upon;
- If there are write-in choices, it is not possible for all voters to rank all candidates.

To win an election, a candidate must have a majority of the first choices or (in extreme cases) be the only remaining option after all others have been eliminated. Of course, this so far is not different from Baldwin's "last man standing," but the general public will likely better understand and be more comfortable with this presentation. Especially when a candidate initially has a majority of first choices as will often be the case, points, weightings and eliminations do not come into play and the system is more obviously equivalent to the familiar plurality.

It might be noted here that successive elimination methodologies should be "safer" than techniques that directly pick a winner using points (e.g., Borda). An "error" in eliminating the weakest alternative does not necessarily mean that the wrong winner will be chosen. See the last example in Appendix B for an illustration of this.

The weighting system to be used to eliminate the weakest option when no option has a majority of first choices deserves further thought and analysis. We know, as is obvious from the ranking, that the second choice has a smaller weight than the first. Let us say that “r” is the weight ratio, where $1 > r >= 0$, so that the weight of the first choice, F, times r equals S, the weight of the second choice. There is no way to know the value of r exactly. Clearly, it varies both from voter to voter and from election to election. Occasionally, there may be a close choice, but usually, people have fairly strong opinions, so r probably is seldom close to 1.

It would be nice if a good way to extract reliable weightings from each voter for each candidate in each election could be devised. Until then, the voting algorithm will have to use a fixed way of assigning weights. Perhaps we need an approximation of r that is an average over many voters and many elections. It might well be possible design some statistical experiments that would shed some light on the values of r in real elections.

The weight of the third choice similarly is less than the second. Lacking any better information, assume that the ratio of the third choice weight to the second choice weight is also r, so that rS = T.

Of course, the choice of r will have an important effect on the behavior of the algorithm. Consider a test case where there are two strong polarizing candidates who split most of the first choice votes evenly and a third candidate who receives only a small decimal fraction, v, of first choices. Here are the votes corresponding to each of three rankings:

(1–v)/2	A, C, B
(1–v)/2	B, C, A
v	C, A, B

Since A and B are tied but do not have a majority, we are interested in which candidate the algorithm will eliminate first. Weighting points are calculated as follows:

$A = (1-v)/2 + r^2(1-v)/2 + rv$

$B = r^2(1-v)/2 + (1-v)/2 + r^2v$

$C = r(1-v) + v$

The weight of A must always be greater than that of B for any v and any r, both between 0 and 1. The weightings for B and C will be equal for v and r where

$$r^2(1-v)/2 + (1-v)/2 + r^2v = r(1-v) + v$$

Solving for v yields, $v = (1-2r+r^2)/(3-2r-r^2)$

At r = 0, B and C will have the same weight if C receives exactly one third of the votes. If C receives less than 1/3 of the votes, it is eliminated first and A wins. If C receives more than 1/3 of the votes, then A is tied with B for elimination, but B goes first (by virtue of A having more second choice votes than B) and C will win.

Of course, this behavior is exactly plurality, since at r = 0, only first choices count.

As the value of r closely approaches 1, the fraction of the votes needed by C to exactly tie B's weight closely approaches 0 and behavior becomes more and more "broad proclivity" (like Borda, and beyond). At 1, we have approval voting.

In the middle, at r = 0.5 (as far as possible from both 0 and 1), C will be eliminated first if it receives less than 1/7 (14.29%) of the votes and A wins; B will go down first when C gets more than 1/7 of the votes and C wins. A value of 0.5 would statistically be the best assumption we could make for the relative weights of the first and second choices averaged over all voters and all elections. An r value of 0.5 means that the weight of one vote at any ranking level is exactly equal to the weight of two votes at the next lower ranking level (if one exists). It also means that the integers 4, 2 and 1 can conveniently be used to weight the first, second and third choices respectively. Again, integer weightings are much more likely to be understood by the general public than decimal fractions. Lacking any hard data and lacking any way to obtain reliable data, 0.5 is the most reasonable assumption that can be made for the average of all voters over all elections.

The MRCV method weighting rule is that the points for each choice level are double the points awarded the next lower level. If a ballot ranks all three levels, the 4, 2, 1 weighting applies. If a voter decides to supply only two choices, they receive a 4 and a 2 weighting. A ballot with a single choice is weighted 4.

