



CALTECH/MIT VOTING TECHNOLOGY PROJECT

A multi-disciplinary, collaborative project of
the California Institute of Technology – Pasadena, California 91125 and
the Massachusetts Institute of Technology – Cambridge, Massachusetts 02139

BALLOT DESIGN OPTIONS

R. Michael Alvarez
Caltech

Key words: *ballot design, ballot order, voting instructions*

VTP WORKING PAPER #4
February 2002

Ballot Design Options

R. Michael Alvarez¹
California Institute of Technology
February 17, 2002

Prepared for *Human Factors Research on Voting Machines and Ballot Designs: An Exploratory Study*

1. Introduction

Does ballot design “matter”? Does the design of ballots influence how voters cast their ballots, and thereby affect the outcome of an election?

Anecdotal evidence indicates that ballot design may be a very important factor in American elections. Probably the most well-known ballot design question is the now infamous “butterfly” ballot design, from the 2000 Florida election.² The “butterfly” ballot design was argued to have confused many voters, especially the elderly (who might have had trouble with the visual layout of the “butterfly” ballot) and low-information voters (who might have been misled by poor instructions and cognitive confusion). Unfortunately, while there was a great deal of rhetoric about the potential impact of the “butterfly” ballot design, there has not been anywhere near as much scientific research in to the exact impact such a ballot design might have had nor whom it might have affected most.

But other examples abound where ballot design issues may have played a role in some recent election. An excellent case is the 2001 June mayoral runoff election in the City of Compton, California.³ The Compton City Clerk, in an apparent misunderstanding of California state elections law regarding ballot design, failed to correctly randomize the name of candidates on the runoff election ballot. Following state elections law, the Compton City Clerk requested and used the appropriate randomized list of candidate names in the March 2001 primary election, but he again used the same randomized list for the June 2001 runoff election (according to the court ruling in this case, the Clerk should have requested and used a second randomized list for the runoff election).

¹ Professor of Political Science, Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125. email rma@hss.caltech.edu.

² A full discussion of the “butterfly” ballot is in Richard A. Posner, *Breaking the Deadlock: The 2000 Election, The Constitution, and the Courts*. Princeton, NJ: Princeton University Press.

³ In the interests of full disclosure, the author was an expert witness in the City of Compton’s defense of this case, and provided testimony regarding this ballot design question. The judge’s decision was issued on February 8, 2002 in Los Angeles County Superior Court, and is currently under appeal.

The sitting incumbent mayor, Omar Bradley, was listed second on the runoff election ballot; his challenger, Eric Perrodin, was listed first and won the election by a slim 261 vote margin. The expert witness for Bradley, Jon Krosnick, testified that this incorrect ordering of candidate names on the runoff ballot could have accounted for at least the 261 vote margin, and perhaps many more ballots for Perrodin. The court was convinced by Krosnick's testimony, and on this basis alone, ruled in Bradley's favor, threw out the results of the June 2001 runoff election, and reinstated Bradley as mayor of the City of Compton.

These are just two examples of ways in which ballot design has been argued to affect two recent elections. But they point to two areas of election administration that have been largely neglected by social science. Much more research on these two areas, and other areas, of ballot design are necessary. At a time in which many election jurisdictions are investing considerable sums of money in the purchase of new voting systems, clearly more insight into how ballots are designed is necessary. In the remainder of this essay, I take up a series of what I consider to be the important general topics regarding ballot design. I conclude the essay by outlining some general principles for scientific study of these ballot design topics.

2. Candidate Name Order

The recent City of Compton court case cited above is an important example about the limitations of current scientific knowledge of ballot design impacts on election outcomes. Importantly, there are an enormous variety of procedures for ordering the names of candidates on ballot across the United States. For example, in the State of California, the election code provides strict rules for randomization and rotation of candidate names on statewide ballots, randomization and limited rotation for state legislative races, and only randomization for local races.

