STATEMENT OF DOUGLAS W. JONES
Regarding the Optical Mark-Sense Vote Tabulators in Maricopa County

1. My name is Douglas W. Jones. I am Associate Professor at the University of Iowa Department of Computer Science, where I have taught since 1980.
2. My expertise in voting technology includes the following:
3. I served on the Iowa Board of Examiners for Voting Machines and Electronic Voting Systems from 1994 to 2004, and chaired the board for 3 terms. This board examines all voting systems offered for sale in the state of Iowa to determine if they meet the requirements of Iowa law.
4. I was invited to testify before the United States Commission on Civil Rights on evaluating voting technology for their January 11, 2001 hearings in Tallahassee Florida. I was invited to testify before the House Science Committee on problems with voting systems and the applicable standards for their May 22, 2001 hearings. I was invited to testify before the Federal Election Commission on voting system standards for their April 17, 2002 hearings.
5. I wrote Chapter 1 of Secure Electronic Voting, edited by Dimitris Gritzalis and published by Kluwer Academic Publishers in 2002.
6. In the summer of 2004, I consulted with Miami-Dade County to assess problems with their touch-screen electronic voting system and to assess their pre-election testing of their touch screen and optical scan voting systems.
7. My paper, Auditing Elections, was published in the Communications of the Association for Computing Machinery in October 2004.

8. I am one of the ten principle investigators in A Center for Correct, Usable, Reliable, Auditable, and Transparent Elections (ACCURATE), a multi-institutional center awarded a 5-year research grant by the National Science Foundation starting in October 2005.
9. In November 2005, I was invited to Kazakhstan by the Office for Democratic Institutions and Human Rights of the Organization for Security and Cooperation in Europe to help assess the Kazakh electronic voting system.

Maricopa County

10. In late November, 2005, following initial contacts by Anton Orlich, Senator Jack Harper of the Arizona Senate contacted me with regard to the possibility that I could investigate the discrepancy between the first count and the recount of the Republican ballots in the September 7, 2004 District 20 primary election; I responded to Senator Harper with a letter dated November 30.
11. On December 20, 2005, I examined the 8 ES&S Optech 4C vote tabulators in the tabulating center of the Maricopa County Elections office, in the presence of the County Recorder, Helen Purcell, the Election Director, Karen Osborne, John Stewart of the Elections department, and several other county employees, as well as Jose Munguia and Steve Wang from Election Systems and Software, and Senator Harper.
12. The 8 machines are named Count 1 through Count 8, inclusive.
13. Of the 8 machines, two were non-functional, Count 6 and Count 8. These machines both suffered from mechanical problems in the ballot feed mechanism that precluded testing. Machine number 5 also suffered from the same feed problem, but to a lesser extent that did not preclude testing.

14. In advance of the test, Maricopa County provided me with laser-printed reprints of all of the ballot styles used on September 7, 2004 in the District 20 Republican primary, both partisan and non-partisan (marked OTH), representing 56 different precincts; this is 112 distinct ballot styles.
15. I marked these ballots in order to perform two distinct tests, first, a test of tabulation accuracy and second, a test of the sensitivity of the mark-sensing mechanism.

Tabulation Tests

16. For the tabulation test, I marked the 8 or 9 ballots per precinct (see Exhibit 1) in a total of 60 precincts, using all 56 non-partisan Republican ballot styles and 4 partisan Republican ballot styles (see Exhibit 2).
17. This marking assigned a different number of votes to each candidate in the race for District 20 state representative (including 2 write-in candidates named Capistrano and Bullmoose) on each of 58 ballot styles, including all 56 non-partisan styles and 2 of the 4 partisan styles for a total of 510 ballots; the other 2 partisan style ballots were not voted in this race, but were used for sensitivity testing.
18. The marking pattern I used allows the test to detect any exchange of votes between the candidates or the failure to count votes for any candidate on any of the 58 ballot styles.
19. For this test, all marks were very dark marks made with a number 2 soft lead pencil, completely connecting the two halves of the arrow used as a voting target for the full width of the arrow.

20. An additional dark mark was made for John McCain on each and every ballot that was part of this test, in order to provide an independent ballot count.
21. This test is somewhat more intensive than the logic and accuracy testing normally done before every election and it is focused on just one race, but in spirit, this is simply a conventional logic and accuracy test.
22. The entire tabulation test was performed twice, once on the tabulating machines named Count 4 and once on the machine named Count 7.
23. The Summary Report provided by the voting system after each test showed exactly the same vote totals as predicted by my test plan (see Exhibits 2 and 3). From this, I conclude (i) that my test ballots conformed with my test plan, (ii)that the tabulating software functioned correctly, and (iii) that the election configuration files were correct.

Sensitivity Tests

24. The sensitivity test was conducted using the ballots for the four partisan Republican ballot styles that were included in the tabulation test.
25. 16 of these ballots, representing two styles (Precincts 14 and 26) were marked as part of the tabulation test in order to show that partisan ballots were counted as well as non-partisan ballots.
26. 20 of these ballots, representing the other two styles (Precincts 7 and 9), were marked to test the scanner's ability to pick up marks made with a variety of different pens and pencils.
27. The races for US Representative, State Senate, Corporation Commissioner, and five county offices were used for sensitivity testing.

28. Votes were cast for 16 candidates in these races, in order to test 6 different ballpoint marking devices, 2 different kinds of pencil, erasure of pencil marks, 4 types of felt-tip marker, and 2 types of glitter pen (See Exhibit 4); note that the number-2 soft lead pencil was tested twice, once in each of two ballot columns.
29. Ten votes were cast for each of these 16 candidates.
30. All votes for any one of these candidates were made with the same pen or pencil, but using a sequence of successively darker marks for each vote; the darkest of these marks connected the full width of the two halves of the arrow, while the lightest mark was a dot (see Exhibit 5).
31. The test of pencil erasures was made by making the darkest pencil mark on each ballot and then making successively more strokes of a new pencil eraser on each successive ballot in an attempt to erase the mark.
32. As a result, after tabulation, each of these candidates received a vote total of between zero and ten in each race used in the sensitivity tests, depending on the ability of the ballot scanner to detect the corresponding pen or pencil; a count of ten indicates that the scanner detected and counted all of the marks made, even a spot or a determined attempt to erase, while a score of one indicates that it counted only the darkest mark.
33. The 36 sensitivity test ballots were run through each of the six operational ballot tabulators 4 times, face up – head first, face up – foot first, face down – head first, and face down – foot first, for a total of 24 experimental runs (see Exhibit 6).