(Note that one could argue that a small penalty should be imposed if fewer than three choices are supplied by a voter to discourage insincere voting; say subtract a point from the first and second choices if a third choice is not provided and subtract two points from the first choice weighting if it is the only one provided. It could well be argued that an even more severe penalty is appropriate. Most logically, points awarded should be the total of points on all lower levels plus 1. It was decide not to do this for simplicity's sake as well as the debatable value of doing so.)

Finally, applying the MRCV method to the above election example we initially have:

Rankings		Points 500 Votes		Points 501 Votes	
250	A, C, D	A	1,000	A	1,000
250	B, C, D	B	1,000	B	1,000
1	F, E, D	C	1,000	C	1,000
		D	500	D	501
				F	4
				E	2

No candidate has a majority of first choices. Eliminate D for the 500-vote case. In the 501-vote case, E, then F, then D are eliminated. Remaining are:

Rankings		Points 500 Votes		Points 501 Votes	
250	A, C	A	1,000	A	1,000
250	B, C	B	1,000	B	1,000
1	(no choices remain)	C	1,000	C	1,000

No candidate has a majority of first choices. All remaining candidates are tied for the lowest point honor. However, C has the fewest first choice points (none), so it is eliminated next.

Rankings		Points 500 Votes		Points 501 Votes	
250	A	A	1,000	A	1,000
250	B	B	1,000	B	1,000
1	(no choices remain)				

In both cases, we are left with an exact tie between A and B which will need to be resolved by random selection. Note that the lone voter who liked F did *not* cause C to become the winner and did not even resolve the tie between A and B. Under these circumstances, this seems altogether fitting and proper, because that's exactly what it reasonably is: a tie between A and B. We must be careful not to fool ourselves by elevating irrelevant or unreliable data to the same status as the more reliable and important stuff!

However, it should be noted that, when balanced on the razor edge of a tie, one vote surely *can* make the difference. If the lone voter includes A, B and/or C within the top three choices, whichever is ranked highest will then win. Here are a few related examples:

Rankings		Points 500 Votes		Points 501 Votes	
250	A, C, D	A	1,000	A	1,001
250	B, C, D	B	1,000	B	1,000
1	F, E, A	C	1,000	C	1,000
		D	500	D	500
A Wins				F	4
				E	2

Rankings		Points 500 Votes		Points 501 Votes	
250	A, C, D	A	1,000	C	1,001
250	B, C, D	B	1,000	A	1,000
1	F, E, C	C	1,000	B	1,000
		D	500	D	500
C Wins				F	4
				E	2

Rankings		Points 500 Votes		Points 501 Votes	
250	A, C, B	A	1,250	A	1,250
250	B, C, A	B	1,250	B	1,250
1	F, E, C	C	1,000	C	1,001
				F	4
A & B Tie				E	2

Rankings		Points 500 Votes		Points 501 Votes	
250	A, C, D	A	1,000	C	1,004
250	B, C, D	B	1,000	A	1,002
1	C, A, B	C	1,000	B	1,001
		D	500	D	500

C Wins

It appears that r = 0.5 is both logical and results in good behavior. If deeper rankings were to be collected, it would take 8 fourth rank votes and 16 fifth rank votes to equal 1 first choice vote. Even if they could be collected and the data were reliable, the influence of lower rankings fades fast (exponentially); so for this and the other reasons cited, we need not shed many tears over their absence.

Appendix B – Some Examples

Note that all examples are shown with complete data sets, so that the Borda and Condorcet methods can be executed as designed. Bear in mind that incomplete data is a virtual certainty in a real election of any size and complexity. Also, rankings below the second or third choice become increasingly suspect as to their validity, if voters are willing to provide them at all.

2000 US Presidential Election, Scenario A

49% prefer Gore, Bush, Nader, Buchanan
2% prefer Nader, Gore, Bush, Buchanan
48% prefer Bush, Buchanan, Gore, Nader
1% prefer Buchanan, Bush, Gore, Nader

Plurality – Gore wins with 49% of the first choices.

Borda – Bush wins with 346 points.