For each election cycle, the Secretary of State's office issues a randomized alphabet. In statewide races (for example, gubernatorial races), the randomized alphabet is used in Assembly District 1 (of 80) for the candidate ordering on the ballot; rotation occurs across Assembly Districts, so that in Assembly District 2, the first letter on the original list is moved to last, effectively rotating the candidate list across Assembly Districts. But for a state legislative race (for example a State Senate election), the Secretary of State's random list is utilized for the entire Senate district, unless that district cuts across county lines; if the district does cut county lines, the list is rotated for each successive county the Senate district lies in. But for local races, no rotation is mandated, and rarely occurs in most local races.⁴

⁴ Exceptions to this rule sometimes occur for countywide races in California. In Los Angeles County, the practice is to rotate the list across the various Assembly Districts that constitute Los Angeles County.

Other states, though, have dramatically different election laws regarding candidate name order. One well-studied state is Ohio, where candidate name order is a matter of county election law.⁵ Many Ohio counties randomize and rotate candidate names by precincts.

Rotation and randomization procedures have been instituted in many election jurisdictions to alleviate a phenomenon that political psychologists call a “primacy effect”. This phenomenon has been well-documented in survey methodology, where there is often a noticeable bias for respondents to provide the first response option from a list, especially in low-information situations or for poorly informed respondents.⁶ While the survey methodologists have studied this phenomenon extensively in recent years, it is clear that candidates running for office, and many previous court decisions, have been aware of the possibility of a bias towards the first-named candidate on a ballot, even if they did not pin a precise psychological term to the phenomena.⁷

Miller and Krosnick compiled an extensive literature review of ballot order effects from social science. Of the thirty or so studies they examined, they argue that only two did not suffer from problematic methodological flaws, but “surprisingly, these (two) investigations found no name-order effects at all” (p. 297).⁸ Miller and Krosnick’s study, a recent and thorough examination of the problem, only focused on three counties in Ohio, and found significant name-order effects in 48% of the 118 races they studied. Importantly, Miller and Krosnick found that the variance across the races could be explained by county-specific differences, by whether races were partisan, involved an incumbent, and had substantial media attention.

Unfortunately, election administrators do spend considerable resources in each election trying to minimize or eliminate candidate name-order effects. The costs of rotating candidate names across all of the precincts in a county, for example,

⁵ Joanne M. Miller and Jon A. Krosnick, 1998. “The Impact of Candidate Name Order on Election Outcomes”, *Public Opinion Quarterly*, 62: 291-330.

⁶ Jon A. Krosnick, 1991. “Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys.” *Applied Cognitive Psychology*, 5:213-36.

⁷ There is an extensive literature on ballot order effects, in particular, the potential bias for the first-named candidate on a ballot. One of the most widely cited pieces from this early literature is the short monograph by Henry M. Bain and Donald S. Hecock, 1957, **Ballot Position and Voter’s Choice**, Detroit: Wayne State University Press. In California, the state Supreme Court issued an important ruling that set the stage for the development of the current randomization and rotation system in 1975 in the case of *Gould v. Grubb*, 14 Cal. 3d 661. This case was relied upon extensively in the Compton decision, and involved a challenge to a Santa Monica law that allowed incumbents to be listed first on a ballot; the court ruled that this practice violated the equal protection clause, but left it up to the trial court to determine the exact extent of the exact effect this practice had on Santa Monica elections.

⁸ The two studies cited by Miller and Krosnick as not being methodologically flawed are Robert Darcy, 1986, “Position Effects with Party Column Ballots,” *Western Political Quarterly*, 39: 648-62 and David Gold, 1952, “A Note on the Rationality of Anthropologists in Voting for Officers,” *American Sociological Review*, 17: 99-100.

involves substantial printing costs in counties using paper-based ballots or programming costs where electronic systems are employed. These costs are incurred, furthermore, in spite of social science research that finds only weak evidence for candidate name-order effects and which has not studied at all the impact that different candidate name-order procedures have on possible primacy effects. For example, is it sufficient to rotate in a county across state legislative districts, or should rotation be done across precincts? Or, should randomization and rotation be done for only certain races?