34. During the sensitivity tests, the total amount of ballot handling was considerably greater than ballots would normally be subjected to during an election and recount.

35. As a result of this intensive handling, the ballots showed significant wear and a significant amount of grayish dust and black crumbs accumulated on the ballots, so much so that, after 18 runs of the ballots through the machines, I opted to wipe the backs of the ballots with a piece of tissue paper.

36. I was told by Maricopa county elections office employees that the regular ballots, which are offset printed, shed significantly less dust than the xerographic (laser printed) copies used for these tests.

37. Despite the accumulation of dirt, no new votes were created in any of the races on the ballots during any of the tests I performed.

Interpretation of Sensitivity Test Results With Regard to Ballot Marking Devices

38. Average sensitivity scores were computed for each marking implement, averaged over the 24 experimental runs in the sensitivity tests (see Exhibit 7).

39. The sensitivity test showed that the Optech 4-C is unable to sense all but the very darkest of marks made using red ink, whether from a ballpoint pen, from a felt-tipped marker, or from a glitter pen.

40. Jose Munguia and Steve Wang of Election Systems and Software informed me, on December 20, that these scanners use and have always used visible red light, not infrared light.

41. The general inability to sense red ink that I found in my testing is in complete agreement with what I would expect from a scanner that used red light for mark sensing.

42. While the auxiliary instructions provided with the ballot apparently warn against using red ink, the actual instructions on the ballot offer no such warning; it would be better to do so, for example, as was done on the Orange County, Florida ballot in November 2000 (see Exhibit 8).
43. The sensitivity test showed that these machines are extraordinarily sensitive to black pencil, to ink from a Jelly Roll brand blue glitter pen, and to black ink from a Sanford Sharpie Extra Fine Point pen.
44. With these pens and pencils, the Optech 4-C scanners routinely counted marks as small as a single dot.
45. While I was performing the tests, several Maricopa county employees commented on the extreme sensitivity of the Optech 4-C scanner; sensitivity to a small pencil dot confirms this.
46. Exhibit 9 presented to Judge Eddward Ballinger, Jr. of the Superior Court in Maricopa County on September 23, 2004 apparently contained several ballots were counted as overvotes because of smudges or other nearly imperceptible marks (See pages 41-43 and 140-141 of the transcript).
47. In general, ballot scanners that detect single dots made with commonly used marking implements are not desirable; they make it too likely that marks will be counted as votes that few voters proofreading their own ballots would notice.
48. On September 5, 2005, Richard Ruelas, writing in the Arizona Republic, quoted Karen Osborne as having said that Sharpie and Glitter pens (among others) were sources of problems; my sensitivity tests contradict this.

49. In my opinion, the problem with Sharpie and some similar markers is that the ink penetrates the ballot paper and can make a mark that is visible from the opposite side of the ballot; this can be a serious source of trouble on a two-sided ballot, but the ballots in question were one-sided.
50. Of all of the non-red pens and pencils tested, the black Bic ballpoint pen was the least effective ballot marking device; on 2 of the 24 tests, a single stroke made with this pen fully spanning the broken arrow was not sensed as a mark.
51. On September 5, 2005, Richard Ruelas, writing in the Arizona Republic, quoted Karen Osborne as having said that she recommended the Bic "Round Stic" pen.
52. The Bic pen I tested was not a "Round Stic," but I have compared marks made by a "Round Stic" with marks made by the Bic pen used in my tests, using a 10x jeweler's magnifier, and I found no visible difference in the lines produced by these two pens; I suspect that the ink and the ballpoint itself were identical.
53. While the average sensitivity of the Optech scanner to the Bic pen was reasonable, my results suggest that ballpoint pens using free-flowing ink and fine-point markers are, in general, better for ballot marking.
54. It is significant that the pen apparently recommended by the county is among the worst of the non-red marking implements I tested; recommendations for ballot marking devices should not be made before broad-based testing of a variety of common marking devices.
55. My tests also demonstrate that the Optech 4C scanners were unable to disregard any of the pencil marks I attempted to erase.

56. In contrast, it should be noted that the Optech I, Optech II and Optech III precinct-count ballot tabulators, as well as the ES&S 650 central count ballot tabulator have all proven to be modestly good at distinguishing careful erasure from deliberate marks.

Interpretation of Sensitivity Test Results With Regard to Machine Calibration

57. Each side of the Optech ballot, as used in Maricopa County, contains four vertical tracks of marking positions, one track on each edge, and two tracks nearer the center, so that the ballot is divided into three equal columns bounded by these tracks.

58. The marking positions tested in the sensitivity tests were in the tracks immediately to the right of each of the leftmost two columns on the ballot, which is to say, they were the two tracks nearest the center of the ballot.

59. As a result, whether the ballots are fed head first or feet first, the sensitivity tests always tested the same two tracks, the two middle tracks on the marked side of the ballot.

60. Average sensitivity scores were computed for each voting machine tested, reading ballots face up and face down; each of these averages combines the data for reading head first and for reading feet first, giving a total of 12 averages (see Exhibit 7).

61. These averages are only useful as a measure of relative sensitivity when comparing different scanners or the two sides of one scanner; the value of this average for any one side of one scanner has no particular meaning.

62. The sensitivity of the Optech 4C scanners measured using these averages ranged 7.5 or higher on 5 of the machines, to a low of 6.0 on one of them.

63. The machine known as Count 4 had top-side sensors with a sensitivity of 7.25 and bottom-side sensors with a sensitivity of 6.00, the greatest difference of any machine I tested.

64. The pattern of low scores on Count 4 suggests that a single photosensor on the bottom sensor assembly of that machine was unusually insensitive to all marks made with anything but a pencil or sharpie marker.

65. The machines known as Count 2 and Count 5 exhibited similar patterns suggestive of problems with single top-side sensor, but to a smaller extent than on Count 4.

66. The problems on Count 4 and Count 2 were severe enough that these machines could not reliably sense single stroke lines made by 6 of the pens tested, where all 6 pens made lines that were reliably sensed as votes by the other machines; the problem pens included the Cross and Bic ballpoints, the only two conventional (viscous ink) non-red ballpoint pens included in my sample of marking devices.

67. Had Count 4 and Count 2 been taken out of service, only one pen, a red felt-tipped marker, would have been problematic.