Gore	Bush	Nader	Buchanan
49 X 4 = 196	49 X 3 = 147	49 X 2 = 98	49 X 1 = 49
2 X 3 = 6	2 X 2 = 4	2 X 4 = 8	2 X 1 = 2
48 X 2 = 96	48 X 4 = 192	48 X 1 = 48	48 X 3 = 144
1 X 2 = 2	1 X 3 = 3	1 X 1 = 1	1 X 4 = 4
----	----	----	----
300	346	155	199

Condorcet – Gore wins with a 51% minimum (that is above 50%).

Gore	Bush	Nader	Buchanan
51% to Bush	98% to Nader	51% to Buchanan	49% to Gore
98% to Nader	99% to Buchanan	2% to Gore	1% to Bush
51% to Buchanan	49% to Gore	2% to Bush	49% to Nader

MRCV – Gore wins with 51% after elimination of Nader.

No candidate initially has a majority of first choices (Gore is closest with 49%), so an elimination is required. From the first three choices ranked, we have:

Gore	Bush	Nader	Buchanan
49 X 4 = 196	49 X 2 = 98	49 X 1 = 49	49 X 0 = 0
2 X 2 = 4	2 X 1 = 2	2 X 4 = 8	2 X 0 = 0
48 X 1 = 48	48 X 4 = 192	48 X 0 = 0	48 X 2 = 96
1 X 1 = 1	1 X 2 = 2	1 X 0 = 0	1 X 4 = 4
----	----	----	----
249	294	57	100

With 57 points, Nader is the weakest and is eliminated. 49% of the voters have their third choice eliminated (they now no longer have a third choice). 2% of the voters have their first choice eliminated, but their second choice is promoted to first and their third choice is promoted to second (leaving no third choice):

49% prefer Gore, Bush
2% prefer Gore, Bush
48% prefer Bush, Buchanan, Gore
1% prefer Buchanan, Bush, Gore

We now have Gore with 49% + 2% = 51% of first choices. Although the election has already been settled, another iteration would boil it down to two candidates (by eliminating Buchanan) and result in Gore 51%, Bush 49%.

49% prefer Gore, Bush
2% prefer Gore, Bush
48% prefer Bush, Gore
1% prefer Bush, Gore

2000 US Presidential Election, Scenario B (IIA Challenge)

(Note that the only change from Scenario A is that the Buchanan/Gore order has been swapped for the 48% Bush voters.)

49% prefer Gore, Bush, Nader, Buchanan
2% prefer Nader, Gore, Bush, Buchanan
48% prefer Bush, Gore, Buchanan, Nader
1% prefer Buchanan, Bush, Gore, Nader

Plurality – Gore wins with 49% of the first choices.

Borda – Gore wins with 348 points.

Gore	Bush	Nader	Buchanan
49 X 4 = 196	49 X 3 = 147	49 X 2 = 98	49 X 1 = 49
2 X 3 = 6	2 X 2 = 4	2 X 4 = 8	2 X 1 = 2
48 X 3 = 144	48 X 4 = 192	48 X 1 = 48	48 X 2 = 96
1 X 2 = 2	1 X 3 = 3	1 X 1 = 1	1 X 4 = 4
----	----	----	----
348	346	155	151

Condorcet – Gore wins with a 51% minimum (that is above 50%).

Gore	Bush	Nader	Buchanan
51% to Bush	98% to Nader	51% to Buchanan	49% to Gore
98% to Nader	99% to Buchanan	2% to Gore	1% to Bush
51% to Buchanan	49% to Gore	2% to Bush	49% to Nader

MRCV – Gore wins with 51% after dropping Buchanan, then Nader.

No candidate initially has a majority of first choices (Gore is closest with 49%), so an elimination is required. From the first three choices ranked, we have:

Gore	Bush	Nader	Buchanan
49 X 4 = 196	49 X 2 = 98	49 X 1 = 49	49 X 0 = 0
2 X 2 = 4	2 X 1 = 2	2 X 4 = 8	2 X 0 = 0
48 X 2 = 96	48 X 4 = 192	48 X 0 = 0	48 X 1 = 48
1 X 1 = 1	1 X 2 = 2	1 X 0 = 0	1 X 4 = 4
----	----	----	----
297	294	57	52

With 52 points, Buchanan is the weakest and is eliminated. 48% of the voters have their third choice eliminated (they now no longer have a third choice); 1% of the voters have their first choice eliminated, but their second choice is promoted to first and their third choice is promoted to second (leaving no third choice):

49% prefer Gore, Bush, Nader
2% prefer Nader, Gore, Bush
48% prefer Bush, Gore, Nader
1% prefer Bush, Gore

We now have Bush with 48% + 1% = 49% of first choices. Since no candidate has a majority of first choices (or is the only candidate), another elimination is necessary.