This is clearly an area of ballot design that needs substantial social science research, in short order. Instead of seeing the vast differences in randomization and rotation practices within and across states as a hurdle for scientific study, researchers clearly need to exploit those differences as important variance to use to explain any name-order effects that can be shown in empirical research. Furthermore, controlled laboratory experiments are necessary to examine with precision the impacts of different name-ordering procedures, as well as the set of mitigating factors on name-ordering effects. For example, maybe California's election law has the situation backwards: if it can be shown that name-ordering effects are most likely in non-partisan and low-information contests, then randomization and rotation should be done for local or municipal races, not for statewide contests!

3. Ballot Design for Individual Candidate Races

In addition to the question about the ordering of candidate names on a ballot, there are other important questions about how ballots are designed for each individual race. Most importantly, how useful is it for ballots to contain information other than the candidate's name --- the party affiliation of the candidate, whether a candidate is the incumbent or not, and the occupation of the candidate? Social science has not provided a great deal of insight into the exact importance of these different cues on ballots, despite the presence of a great deal of social science theory about why such cues might be helpful for voter decisionmaking.⁹ But which cues are better to provide for voters, and in which order should they be used? And when it comes to information like the candidate's occupation, who should be given the task of compiling that information --- the election official or the candidate?¹⁰

⁹ The literature about how partisan, incumbency, and other cues can help poorly informed voters is too extensive to even summarize in this context. Much of the relevant literature is summarized in R. Michael Alvarez, 1997, **Information and Elections**, Ann Arbor: The University of Michigan Press.

¹⁰ For example, the 2002 Democratic Official Sample Ballot for the March 5, 2002 Democratic primary in California is instructive. The primary race for Insurance Commissioner, a race where there is no incumbent, involves John Garamendi, and the line below his name in the sample ballot reads "Businessman/Rancher"; in fact, Garamendi had been California's first elected Insurance Commissioner and most recently was Deputy Secretary of Interior in the Clinton administration. His opponents include Bill Winslow ("Attorney/Insurance Consultant"), Thomas Calderon

Furthermore, there are many typographical questions to ask about the design of individual candidate races. What is an appropriate font size to use? Should the candidate names be in bold face --- and what about their partisan affiliation and other potential voter cues? How close should their name be to the place the voter makes her mark? Should an individual candidate race be kept on one page, or should it span facing pages (for example, the “butterfly” ballot)? These are all important ballot design questions, albeit very specific ones. Each deserves careful study, with the idea being the elaboration of clear design principles for ballot construction.

4. Design of Entire Ballots

This is a critical question for the design of both paper-based and electronic voting systems. How should the entire ballot be laid out? How can a ballot be developed that is useable, minimizes voter errors, and is cost effective? This is important in many jurisdictions, as Americans are being called upon to vote on more issues and in more candidate races, thus making for longer and more complicated ballots.

With paper-based voting systems, the layout of the entire ballot can be one of the most important administrative decisions made by election officials. In jurisdictions with paper or optically scanned ballots, officials try to reduce costs by packing long ballots onto single pages, even if they have to use small type face, oversized paper stock, and virtually no instructions. Unfortunately, decisions made to pack a long ballot onto a single page of paper for cost considerations might lead to a higher risk of voter mistakes.

Thus, for paper-based voting systems, scientific studies of voter responses to the layout of the entire ballot are necessary. Technical questions of layout must be examined, including appropriate type faces (especially for voters with diminished eyesight), design for differentiating one race from other races, an appropriate number of columns per page.

The design of the entire ballot, though, is equally important for electronic voting systems, especially “touchscreen” voting systems. In sharp contrast to contemporary paper ballots where many races and ballot measures are packed onto a single page of paper, most touchscreen voting systems are constructed so that each race is on a single screen. A voter thus navigates through the ballot, usually in a linear manner, by viewing one race on a screen and by then touching an icon on the screen to move to the next race or screen.

There has been no research that I am aware of regarding if the page-by-page design of touchscreen voting machines leads to a different voting experience

(“Member, California Legislature”), and Tom Umberg (“Prosecutor”). Are these descriptions used by California voters? Are they informative? Can they be more informative?

relative to the typical paper-based system where all of the races are on a single page. There have been some indications in research that undervote rates, especially on down-ticket races, might be quite high on electronic voting machines.¹¹ However, systematic studies (while necessary) have not been done on the exact differences between the impact of each ballot design.