68. Note that the ballot marking instructions printed on the ballot ask the voter to "complete the arrow ... with a single line" (see Exhibit 8) and that this form of mark is precisely the form that caused problems on Count 4 and Count 2 when made with conventional ballpoint pens.

69. Note that, during the tests on December 20, ES&S technicians were on hand; they said that they had checked the machines before the tests and they

performed normal maintenance tasks on the machines during testing, just as they would have done during a real election.

70. It is reasonable to conclude that the testing and calibration done by ES&S were inadequate, and that the oversight of testing and calibration by the county were inadequate.

71. Had the ballot marking instructions recommended a darker mark than "a single line," the impact of these inadequacies might have been reduced; I would recommend, at the very minimum, using the phrase "a dark line" in order to encourage voters to make multiple strokes if they have any doubt about the darkness of their markings.

72. Note, however, that we have very little data about how voters actually respond to ballot marking instructions; the best way to get this information would be to gather statistics from a random sample of ballots actually cast during a real election.

73. I would recommend that the state require that pre-election testing include not only testing of the accuracy with which the tabulating machine can tabulate prescribed marks made with the prescribed pen, but also tests to determine whether the voting machine's mark discrimination thresholds are set reasonably based on the way real voters respond to the ballot marking instructions.

Repeatability of Test Results

74. It is reasonable to ask whether the results of my tests are repeatable.

75. Because the calibration test ballots were included in the set of 510 ballots used for the tabulation tests, it is possible to compare the totals from the two runs of

the tabulation test with the totals obtained during the tabulation tests (see Exhibit 9).

76. These two repeated runs showed only a few differences, and these differences were differences of only a single mark.

77. Such off-by-one differences are most likely explained by the fact that one of the test markings was very close to the sensing threshold for one of the scanners under test.

78. Given that the test marks included marks that were deliberately intended to be marginal, the occurrence of a few off-by-one differences is to be expected.

79. Therefore, these results suggest that the calibration test measurements reported here reflect real differences in the calibration of different machines.

Voting System Standards and the State Voting System Certification Process

80. Karen Osborne testified, before Judge Eddward Ballinger, Jr. of the Superior Court in Maricopa County on September 23, 2004 that, while there are slight differences between the machines, the calibration of these machines conforms to an industry standard set by national testing laboratories and approved by the Secretary of State's office (See pages 114 and 115).

81. I understand this to be a reference to the Federal Voluntary Voting System Standards, to the Independent Testing Authorities that certify voting system conformance to these standards, and to the state requirement that systems used in Arizona conform to these Federal standards.

82. A longstanding problem with the Federal Voluntary Voting System Standards, both the 1990 and the 2002 versions, is that they do not dictate sensing thresholds; they merely dictate that the machine be adjusted in such a way that

it can reliably count marks made in the recommended way with the recommended marking device (see Section 3.2.4.2.3 of the 2002 Federal Election Commission Voting System Standards) .

83. The United States Election Assistance Commission report on Improving the Usability and Accessibility of Voting Systems and Products, dated April 2004, does not address optical mark-sense systems to any significant extent (see http://www.eac.gov/finalhfvotereport-4-29-041wcover.asp).

84. The United States Election Assistance Commission material on best practices provides some guidance for pre-election tests of mark-sense scanners, in the form of a link to the Pasco County Florida Election Security Procedures; these do not provide for testing of ballot tabulator thresholds (see http://www.eac.gov//bp/docs/PascoSecurity.pdf).

85. The new 2005 Voting System Standards issued by the Election Assistance Commission appears to make no substantial change to the 2002 standard with regard to how mark-sensing thresholds are set (see http://guidelines.kennesaw.edu/vvsg/).

86. I should note that I performed tests on several other mark-sense voting machines that are similar to but less extensive than those reported here.

87. When I tested the ES&S Model 650 central count ballot tabulator for Miami-Dade County, Florida, I found that that tabulator had thresholds set to a very reasonable level, counting marks made with reasonable pens and pencils, excluding erasures, and excluding dots and short dashes (see Observations and Recommendations on Pre-election testing in Miami-Dade County, http://www.cs.uiowa.edu/~jones/voting/miamitest.pdf).

88. When I tested the Global AccuVote ES-2000 mark-sense scanner (now supported by Diebold) for the state of Iowa, I found that the default calibration would only reliably count marks made using a Sharpie or equivalent marker, while only the very darkest of pencil marks were recognized.
89. At the time of my tests, both of these machines were certified to conform to the applicable Voluntary Voting System Standards by the appropriate independent testing authorities.
90. I believe that the results reported here for the Optech 4C, combined with my previous observations of the ES&S Model 650 and the AccuVote ES-2000 clearly demonstrate the lack of an effective Federal standard for setting the sensitivity thresholds for mark-sense vote tabulating systems.
91. Note that when voters check their own ballots, they refer only to the ballot itself and to a few sentences of instructions, not to complex technical documentation or codes of law describing what is an acceptable ballot marking.
92. Therefore, I strongly suggest that the technical specification for what constitutes an acceptable ballot marking, both in law and in voting system testing standards, must reflect the interpretations of typical voters informed only by the ballot itself and the few sentences of instruction provided on or with the ballot.
93. My tests identify a clear problem with the scanner calibration; it is too sensitive to pencil lead, yet too insensitive to ballpoint pen marks made in conformance to the current instructions.

94. The above-stated observations identify a clear problem with the knowledge expressed by Maricopa County election officials about their vote tabulating machines and about the standards to which these machines are certified.

95. I should note that these problems are by no means unique to Maricopa County or the State of Arizona, but rather, that they are typical of election administration across the United States.

96. I strongly recommend that prior to certification of an optical mark-sense voting system for use in the state, the state should check to see that the mark-sensing thresholds conform to the way reasonable people interpret marks on the same ballot; this check will remain essential until this issue is clearly addressed in the Federal Voting System Standards.

97. Furthermore, since the way voters interpret their marks depends, in part, on the instructions printed on the ballot, any change to these instructions should require recertification by the state before the changed wording is permitted to be used in an election.

A Hypothesis to Explain the District 20 Recount Results: Ballot Wear

98. One hypothesis put forward to explain the anomalous results of the District 20 recount is that marginal marks were created at random by ballot handling between the first and second count (this is implicit in Karen Osborne's testimony before Judge Eddward Ballinger, Jr. of the Superior Court in Maricopa County on September 23, 2004 on pages 102-104).