Gore	Bush	Nader
49 X 4 = 196	49 X 2 = 98	49 X 1 = 49
2 X 2 = 4	2 X 1 = 2	2 X 4 = 8
48 X 2 = 96	48 X 4 = 192	48 X 1 = 48
1 X 1 = 1	1 X 3 = 3	1 X 0 = 0
----	----	----
297	295	105

With 105 points, Nader is the weakest and is eliminated. 49% and 48% of the voters have their third choice eliminated (they now no longer have a third choice); 2% of the voters have their first choice eliminated, but their second choice is promoted to first and their third choice is promoted to second (leaving no third choice):

49% prefer Gore, Bush
2% prefer Gore, Bush
48% prefer Bush, Gore
1% prefer Bush, Gore

Gore now wins with 49% + 2% = 51% of first choices (identical to result of Scenario A demonstrating IIA).

2002 French Presidential Election

30% prefer Jospin, Chirac, Le Pen
36% prefer Chirac, Jospin, Le Pen
34% prefer Le Pen, Jospin, Chirac

French law requires a Chirac/Le Pen runoff that Chirac (actually) won easily.

Plurality – Chirac wins with 36% of the first choices.

Borda – Jospin wins with 230 points.

Jospin	Chirac	Le Pen
30 X 3 = 90	30 X 2 = 60	30 X 1 = 30
36 X 2 = 72	36 X 3 = 108	36 X 1 = 36
34 X 2 = 68	34 X 1 = 34	34 X 3 = 102
----	----	----
230	202	168

Condorcet – Jospin wins with a 64% minimum (that is above 50%)

Jospin	Chirac	Le Pen
64% to Chirac	66% to Le Pen	34% to Jospin
66% to Le Pen	36% to Jospin	34% to Chirac

MRCV – Jospin wins with 64% after elimination of Le Pen.

No candidate initially has a majority of first choices (Chirac is closest with 36%), so an elimination is required. From the first three choices ranked, we have:

Jospin	Chirac	Le Pen
30 X 4 = 120	30 X 2 = 60	30 X 1 = 30
36 X 2 = 72	36 X 4 = 144	36 X 1 = 36
34 X 2 = 68	34 X 1 = 34	34 X 4 = 136
----	----	----
260	238	202

With 202 points, Le Pen is the weakest and is eliminated. 30% and 36% of the voters have their third choice eliminated (they now no longer have a third choice); 34% of the voters have their first choice eliminated, but their second choice is promoted to first and their third choice is promoted to second (leaving no third choice):

30% prefer Jospin, Chirac
36% prefer Chirac, Jospin
34% prefer Jospin, Chirac

We now have Jospin as the winner with 30% + 34% = 64% of first choices.

2000 US Presidential Election, Scenario C (transitivity challenge)

35% prefer Gore, Bush, Nader
33% prefer Bush, Nader, Gore
32% prefer Nader, Gore, Bush

Plurality – Gore wins with 35% of the first choices.

Borda – Gore wins with 202 points.

Gore	Bush	Nader
35 X 3 = 105	35 X 2 = 70	35 X 1 = 35
33 X 1 = 33	33 X 3 = 99	33 X 2 = 66
32 X 2 = 64	32 X 1 = 32	32 X 3 = 96
----	----	----
202	201	197

Condorcet – Gore has the highest minimum (35%), but it is not greater than 50%. Condorcet yields no winner in this case.

Gore	Bush	Nader
67% to Bush	68% to Nader	65% to Gore
35% to Nader	33% to Gore	32% to Bush

MRCV – Gore wins with 67% after dropping Nader.