One common criticism of touchscreen voting layout designs, furthermore, is that they sometimes have poorly designed summary screens. In most touchscreen voting systems on the market today, once a voter has navigated through the entire ballot they can receive a summary screen to examine their entire set of votes before finally casting their ballot. These summary screens vary dramatically between different touchscreen systems; some list only the specific candidates or issues voted on, others highlight undervoted races, and they use different typefaces and colors. Furthermore, exactly how a voter navigates their way back into the ballot to change a vote or undervote also varies between systems.

Each of these ballot layout features in touchscreen systems requires scientific study. To my knowledge, neither election administrators nor election system vendors undertake serious studies of the impact that these design decisions in their touchscreen systems have on the quality of votes cast or on the quality of the voting experience. Also, the basic design of a page-by-page ballot, relative to having the entire ballot on a single page, needs to be studied carefully.

There is a second important question about entire ballot design that deserves further research --- the impact of party-line versus non-party line ballots, and how party-line ballots themselves are designed. The impact of party-line ballots has been researched in political science, both in the context of the Australian ballot reform and in regarding the rise in split-ticket voting.¹² However, much of this research has been historical, and quasi-experimental; there have been no studies that I am aware of that examine the impact of party-line ballots in carefully controlled experimental conditions.

Furthermore, there are vast differences across election jurisdictions that allow for party-line ballots, both in the placement of the party choice on the ballot and in the instructions that are given to voters about how to vote a party-line. These

¹¹ See Caltech/MIT Voting Technology Project, July 2001, *Votign: What Is, What Could Be*. Pasadena, CA: California Institute of Technology.

¹² There is a lengthy literature on party-line ballots, and in particular, on how the Australian ballot reform and the decline in the use of party-line ballots has influenced voting behavior. For representative work, see Walter Dean Burnham, 1965, "The Changing Shape of the American Political Universe," *American Political Science Review* 59: 7-28; Angus Campbell and Warren E. Miller, 1957, "The Motivational Basis of Straight and Split Ticket Voting", *American Political Science Review* 51: 293-312; Jonathan N. Katz and Brian R. Sala, 1996, "Careerism, Committee Assignments, and the Electoral Connection", *American Political Science Review*, 90: 21-33; Jerrold G. Rusk, 1970, "The Effects of the Australian Ballot Reform on Split-Ticket Voting", *American Political Science Review* 64: 1220-1238.

design differences need examination and study, so that their impact on voter errors and the quality of the voting experience can be understood.

5. Voting Instructions

Unfortunately, I am aware of no research on voting instructions. Throughout my own research on voting systems since the 2000 elections, I have encountered no academic research on this important aspect of ballot design. This research vacuum needs to be filled, quickly.

Currently, voting instructions seem to be developed largely by election administrators, sometimes in collaboration with election system vendors, sometimes in collaboration with other interested parties (like voter education groups, or groups representing certain classes of voters with special needs like language concerns or disabilities). Rarely, and possibly never, are proposed voting instructions subjected to any type of experimental or field testing before their implementation.

An important example comes from the 2001 mayoral election in the City of Los Angeles. Working closely with staff from the Los Angeles County Registrar-Recorder's Office (the governmental entity that owns and operates the Votomatic punchcard voting and tabulation machines that are typically then leased to municipalities and other governmental agencies in Los Angeles County for their elections), the Los Angeles City Clerk's Office launched a "Got Chad" voter education campaign. This was an advertising and voter education campaign about how to correctly use the Votomatic punchcard system, for the 2001 mayoral election. The City Clerk's office launched a clever pre-election advertising campaign, a voter education campaign in the mass media, distributed materials to voter education and other groups, developed inserts in the ballot books, and produced posters and instruction cards for polling places.