99. The fact that I did not observe any new votes being created on my test ballots, despite the accumulation of dust and crumbs from the intensive handling, suggests that this hypothesis is unlikely.

100. The random addition of marks resulting from ballot handling would be expected to add marks with equal likelihood to ballots where voters had voted for two candidates in the State Representative race as it does on ballots where voters had voted for only one candidate, and it would be expected to be equally likely to add marks to partisan Republican ballots as it does to non-partisan Republican ballots.

101. If a vote were added to a ballot where a voter had already voted for two candidates in the District 20 race, this would be detected and counted as an overvote by the tabulating machinery; where a voter votes for fewer than the allowed number of candidates, this is counted as an undervote.

102. The statistics from the recount, as presented in Karen Osborne's testimony before Judge Eddward Ballinger, Jr. of the Superior Court in Maricopa County on September 23, 2004 (see pages 91-92) do not conform to the patterns that ballot wear hypothesis predicts; instead, it appears that undervoted partisan Republican ballots were unusually likely to pick up votes, while few overvotes were created on any ballots and few votes were added to non-partisan Republican ballots.

103. The results of my tests taken together Karen Osborne's testimony cited above suggest that this hypothesis is very unlikely to explain more than a small part of the difference between the first and second counts.

A Hypothesis to Explain the District 20 Recount Results: Marginal Marks

104. Another hypothesis put forward to explain the anomalous results of the District 20 recount is that "it's a bit of a crapshoot whether your ballot got counted. It depends on what machine read your ballot, and whether you used

a Bic pen." (Quoted from the Arizona Republic, September 5, 2005, Richard Ruelas.)

105. My experiments allow this hypothesis to be refined as follows:

106. First, we assume that one or more of the ballot scanners used on the first count for partisan Republican ballots exhibited the kind of insensitivity that I observed on the machines known as Count 4 and Count 2.

107. Second, we assume that the machine or machines used to count the nonpartisan Republican ballots did not suffer from this problem.

108. Third, we assume that the machine or machines used in the recount did not suffer from this problem.

109. And finally, we assume that a significant number of voters voted with marginal marks, that is, marks with a darkness between the sensing thresholds of the "normal" machines and the insensitive machine; note that in my tests, a voter making the prescribed mark, "a single line," with the recommended Bic pen, would have made such a marginal mark.

110. If we assume that one out of something like four scanners used on the first count of the partisan Republican ballots was insensitive, a vote gain in the recount of approximately 4 percent would require that approximately 16 percent of the ballots be marked using marginal marks.

111. My testing does not allow any conclusion to be drawn about which ballot scanner might have been insensitive during the first count because the scanners have been subject to normal maintenance, including recalibration, between the time of the first count and the time of my testing.

112. An examination of a representative sample of the actual partisan Republican early voting ballots from the election would allow this hypothesis to be confirmed or ruled out; it is important that the sample be selected at random.

113. If the examination shows that around 16 percent of the ballots were marked with a single stroke using a ballpoint pen, as opposed to use of pencil or votes cast with a deeply scribbled mark, this hypothesis would become very likely.

114. Examination of a few hundred ballots out of the thousands cast would expose the frequency of such markings with very high confidence.

115. In general, whenever a machine recount produces a significant change from the first count, examination of such a sample of the ballots is an appropriate measure to help determine if the discrepancy was caused by the kind of marginal mark problems described here.

A Hypothesis to Explain the District 20 Recount Results: Fraud

116. A final hypothesis can be put forward to explain the gain in votes during the recount. This hypothesis is that ballots were altered between the time of the first count and the time of the recount.

117. Karen Osborne testified, before Judge Eddward Ballinger, Jr. of the Superior Court in Maricopa County on September 23, 2004 that the ballots for the District 20 primary were subject to significant handling between the first count and the recount, and that there were times when ballots were handled in anticipation of the recount without first informing the candidates or others who might wish to observe this handling (see pages 57-62 and particularly page 60 lines 6-14).

118. This may have created an opportunity that could have been used to surreptitiously add marks to ballots; I want to emphasize that I do not allege that any such ballot alteration occurred, but only that the opportunity may have existed.

119. The pattern of ballot alteration necessary to create the pattern of new votes observed in the recount would require that someone search out undervoted partisan Republican early voting ballots and add single new votes to a significant fraction of these ballots.

120. If someone is in a hurry to surreptitiously alter large numbers of ballots, I would expect them to mark those ballots all with the same pen or pencil and to make no serious effort to match the style of marks made by the original voter.

121. Therefore, there is a reasonable chance that an examination of a sample of the partisan Republican early-voting ballots would reveal the presence of surreptitiously made marks, so long as this sample is indeed a representative random sample.

122. In general, such an examination would seek out ballots where one of the two votes cast in the state representative race was different in style or marking implement from the style and marking implement used for all other votes on the same ballot.

123. Since approximately four percent of new votes were found in the recount of the partisan Republican early-voting ballots, an examination of several hundred ballots would be required to rule out this hypothesis.

124. In general, whenever a machine recount produces a significant change from the first count, examination of such a sample of the ballots is an appropriate

measure to help determine if there is any evidence for the kind of ballot tampering described here.

Concluding Remarks

125.I have reached the following conclusions in the above:

126.My testing did not uncover any errors in election configuration or vote tabulation software (point 23), there was no evidence that wear and tear on the ballots was likely to have changed how they read (point 103), and there was no evidence that the results of scanning are likely to change if ballots are re-scanned with no change in ballot orientation (point 79).

127.The ballot marking instructions currently used in Maricopa County should be revised (points 42, 71) and ballot marking instructions should not be left to the discretion of individual counties (point 97).

128.We do not know enough about how real voters mark their ballots; empirical data obtained from examination of real ballots can aid in setting standards for mark-sense voting machine calibration (point 96), mark interpretation (point 92), and formulating appropriate ballot marking instructions (point 72).

129.Currently, no effective standard regulates the setting of the sensing thresholds of mark-sense ballot tabulators; until the Federal Voluntary Voting System Standards are modified to include such regulation, the state must do so (point 96).

130.Currently, voting system vendors set the sensing thresholds, but the setting of these thresholds is not adequately monitored by the county; pre-election testing must be augmented to check these thresholds (point 73).