No candidate initially has a majority of first choices (Gore is closest with 35%), so an elimination is required. From the first three choices ranked, we have:

Gore	Bush	Nader
35 X 4 = 140	35 X 2 = 70	35 X 1 = 35
33 X 1 = 33	33 X 4 = 132	33 X 2 = 66
32 X 2 = 64	32 X 1 = 32	32 X 4 = 128
----	----	----
237	234	229

With 229 points, Nader is the weakest and is eliminated. 35% of the voters have their third choice eliminated (they now no longer have a third choice); 33% of the voters have their second choice eliminated, but their third choice is promoted to second (leaving no third choice); 32% of the voters have their first choice eliminated, but their second choice is promoted to first and their third choice is promoted to second (leaving no third choice):

35% prefer Gore, Bush
33% prefer Bush, Gore
32% prefer Gore, Bush

Gore now wins with 35% + 32% = 67% of first choices.

2000 US Presidential Election, Scenario D (IIA challenge)

51% prefer Bush, Gore, Nader
49% prefer Gore, Nader, Bush

Plurality – Bush wins with 51% of the first choices.

Borda – Gore wins with 249 points.

Gore	Bush	Nader
51 X 2 = 102	51 X 3 = 153	51 X 1 = 51
49 X 3 = 147	49 X 1 = 49	49 X 2 = 98
----	----	----
249	202	149

Condorcet – Bush wins with 51% minimum (that is above 50%).

Gore	Bush	Nader
49% to Bush	51% to Nader	0% to Gore
100% to Nader	51% to Gore	49% to Bush

MRCV – Bush wins with 51% of the initial first choices.

2000 US Presidential Election, Scenario E (IIA challenge)

(Note that the only change from Scenario D is that the Nader/Bush order has been swapped for the 49% Gore voters.)

51% prefer Bush, Gore, Nader
49% prefer Gore, Bush, Nader

Plurality – Bush wins with 51% of the first choices.

Borda – Bush wins with 251 points (note IIA violation).

Gore	Bush	Nader
51 X 2 = 102	51 X 3 = 153	51 X 1 = 51
49 X 3 = 147	49 X 2 = 98	49 X 1 = 49
----	----	----
249	251	100

Condorcet – Bush wins with 51% minimum (that is above 50%).

Gore	Bush	Nader
49% to Bush	100% to Nader	0% to Gore
100% to Nader	51% to Gore	0% to Bush

MRCV – Bush wins with 51% of the initial first choices (same result as Scenario D demonstrating IIA).

Fictitious Five-Candidate Free-for-all (instructive example)

(It will be noted that this "election" has an engineered complexity that causes the methods to produce different results.)

28% prefer A, D, E, C, B
20% prefer B, E, D, C, A
16% prefer C, B, E, D, A
14% prefer D, C, E, B, A
12% prefer A, C, B, D, E
7% prefer E, D, B, C, A
3% prefer E, C, B, D, A

Plurality – Candidate A wins with 28% + 12% = 40% of first choices.

Borda – Candidate D wins with 332 points.

	A	B	C	D	E
28 X	5 =140	1 = 28	2 = 56	4 =112	3 = 84
20 X	1 = 20	5 =100	2 = 40	3 = 60	4 = 80
16 X	1 = 16	4 = 64	5 = 80	2 = 32	3 = 48
14 X	1 = 14	2 = 28	4 = 56	5 = 70	3 = 42
12 X	5 = 60	3 = 36	4 = 48	2 = 24	1 = 12
7 X	1 = 7	3 = 21	2 = 14	4 = 28	5 = 35
3 X	1 = 3	3 = 9	4 = 12	2 = 6	5 = 15
	----	----	----	----	----
	260	286	306	332	316

Condorcet – No winner (no candidate is preferred by more than 50% to each of the other four).

To	A	B	C	D	E
A	-	60%	60%	60%	60%
B	40%	-	73%	49%	52%
C	40%	27%	-	69%	58%
D	40%	51%	31%	-	46%
E	40%	48%	42%	54%	-
	----	----	----	----	----
Minimum	40%	27%	31%	49%	46%

MRCV – Candidate E wins with 52% of first choices after eliminating C, D and A (in that order).