There is some evidence that the "Got Chad" voter education and instruction campaign was productive in instructing Los Angeles City voters how to use the Votomatic machines. I compared the over- and undervote rates in the 2000 presidential and 2001 mayoral election, in all of the City of Los Angeles precincts; only 16% of precincts had higher overvote rates in 2001, while 22% of precincts had higher undervote rates, in these top-of-the-ticket races. Moreover, in heavily Black precincts, the overvote rate increased in only 10% of precincts, while the undervote rate increased in 5% of precincts; in contrast, in heavily Latino precincts, the overvote rate increased by 8%, and the undervote rate increased by 6%.¹³ Clearly, both over- and undervote rates decreased dramatically

¹³ Heavily Black or heavily Latino precincts are defined as those in the 90% percentile of Black or Latino population. These data are for precinct voters only.

between these two elections, in the same precincts, especially for nonwhite voters.¹⁴

Thus while both common sense and some anecdotes indicate that voter instructions on ballots might lead voters to more effectively interact with voting systems, again there is little scientific work on how and why voter instructions influence the voting process. There are a number of open research topics on voter instructions that deserve careful attention:

1. The effectiveness of pre-election distribution of voting instructions, in ballot books and voter information manuals.
2. The effectiveness of mass media distribution of voting instructions, and of education in voting instructions by interest groups.
3. The wording of voting instructions.
4. Technical issues about voting instructions, in particular, typographical issues (point size, color, etc.) and placement issues (in the voting booth, on the ballot, where on the ballot, etc.)

These, and other important issues about voter instructions, deserve scientific study.

6. Language and Images

As American continues to become a more diverse nation, many election jurisdictions are finding they have to provide ballots and voter instructions in multiple languages. Many jurisdictions in California, Texas, Florida, and New York run elections in both English and Spanish; there are now precincts in Los Angeles County where ballots and voter instructions are provided in a number of Asian and European languages.

However, ballot and voter instructions may not directly and clearly translate from English to many of these other languages; this leaves open the question as to whether non-English ballot design and voter instructions are understood and utilized by non-English voters in the same way that English voters understand and use the ballot design and instructions. The impact of language used for ballot design and voter instructions clearly needs scientific examination.

Some nations have devised a different solution to the problem of a diversity of languages, and to alleviate possible problems associated with high rates of voter illiteracy, by not using candidate or party names in their ballot designs. Instead,

¹⁴ Of course, other factors might have produced, or contributed to, these dramatic declines in over- and undervotes between these two elections. Many voters might have become more aware of “pregnant”, dimpled, or hanging chads in the wake of the 2000 election, and thereby were more careful in their use of the Votomatic punchcards in 2001. Other groups, especially groups representing nonwhite voters, worked to educate their constituencies about the Votomatic system in the wake of the 2000 elections. Last, Los Angeles City election administrators were more proactive in their “cleaning” of punchcards in 2001, and were observed by the author removing hanging chads from punchcard ballots before they were run through tabulation machines on election evening.

candidates and especially political parties devise specific pictorial images that become associated with the party through pre-election education campaigns. Voters then cast a vote for a party by selecting an icon on an electronic voting machine (in the case of Brazil, for example), rather than touching a square or circle next to the name of a candidate or party.

It is not clear whether pictorial icons would easily translate to the American system of elections, especially given the long length of ballots in the United States. A proliferation of pictorial icons would have to be created, not just for each party but also for votes on ballot measures, and for candidates in non-partisan races. But pictorial icon ballot design should be tested, relative to more typical ballot designs, to determine whether they lead to a more accurate and higher quality voting experience for voters (or certain classes of voters).

7. A Scientific Agenda for Studying Ballot Design

Each of the dimensions of ballot design discussed above can, and should, be examined within the context of careful scientific analysis. Research agendas based on both controlled experimental studies and on quasi-experimental designs should be developed and implemented.

A prototype of a controlled experimental study on the “butterfly” ballot could take the following form. Two small groups of randomly selected registered voters from an election jurisdiction could be contacted about participation in a research project on election technologies. Subjects could be offered small inducements for their participation. One group would be the control group, and they would be asked to participate in a hypothetical election, using the same ballot style used in that election jurisdiction in the most recent election. The experimental group would be asked to participate in a hypothetical election, “voting” for the same set of candidates as the control group, but using a “butterfly” ballot. Both groups would be interviewed, either using a quantitative or qualitative approach, before and after their “voting” about their experience and opinions about the ballot. The ballots cast by both groups could then be examined for errors, and for deviations in the vote cast from the voter’s stated “vote” in the follow-up interview. Such a study design could produce a powerful, and clear, analysis of the precise effects of the butterfly ballot on voter behavior in the ballot booth.