131. Voting officials have made inaccurate comments about the best marking implements to use for voting; to avoid such comments, a variety of common marking implements should be routinely tested (point 54).

132. Without empirical examination of a random sample of voted ballots, there is no way to decide between the hypothesis that ballots have been altered and the hypothesis that ballots were miscounted by poorly calibrated machines after a recount identifies such discrepancies (points and 115 and 124).

133. Note in conclusion that, while my work points to some inadequacies in work by the voting system vendor and by the county, these inadequacies are typical of current practices nationwide; similarly, while I have pointed out inadequacies in the regulation of voting technology by the state of Arizona, these problems are also typical of most of the country.

134. The primary burden of this work therefore falls on the states, which must allocate adequate resources and provide the necessary legislative framework to properly regulate mark-sense voting, and on the Federal government, which needs to revise the current Voluntary Voting System Standards to account for the realities of mark-sense voting.

I assert that the above statements are true, to the best of my knowledge.

This 12th day of January, 2006

Douglas W. Jones

Exhibit 1: Ballot marking patterns

The following ballot marking patterns involve 8 (patterns A through J) or 9 (patterns K through T) ballots per precinct. Each column of each pattern corresponds to one ballot. The 5 different rows of patterns correspond to the 5 different ballot rotations that apply to the District 20 State Representative race. The candidate names Capistrano and Bullmoose are arbitrary names used for write-in candidates.