No candidate initially has a majority of first choices (A is closest with 40%), so an elimination is required. From the first three choices ranked, we have:

	A	B	C	D	E
28 X	4 =112	0 = 0	0 = 0	2 = 56	1 = 28
20 X	0 = 0	4 = 80	0 = 0	1 = 20	2 = 40
16 X	0 = 0	2 = 32	4 = 64	0 = 0	1 = 16
14 X	0 = 0	0 = 0	2 = 28	4 = 56	1 = 14
12 X	4 = 48	1 = 12	2 = 24	0 = 0	0 = 0
7 X	0 = 0	1 = 7	0 = 0	2 = 14	4 = 28
3 X	0 = 0	1 = 3	2 = 6	0 = 0	4 = 12
	----	----	----	----	----
	160	134	122	146	138

With 122 points, C is the weakest and is eliminated. After promotion into vacancies created, we then have:

	A	B	C	D	E
28 X	4 =112	0 = 0	0 = 0	2 = 56	1 = 28
20 X	0 = 0	4 = 80	0 = 0	1 = 20	2 = 40
16 X	0 = 0	4 = 64	0 = 0	0 = 0	2 = 32
14 X	0 = 0	0 = 0	0 = 0	4 = 56	2 = 28
12 X	4 = 48	2 = 24	0 = 0	0 = 0	0 = 0
7 X	0 = 0	1 = 7	0 = 0	2 = 14	4 = 28
3 X	0 = 0	2 = 6	0 = 0	0 = 0	4 = 12
	----	----	----	----	----
	160	181	0	146	168

Still, no candidate has a majority and A is still closest with 40%. Another elimination is required and D is now the weakest with 146 points. After eliminating D and promoting into vacancies, we then have:

	A	B	C	D	E
28 X	4 =112	0 = 0	0 = 0	0 = 0	2 = 56
20 X	0 = 0	4 = 80	0 = 0	0 = 0	2 = 40
16 X	0 = 0	4 = 64	0 = 0	0 = 0	2 = 32
14 X	0 = 0	0 = 0	0 = 0	0 = 0	4 = 56
12 X	4 = 48	2 = 24	0 = 0	0 = 0	0 = 0
7 X	0 = 0	2 = 14	0 = 0	0 = 0	4 = 28
3 X	0 = 0	2 = 6	0 = 0	0 = 0	4 = 12
	----	----	----	----	----
	160	188	0	0	224

Still, no candidate has a majority. Candidate A is still closest with 40%. Another elimination is required and A is now the weakest with 160 points. After eliminating A and promoting into vacancies, we then have:

	A	B	C	D	E
28 X	0 = 0	0 = 0	0 = 0	0 = 0	4 =112
20 X	0 = 0	4 = 80	0 = 0	0 = 0	2 = 40
16 X	0 = 0	4 = 64	0 = 0	0 = 0	2 = 32
14 X	0 = 0	0 = 0	0 = 0	0 = 0	4 = 56
12 X	0 = 0	4 = 48	0 = 0	0 = 0	0 = 0
7 X	0 = 0	2 = 14	0 = 0	0 = 0	4 = 28
3 X	0 = 0	2 = 6	0 = 0	0 = 0	4 = 12
	----	----	----	----	----
	0	212	0	0	280

Candidate E now has a majority of the first choices with 52% and is the winner. It might be noted that, if yet another iteration were performed to eliminate candidate B, candidate E would end up with 88% of the first choices. This happens because E is not one of the top three choices for 12% of voters and thus has no promotion path. It is a consequence of incomplete data. In fact, in a situation where there are a large number of candidates, say about 9, *and* there are at least 3 major groups of voters *and* the voter groups are strongly polarized (the groups do not include each others' candidates within their top three choices), it is quite possible that the winning candidate may not have a majority of the first choices. Short of the possibility that the 3 polarized groups go their separate ways, this is still a reasonable result under such circumstances and would match plurality.

The resilience of this method to the order of eliminations is comforting. The MRCV weighting based on the top three choices eliminated C first. The Borda count with complete data would eliminate A first, but this path also arrives at exactly the same conclusion (E wins). Both Borda and MRCV counts indicate B as the second weakest; even if B is eliminated first, the method still converges on E as the winner. For completeness, a tie-break method is provided should more than one candidate have the same lowest total points. Indications are that the rules used to break such ties are not critical to the process in most cases and the probability that the ultimate winner will be affected by them is small.