However, many important quasi-experimental studies can also be conducted using data from historical and contemporary elections. First, there are collections in various archives of election ballots; an interesting example is the Munro collection of ballots at the Huntington Library in San Marino, California. One study has been conducted using this archive, examining the impact of the Australian ballot reform in the United States.¹⁵ Archival research should be

¹⁵ Lisa A. Reynolds, 1995, “Reassessing the Impact of Progressive Era Ballot Reform”, Ph.D. Dissertation, University of California, San Diego.

undertaken to study the details of historical ballot designs, and to link the designs with voting behavior and election outcomes.

Second, there is a great deal of heterogeneity in the United States in ballot designs, even for election jurisdictions using the same type of voting system. For a prime example, one important difference in the ballot designs currently in use for optically scanned ballots is whether the voter indicates her vote by filling in a circle or by connecting two broken lines to make a solid unbroken line. Some jurisdictions use the first type of ballot, others the second type. Data about overvotes, undervotes, and ballot spoilage rates could be collected for contemporary elections for each type of ballot design, and could be linked with other demographic and political data from each jurisdiction. A statistical analysis of such databases could demonstrate what impact these two different optical scan ballot designs have on election outcomes and voting behavior.

Third, by working more closely with election administrators, researchers could conduct detailed “before and after” studies of changes of ballot designs in particular election jurisdictions. For example, if researchers learn about an important change in voter instructions --- or about a voter education campaign about how to use a particular voting system --- they can work with the particular election jurisdiction to track and collect specific data so as to measure the impact of the change in voter instructions by studying changes in aggregated election statistics. Also, researchers could work with election administrators to conduct surveys of voters leaving polling places after some change in ballot design was implemented to elicit opinions and perceptions of the new design. Last, polling place workers themselves can be interviewed within the context of this type of quasi-experimental study to obtain their assessment of the ballot design change, and to gather data on the problems that voters had with the new ballot design.

Within each type of study, though, there are many evaluative dimensions to consider. The first and most obvious are indicators of voter error. Unfortunately, while the concept is obvious, a precise measure for voter error may be difficult to operationalize. Most studies have taken the “residual” or roll-off rate (the difference between the number of votes cast and the number counted for a specific race) as a measure of error, while others look at more specific measures like over- or undervotes. One problem with these approaches is that they do not differentiate between intentional and unintentional actions by voters. This is an area where social scientists have done little study --- but where more research is obviously needed and a place where controlled experimental studies might be very helpful.¹⁶ Secondly, there is what I have termed the “quality of the voting experience.” This is a broad concept, and should incorporate measures of the

¹⁶ The only study that attempts to differentiate intentional from unintentional voter errors that I am aware of is Stephen Knack and Martha Kropf’s 2001 unpublished analysis, “Roll Off at the Top of the Ballot: Intentional Undervoting in American Presidential Elections.” Knack and Kropf examine presidential exit poll data and compare it to aggregate voting statistics to estimate the rate of unintentional undervoting in presidential elections.

voter's satisfaction, efficacy, and perception that his or her cast vote would be counted. Third, we should focus attention on studying the impact of ballot design on pollworkers. People who work in polling places on election day are typically most clearly exposed to the problems with existing ballot designs; quantitative and qualitative studies of their perceptions of the problems voters face are critical for quasi-experimental studies of ballot design.

Last, studies of ballot design should include as important independent variables attributes of voters. Important and obvious attributes to study are disabilities, especially vision impairment and physical disabilities. But not as obvious, but perhaps as important, are the many factors that lead some voters to be considered poorly informed (factors like educational attainment, income, weak social connectedness and capital). Given that many ballot design issues revolve around the cognitive capabilities of voters, variance in these capabilities must be factored into studies of ballot design.