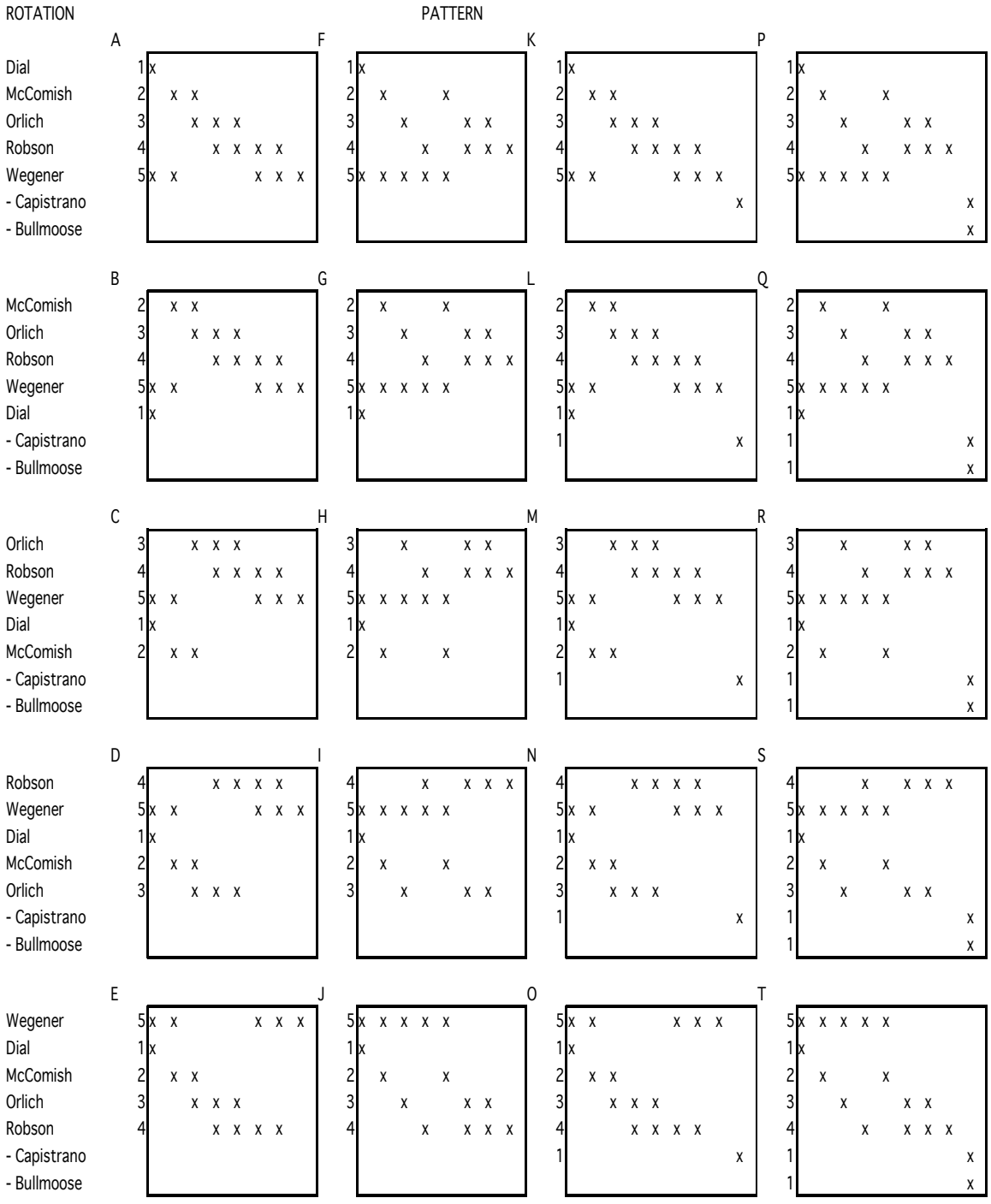

Exhibit 2: Marking Plan for all 62 precincts in the tabulation test

PRECINCT	PATTERN	BALLOTS	DIAL	MCCOMISH	ORLICH	ROBSON	WEGENER	CAPISTRANO	BULLMOOSE
7	sensitivity	10							
7 OTH	A	8	1	2	3	4	5		
9	sensitivity	10							
9 OTH	C	8	1	2	3	4	5		
14	J	8	1	2	3	4	5		
14 OTH	E	8	1	2	3	4	5		
26	E	8	1	2	3	4	5		
26 OTH	J	8	1	2	3	4	5		
68 OTH	D	8	1	2	3	4	5		
88 OTH	H	8	1	2	3	4	5		
93 OTH	O	9	1	2	3	4	5	1	
103 OTH	B	8	1	2	3	4	5		
109 OTH	G	8	1	2	3	4	5		
112 OTH	T	9	1	2	3	4	5	1	1
117 OTH	E	8	1	2	3	4	5		
118 OTH	J	8	1	2	3	4	5		
123 OTH	O	9	1	2	3	4	5	1	
124 OTH	M	9	1	2	3	4	5	1	
125 OTH	R	9	1	2	3	4	5	1	1
126 OTH	C	8	1	2	3	4	5		
127 OTH	T	9	1	2	3	4	5	1	1
132 OTH	E	8	1	2	3	4	5		
133 OTH	D	8	1	2	3	4	5		
134 OTH	I	8	1	2	3	4	5		
135 OTH	F	8	1	2	3	4	5		
140 OTH	K	9	1	2	3	4	5	1	
142 OTH	N	9	1	2	3	4	5	1	
150 OTH	P	9	1	2	3	4	5	1	1
163 OTH	A	8	1	2	3	4	5		
283 OTH	F	8	1	2	3	4	5		
290 OTH	K	9	1	2	3	4	5	1	
413 OTH	L	9	1	2	3	4	5	1	
426 OTH	S	9	1	2	3	4	5	1	1
432 OTH	H	8	1	2	3	4	5		
455 OTH	Q	9	1	2	3	4	5	1	1
467 OTH	P	9	1	2	3	4	5	1	1
617 OTH	M	9	1	2	3	4	5	1	
661 OTH	B	8	1	2	3	4	5		
720 OTH	G	8	1	2	3	4	5		
721 OTH	R	9	1	2	3	4	5	1	1
751 OTH	L	9	1	2	3	4	5	1	
772 OTH	J	8	1	2	3	4	5		
812 OTH	A	8	1	2	3	4	5		
816 OTH	I	8	1	2	3	4	5		
830 OTH	C	8	1	2	3	4	5		
848 OTH	N	9	1	2	3	4	5	1	
862 OTH	S	9	1	2	3	4	5	1	1
914 OTH	D	8	1	2	3	4	5		
915 OTH	Q	9	1	2	3	4	5	1	1
919 OTH	B	8	1	2	3	4	5		
920 OTH	G	8	1	2	3	4	5		
921 OTH	H	8	1	2	3	4	5		
922 OTH	F	8	1	2	3	4	5		
923 OTH	I	8	1	2	3	4	5		
933 OTH	N	9	1	2	3	4	5	1	
972 OTH	L	9	1	2	3	4	5	1	
991 OTH	K	9	1	2	3	4	5	1	
993 OTH	O	9	1	2	3	4	5	1	
1016 OTH	T	9	1	2	3	4	5	1	1
1020 OTH	M	9	1	2	3	4	5	1	
TOTALS		510	58	116	174	232	290	26	11

Exhibit 3: Results of tabulation tests

This data is summarized from 2 summary reports generated by the GEMS election management system software, each 106 pages long.

	Time	10:16	12:20
	Machine	4	7
McCain	Count ballots, mark every one	510	510

Dial	Pattern	58	58
McComish	Pattern	116	116
Orlich	Pattern	174	174
Robson	Pattern	232	232
Wegener	Pattern	290	290
WRITE-IN	Pattern	37	37

Exhibit 4: Sensitivity test races and the marking implements tested

RACE	SENSITIVITY TESTING	PRECINCTS
	TESTING IN BALLOT COLUMN 1	
US SENATE		
McCain	Count ballots, mark every one darkly	all
US REP D5		
O'Connell	Sensitivity, #2 soft lead pencil	7
Hayworth	Sensitivity, #2 pencil erasure	9
STATE SEN D20		
Mead	Sensitivity, Sanford sharpie XF pt	7
Huppenthal	Sensitivity, Bic black ballpoint	9
STATE REP D20		
Dial	Not involved in sensitivity tests	
McComish		
Orlich		
Robson		
Wegener		
CORP COM 07		
Seel	Sensitivity, Cross blue ballpoint	7
Mayes	Sensitivity, Papermate red ballpoint Med pt	9
CORP COM 09		
Gleason	Sensitivity, Pilot Precise blue rolling ball XF pt	7
Hatch-Miller	Sensitivity, Pilot Precise violet rolling ball XF pt	7
Mundell	Sensitivity, Jelly Roll blue glitter	7
CO BOARD D1		
Brock	Sensitivity, Jelly Roll red glitter	7
	TESTING IN BALLOT COLUMN 2	
CO ASSESSOR		
Russell	Sensitivity, #2 soft lead pencil	7
Ross	Sensitivity, Pilot Razor Point marker	9
CO ATTORNEY		
McCauley	Sensitivity, Millennium purple .05mm	7
Baily	Sensitivity, Pilot Easy Touch red ballpoint	9
CO RECORDER		
Purcell	Sensitivity, Sanford Vis-a-Vis red F pt marker	7
CO SCHOOL		
Dowling	Sensitivity, #3H pencil	7

Exhibit 5: Markings used in sensitivity testing

These markings must be made by hand. The exact spot size and line width will depend on the particular marking implement. Markings 1 through 4 below should be made as scribbles, which is to say, multiple rapidly made pen or pencil strokes. The defining characteristic of mark 3 is that it is wider than mark 4 but with an attempt to keep it under half of the arrow width. Marks 5 through 10 should be made as single strokes; in cases where the stroke comes up short, for example, because of skips with a ballpoint pen, non-overlapping strokes or spots may be added at the ends of the initial stroke in order to fill out the length – but in no case should this produce parts of the composite stroke that are darker than typical of a single stroke.

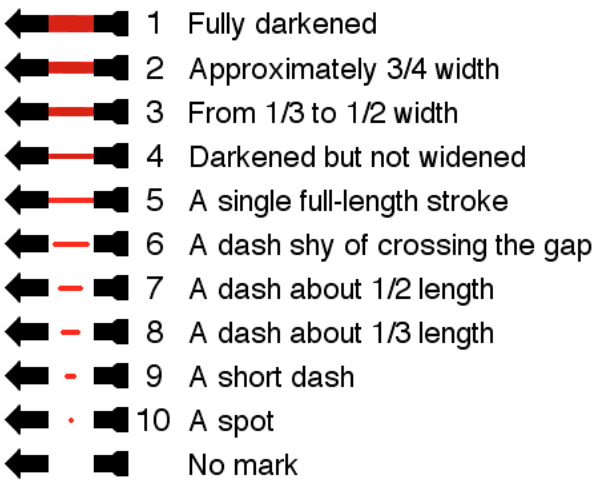

Exhibit 6: Results of sensitivity tests

FUHF = Face-Up Head-First
FUFF = Face-Up Feet-First
FDHF = Face-Down Head-First
FDFF = Face-Down Feet-First