Appendix C – Implementation

Overview – The crux of the system is the data file produced by each polling place that contains all voter inputs as well as identification information about the polling place and the election. Voter inputs are identified by random ballot numbers assigned to each ballot and (optionally) revealed only to the voter. The voting process is managed by a standard, publicly available piece of software that produces the file and can (later) accept multiple such files as input upon which the voting algorithm is performed to produce the final election results. All files of voter inputs are made publicly available. The files can be used by anyone to independently verify election results. The same software running on inexpensive standard equipment could be used for any election: national, state, local or private. Everything about this system (except the real identity of each individual voter) is completely visible and open to public scrutiny. Machine printed human readable ballots serve as an audit trail and are the ultimate authority on voters' intents. A reliable audit trail is critical as it provides a way to detect and correct any system errors, either accidental or intentional.

Data Files – A standard format must be defined for the data files using a technology that is very widely used and understood. The obvious choice at the present time is XML (Extensible Markup Language). XML files can be built using 7-bit ASCII (American Standard Code for Information Interchange) characters as a subset of UTF-8 Unicode, which can also handle any language in the world. A separate file, also in XML format, called an XML Schema Definition (XSD) is used to define and control the standard format of the data file itself. All of this is based upon bedrock standards that are understood, accepted and used worldwide. There are thousands of tools widely available that facilitate examination of XML files. In the worst case, an XML file can be printed out (or viewed with a text editor) and read (albeit, more tediously) by humans. There also are standard ways (e.g., checksums) that can be utilized to verify that official versions of files (both XML schema definitions and XML data files) have not been modified. Obviously, computers are very happy and comfortable reading and/or writing XML files as well.

An XML Schema definition would be developed for the output file of voter inputs. Once developed, the same Schema definition can be used for any election. Optionally, an XML Schema definition could be developed for a separate type of XML file that would be used to describe an election, its issues, its races and its candidate lists. Such a file could be distributed and loaded to facilitate quick set up of the software for specific elections. Again, once the Schema definition has been developed, it can be used for any election. The Schema definitions would be published and available to anyone.

Software – The programming language for the software should be very widely used and understood. It should also be possible to run the software on almost any computer and have it work identically. The obvious choice at the present time is Java. The Java language was developed in the mid-1990's by Sun Microsystems (now owned by Oracle), has been very widely adopted, is stable and runs identically on substantially any computer regardless of hardware or operating system. An official zipped and checksummed version of the software would be publicly available to anyone. Source code would also be published.

A widely used RDBMS (Relational DataBase Management System) of established reliability should be employed. The RDBMS should implement the SQL92 standard to a reasonably high level. There are many suitable DBMSs available. At the present time, the open source

MySQL RDBMS would be one good recommendation. Another very suitable open source DBMS is “Derby,” an embeddable DBMS that is available from the Apache Software Foundation. For security reasons, an embeddable DBMS is to be preferred.

In addition to the control console used by poll workers, the server software should support from 1 to 99 clients and maintain a detailed log of all operations. Four modes of operation are supported. Only one mode can be operational at any given time. The functionality of each mode is generally defined as follows:

- Jurisdiction and Polling Place Set Up – The locale (country, language, time zone), state, county, postal code, street address and polling place number are maintained in this mode.
- Election Definition – This mode supports definition and set up for a specific election, including issues with accompanying text, candidate races and the candidates. A part of the definition of each race or issue is the “territory” eligible to vote (a specification ultimately resolvable to a specific set of polling places). An XML file for all or part of the election definition may optionally be imported. Validity of the checksum will be verified and its value displayed for any imported file. An XML representation of the current definition of the election can be exported at any time; any file produced will automatically contain the polling place set up information, a timestamp and have a checksum appended.
- Voting – The software is in this mode for the duration of the polling place’s voting hours. When first entering this mode, the database is “cleared” for the beginning of a count. Input is now accepted from the defined clients in each voting booth until the end of the voting period. The server will display a “Vacant” or “Occupied” status for each booth. The number of voters who have completed voting, the number currently voting and the total will also be displayed. When this mode is being shut down at the conclusion of the voting period, the XML file of voter inputs for the polling place is exported with checksum appended. All access to the database other than by one user (which is this program) are locked out while in this mode.
- Tally Votes – Viewing and searching of voter input files is supported in this mode. The checksum is always verified and displayed. Also, the voting algorithm can be executed using one or many voter input files. When running the algorithm, the checksums are automatically verified for any/all input files. The software will determine from the election definition which files are applicable to each issue or race. It will report the results of each race, including the number of polling places reporting out of the total number that should be reporting. A detailed list of reporting and missing polling places should also be available.

Supervision of each voter by the software would work approximately like this:

- The voter initiates voting by means of a “Begin Vote” button. The server uses this exact time in milliseconds and the exact time in milliseconds when the booth was enabled to generate a seven-digit random number for the voter. (Note that pseudo-random numbers cannot be used.) Because the number is truly random, there is a (miniscule) chance that the same number could be generated more than once. Therefore, the database is checked to make sure the new number is unique. If it is not, another number is generated. When a verified unique random number has been obtained, it is (optionally) displayed for the voter until s/he activates the “Cast Ballot” button (terminates the voting session, records the inputs and clears everything from the display). The assigned random number is the *only* information relating to the voter that is recorded or retained in any way. A log should show the ordinal voter number, date and time (to the nearest minute) for when each voter recorded his or her vote, but obviously cannot contain the random number or the voters' choices.

- The voter is presented a list of all issues and races. There is an indicator for each as to whether or not its selection(s) has been made/completed. Any issue or race can be selected to either initially enter, or to change already entered, selection(s) for it.
- When an issue or race is selected, its detail is displayed as a separate screen with any associated text. Some issues may require a simple yes/no selection. Where candidates are involved, the voter may enter a first, second and third choice name. Each name is entered in separate fields for "First Name," "Middle Initial," "Last Name" and "Name Suffix". Any name may be entered or a name can be selected from the consolidated list of candidates. A name selected from the list will automatically fill in the F/M/L/S name fields with that name. The voter can return to the top level list of issues and races at any time by pressing a "Return to Issues/Races Screen" button.
- The voter may end the voting session at any time by means of the "Cast Ballot" button on the Issues/Races screen. If this button is pressed when selections have not yet been made for all issues/races, the voter is warned of this fact and offered the opportunity to return and complete her or his selections or to confirm that s/he really wishes to end the vote. When the session is ended, the voter's inputs are inserted into the database and immediately re-sorted using the assigned random ballot number and the screen is cleared in preparation for the next voter. Even if the voter made no selections at all, the database must record that a voter with the assigned random number did vote. A human and machine readable copy of the completed ballot is printed which the voter verifies before depositing into a ballot box which serves as an audit trail and final authority.

<u>Polling Place Procedures</u> – The polling place should be set up the day before the election. All cables, connectors and equipment are in open view (except for that which must be inside each voting booth). Only one Ethernet cable and one power cable enters each voting booth (hard wired connections only, no RF is permitted). Officials and/or officially appointed observers from opposing factions are permitted to do supervised inspection and testing of the equipment if they so desire. The formidable security challenges of electronic voting are rendered considerably more manageable by keeping the system completely isolated and, of course, by the ultimate control of the paper ballots.

During voting hours, voters are processed thusly:

- Each voter first checks in and is properly identified against the list of registered voters for the polling place. A tally of the number of voters is maintained.
- When a voting booth becomes available the next voter is assigned to it. A check is made that the voter's ordinal number is one greater than the total voters who have voted or are in the process of voting (as displayed). Any discrepancy must be immediately investigated and rectified. A poll worker enables voting in the booth to which the voter has been directed or ushered.
- The voter enters booth, starts vote, enters inputs, ends vote, deposits printed ballot and departs. It should be verified that each voter does indeed press the "Cast Ballot" button and deposit a ballot before leaving. A warning status message would be displayed on the control console when voting is active and there has been no activity for more than, say, a minute.

At the conclusion of voting hours, the XML data file that is produced is copied for each observer and judge of election as well as being forwarded to a central location for publication and processing. Its checksum is also officially recorded. A printed version of the file is also posted at the polling place. Random audits would be conducted or, ideally, each polling place would later verify and certify its XML ballot file against the paper ballots. Turning on the option to print the random ballot numbers on the ballots would make this quick and easy (but, of course, would crack the door to the possibility of vote buying/selling). Machine counting of the paper ballots is also a (somewhat less satisfying) possibility.