	Time (chronological order)	10:27	10:29	10:30	10:32	10:35	10:45	10:46	10:47	10:49	10:54	10:55	10:57
	Machine	4	4	4	4	5	1	1	1	1	2	2	2
Candidate	Orientation	FUHF	FUFF	FDHF	FDFF	FUHF	FUHF	FUFF	FDHF	FDFF	FUHF	FUFF	FDHF
McCain	Count ballots, mark every one	36	36	36	36	36	36	36	36	36	36	36	36
O'Connell	#2 soft lead pencil	10	10	10	10	10	10	10	10	10	10	10	10
Hayworth	#2 pencil erasure	10	10	10	10	10	10	10	10	10	10	10	10
Mead	Sanford sharpie XF pt	10	9	9	9	9	9	10	9	10	9	10	9
Huppenthal	Bic black ballpoint	8	6	3	8	8	7	8	7	8	3	8	8
Dial	Pattern	2	2	2	2	2	2	2	2	2	2	2	2
McComish	Pattern	4	4	4	4	4	4	4	4	4	4	4	4
Orlich	Pattern	6	6	6	6	6	6	6	6	6	6	6	6
Robson	Pattern	8	8	8	8	8	8	8	8	8	8	8	8
Wegener	Pattern	10	10	10	10	10	10	10	10	10	10	10	10
Steel	Cross blue ballpoint	9	7	4	10	8	5	10	8	10	4	9	9
Mayes	Papermate red ballpoint Med pt	2	1	0	2	2	2	2	1	2	1	2	1
Gleason	Pilot Precise blue rolling ball XF pt	9	8	4	9	9	8	9	9	10	8	8	9
Hatch-Miller	Pilot Precise violet rolling ball XF pt	10	6	0	10	9	6	10	9	10	4	10	10
Mundell	Jelly Roll blue glitter	10	10	6	10	10	10	10	10	10	10	10	10
Brock	Jelly Roll red glitter	0	1	0	0	0	0	0	0	0	0	0	0
Russell	#2 soft lead pencil	10	10	10	10	10	10	10	10	10	10	10	10
Ross	Pilot Razor Point blue marker	9	10	10	1	7	10	10	10	10	10	6	10
McCauley	Millenium purple .05mm	8	9	10	3	8	10	8	10	9	9	7	10
Baily	Pilot Easy Touch red ballpoint	0	0	1	0	0	1	0	1	1	0	0	0
Purcell	Sanford Vis-a-Vis red F pt marker	3	7	3	0	3	7	3	3	1	4	1	3
Dowling	#3H pencil	10	10	10	10	10	10	10	10	10	10	9	10

Exhibit 6, continued

FUHF = Face-Up Head-First
FUFF = Face-Up Feet-First
FDHF = Face-Down Head-First
FDFF = Face-Down Feet-First

	Time (chronological order)	10:58	11:03	11:05	11:06	11:07	11:17	11:18	11:19	11:29	11:30	11:31	11:32
	Machine	2	3	3	3	3	5	5	5	7	7	7	7
Candidate	Orientation	FDFF	FUHF	FUFF	FDHF	FDFF	FUFF	FDHF	FDFF	FUHF	FUFF	FDHF	FDFF
McCain	Count ballots, mark every one	36	36	36	36	36	36	36	36	36	36	36	36
O'Connell	#2 soft lead pencil	10	10	10	10	10	10	10	10	10	10	10	10
Hayworth	#2 pencil erasure	10	10	10	10	10	10	10	10	10	10	10	10
Mead	Sanford sharpie XF pt	9	9	9	9	10	9	9	10	10	9	10	9
Huppenthal	Bic black ballpoint	8	8	6	8	8	5	8	8	9	9	8	8
Dial	Pattern	2	2	2	2	2	2	2	2	2	2	2	2
McComish	Pattern	4	4	4	4	4	4	4	4	4	4	4	4
Orlich	Pattern	6	6	6	6	6	6	6	6	6	6	6	6
Robson	Pattern	8	8	8	8	8	8	8	8	8	8	8	8
Wegener	Pattern	10	10	10	10	10	10	10	10	10	10	10	10
Steel	Cross blue ballpoint	10	8	7	10	10	5	9	10	10	10	10	10
Mayes	Papermate red ballpoint Med pt	2	1	1	2	2	1	2	2	2	2	2	0
Gleason	Pilot Precise blue rolling ball XF pt	9	9	8	10	10	8	8	9	10	10	10	9
Hatch-Miller	Pilot Precise violet rolling ball XF pt	10	9	5	10	10	5	9	9	10	10	10	10
Mundell	Jelly Roll blue glitter	10	10	10	10	10	10	10	10	10	10	10	10
Brock	Jelly Roll red glitter	0	0	0	0	0	1	0	2	1	0	0	0
Russell	#2 soft lead pencil	10	10	10	10	10	10	10	10	10	10	10	10
Ross	Pilot Razor Point blue marker	10	8	10	10	10	10	10	9	10	10	10	10
McCauley	Millenium purple .05mm	9	8	9	10	10	9	10	10	10	10	10	10
Baily	Pilot Easy Touch red ballpoint	1	0	0	0	1	0	0	0	0	1	0	1
Purcell	Sanford Vis-a-Vis red F pt marker	3	3	4	4	5	4	6	4	4	3	2	3
Dowling	#3H pencil	10	10	10	10	10	10	10	10	10	10	10	10

Exhibit 7: Sorted results of sensitivity tests

Averages by marking device are given in the rightmost column.
Averages by machine and orientation are given in the bottom rows.

	Time	10:32	10:30	10:55	10:54	11:17	10:35	11:05	11:03	10:29	10:27	10:46	10:45	
	Machine	4	4	2	2	5	5	3	3	4	4	1	1	
Candidate	Orientation	FDFF	FDHF	FUFF	FUHF	FUFF	FUHF	FUFF	FUHF	FUFF	FUHF	FUFF	FUHF	
Russell	#2 soft lead pencil	10	10	10	10	10	10	10	10	10	10	10	10	
Hayworth	#2 pencil erasure	10	10	10	10	10	10	10	10	10	10	10	10	
O'Connell	#2 soft lead pencil	10	10	10	10	10	10	10	10	10	10	10	10	
Dowling	#3H pencil	10	10	9	10	10	10	10	10	10	10	10	10	
Mundell	Jelly Roll blue glitter	10	6	10	10	10	10	10	10	10	10	10	10	
Mead	Sanford sharpie XF pt	9	9	10	9	9	9	9	9	9	10	10	9	
Ross	Pilot Razor Point blue marker	1	10	6	10	10	7	10	8	10	9	10	10	
McCauley	Millenium purple .05mm	3	10	7	9	9	8	9	8	9	8	8	10	
Gleason	Pilot Precise blue rolling ball XF pt	9	4	8	8	8	9	8	9	8	9	9	8	
Steel	Cross blue ballpoint	10	4	9	4	5	8	7	8	7	9	10	5	
Hatch-Miller	Pilot Precise violet rolling ball XF pt	10	0	10	4	5	9	5	9	6	10	10	6	
Huppenthal	Bic black ballpoint	8	3	8	3	5	8	6	8	6	8	8	7	
Purcell	Sanford Vis-a-Vis red F pt marker	0	3	1	4	4	3	4	3	7	3	3	7	
Mayes	Papermate red ballpoint Med pt	2	0	2	1	1	2	1	1	1	2	2	2	
Baily	Pilot Easy Touch red ballpoint	0	1	0	0	0	0	0	0	0	0	0	1	
Brock	Jelly Roll red glitter	0	0	0	0	1	0	0	0	1	0	0	0	
	Average (sorted by sensitivity)		6.00		6.63		6.88		6.94		7.25		7.34	

	Time	10:49	10:47	10:58	10:57	11:32	11:31	11:19	11:18	11:07	11:06	11:30	11:29	
	Machine	1	1	2	2	7	7	5	5	3	3	7	7	
Candidate	Orientation	FDFF	FDHF	FDFF	FDHF	FDFF	FDHF	FDFF	FDHF	FDFF	FDHF	FUFF	FUHF	Average
Russell	#2 soft lead pencil	10	10	10	10	10	10	10	10	10	10	10	10	10.00
Hayworth	#2 pencil erasure	10	10	10	10	10	10	10	10	10	10	10	10	10.00
O'Connell	#2 soft lead pencil	10	10	10	10	10	10	10	10	10	10	10	10	10.00
Dowling	#3H pencil	10	10	10	10	10	10	10	10	10	10	10	10	9.96
Mundell	Jelly Roll blue glitter	10	10	10	10	10	10	10	10	10	10	10	10	9.83
Mead	Sanford sharpie XF pt	10	9	9	9	9	10	10	9	10	9	9	10	9.33
Ross	Pilot Razor Point blue marker	10	10	10	10	10	10	9	10	10	10	10	10	9.17
McCauley	Millenium purple .05mm	9	10	9	10	10	10	10	10	10	10	10	10	9.00
Gleason	Pilot Precise blue rolling ball XF pt	10	9	9	9	9	10	9	8	10	10	10	10	8.75
Steel	Cross blue ballpoint	10	8	10	9	10	10	10	9	10	10	10	10	8.42
Hatch-Miller	Pilot Precise violet rolling ball XF pt	10	9	10	10	10	10	9	9	10	10	10	10	8.38
Huppenthal	Bic black ballpoint	8	7	8	8	8	8	8	8	8	8	9	9	7.29
Purcell	Sanford Vis-a-Vis red F pt marker	1	3	3	3	3	2	4	6	5	4	3	4	3.46
Mayes	Papermate red ballpoint Med pt	2	1	2	1	0	2	2	2	2	2	2	2	1.54
Baily	Pilot Easy Touch red ballpoint	1	1	1	0	1	0	0	0	1	0	1	0	0.33
Brock	Jelly Roll red glitter	0	0	0	0	0	0	2	0	0	0	0	1	0.21
	Average (sorted by sensitivity)		7.44		7.50		7.56		7.63		7.78		7.81	

Exhibit 8: Contrasting Instructions from Orange and Maricopa Counties

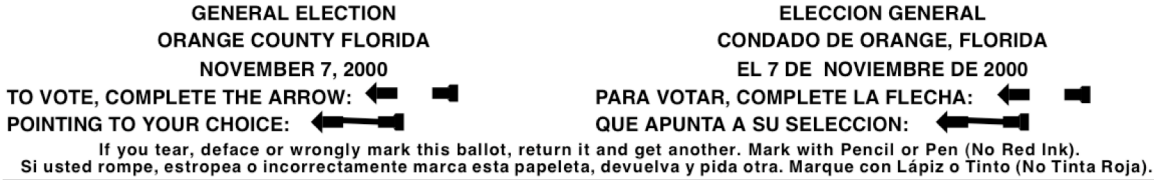

Exhibit 9: Repeatability

	Time	10:27	10:16	11:29	12:20
	Machine	4	4	7	7
Race	Orientation	FUHF	FUHF	FUHF	FUHF
McCain	Count ballots, mark every one	36	510	36	510
O'Connell	Sensitivity, #2 soft lead pencil	10	10	10	10
Hayworth	Sensitivity, #2 pencil erasure	10	10	10	10
Mead	Sensitivity, Sanford sharpie XF pt	10	10	10	10
Huppenthal	Sensitivity, Bic black ballpoint	8	9	9	9
Dial	Pattern	2	58	2	58
McComish	Pattern	4	116	4	116
Orlich	Pattern	6	174	6	174
Robson	Pattern	8	232	8	232
Wegener	Pattern	10	290	10	290
WRITE-IN	Pattern	0	37	0	37
Steel	Sensitivity, Cross blue ballpoint	9	10	10	10
Mayes	Sensitivity, Papermate red ballpoint Med pt	2	2	2	2
Gleason	Sensitivity, Pilot Precise blue rolling ball XF pt	9	9	10	10
Hatch-Miller	Sensitivity, Pilot Precise violet rolling ball XF pt	10	9	10	10
Mundell	Sensitivity, Jelly Roll blue glitter	10	10	10	10
Brock	Sensitivity, Jelly Roll red glitter	0	1	1	1
Russell	Sensitivity, #2 soft lead pencil	10	10	10	10
Ross	Sensitivity, Pilot Razor Point blue marker	9	10	10	10
McCauley	Sensitivity, Millenium purple .05mm	8	9	10	10
Baily	Sensitivity, Pilot Easy Touch red ballpoint	0	0	0	0
Purcell	Sensitivity, Sanford Vis-a-Vis red F pt marker	3	3	4	4
Dowling	Sensitivity, #3H pencil	10	10	10	10