HOW FIRST AND SECOND GRADE CHILDREN ADD WHOLE NUMBERS WITH CALCULATORS

Patricia Baggett & Andrzej Ehrenfeucht

CU-CS-748-94

University of Colorado at Boulder
DEPARTMENT OF COMPUTER SCIENCE

How First and Second Grade Children Add Whole Numbers with Calculators

Patricia Baggett
University of Michigan

Andrzej Ehrenfeucht
Computer Science Department
Box 430
University of Colorado
Boulder, CO 80309-0430
(303) 492-7212
andrzej@piper.cs.colorado.edu

Please send correspondence to:
Patricia Baggett
School of Education
610 East University
University of Michigan
Ann Arbor, MI 48109-1259
(313) 747-0612
fax (313) 763-1229
baggett@umich.edu

Running title: Adding with calculators

HOW FIRST AND SECOND GRADE CHILDREN ADD WHOLE NUMBERS WITH CALCULATORS

Patricia Baggett & Andrzej Ehrenfeucht

CU-CS-748-94 1994

Department of Computer Science
University of Colorado at Boulder
Campus Box 430
Boulder, Colorado 80309-0430 USA

How First and Second Grade Children Add Whole Numbers with Calculators

Abstract

Ninety-six children in five first and second grades used calculators in their math lessons for at least a year, and were given a test consisting of six rather complex whole number addition problems. They had calculators available as they worked the problems, and they wrote their answers rather than selecting answers on a multiple choice test. Children in first and in second grade with experienced teachers performed equally well. For 151 of the 170 incorrect answers (out of 576), a fairly straightforward analysis explained the causes of errors. Of the 151 explainable errors, only 22 were conceptual, meaning that the child showed a lack of understanding of addition. The analysis of errors that children make when they use calculators has the potential to give a considerable amount of information about their understanding of arithmetic, and could lead to individual diagnostics and plans for remediation.

How First and Second Grade Children Add Whole Numbers with Calculators

There has been an interest recently in the United States, as well as in England, Australia, and other countries, in using calculators in elementary mathematics. This has come about in the United States in large measure because of a suggestion in the 1989 Curriculum and Evaluation Standards of the National Council of Mathematics that "....appropriate calculators should be available to all students at all times," together with the publication of the NCTM 1992 Yearbook, titled Calculators in Mathematics Education (Fey & Hirsch, 1992), which examined the potential for calculators to transform the teaching and learning of school mathematics. When four-operation calculators are available in schools, an important question arises: how are children at different ages and grade levels able to use them? Baggett & Ehrenfeucht (1992) found that second graders do not need to be drilled in calculator use, but can learn to use calculators competently in the context of meaningful problem solving. Children in the study had used calculators for almost a year, as a part of their math instruction for project-like, non-drill activities, and errors they made in decimal addition problems with more than two addends were for the most part slips in correct procedures, rather than mistakes that were conceptual in nature.

In this paper we examine answers children in first and second grade, who had used four-operation calculators (Texas Instruments TI-108s) in their math lessons for at least a year, gave to six whole number addition problems. The children had calculators available as they worked the problems, and they wrote their answers on sheets of paper (the problems did not have multiple choice answers). We focus on two questions: (1) How do first and second graders compare in their performance on the problems? (2) When a child has written an incorrect answer, can we determine what process the child went through to make the error? In particular, can we tell if the child's answer shows a lack of understanding of the process of addition, or whether the child seems to understand the process but has made an error in pressing a button on the calculator or in copying the answer from the calculator display to the paper? Are errors indicating a lack of understanding more frequent among first graders than among second graders?

Some reading researchers in the developmental literature suggest that there are differences among children of ages five, six, and seven. For example, Clay (1991, p. 204) presents an idealized graph showing that scores on measures of reading attainment improve as a child's age increases from 5 to 7. She also gives graphical data (pp. 214-215) on children from ages 5 to 7, indicating an increase in the level of the reading book a child attempts, as the child's age increases. Chall (1983) presents a proposal for stages in reading development, with stage 1 (initial reading, or decoding) placed in grades 1 and 2, ages 6 and 7; and stage 2 (confirmation, fluency, ungluing from print) placed in grades 2 and 3, ages 7 and 8. Piaget (1952) found differences in the kinds of errors children of different ages made while performing logical and mathematical tasks, and his theory of stages of cognitive development suggests that children of different ages are at different levels of readiness for learning particular concepts or principles. However, Stallard (1982; cited in Hughes, 1986) found no relation at all between how adequately 6 to 10 year old children demonstrated what symbols such as '6' or equations such as '3 + 1 = 4' mean, and how old the children were. Whether there are differences in performance between first and second graders using calculators has not been investigated in the literature.

Students are now allowed to use calculators on parts or all of some national and state standardized tests, such as the SAT (Scholastic Assessment Test) and the MEAP (Michigan Assessment of Educational Progress). We will also consider implications of our rather straightforward analysis of children's errors for the current practice of requiring children to mark one of four boxes on standardized multiple choice mathematics tests, rather than having them write their answers.

Method

Ninety-six children from five classes, three first grades and two second grades, in an elementary school in a rural south-central Michigan town participated in the study. The classes will be designated by 1a, 1b, 1c, 2a, and 2b ('1' indicating first grade and '2' indicating second grade). There was no tracking in the classes; children of all abilities were present in each class. These five classes were the entire population of first and second graders in the school. Four of the teachers were experienced, having taught at the elementary level at least 15 years. It was the first year of teaching for one of the first grade teachers (who taught class 1a).

Each classroom had a set of four-function calculators, one calculator per child, supplied by the school district (all were Texas Instruments 108 calculators). Calculators were used for mathematics instruction in each classroom once or more per week throughout the year. The calculator lessons were mostly led by the teachers, although at least one demonstration lesson using calculators was given in each class by one of the authors of this paper. Teachers had a total of at least 12 hours of workshops in calculator use and were meeting biweekly with one of the authors to discuss both calculator activities they were trying in their classrooms and new lesson plans incorporating calculators provided to them by the authors of this paper.

In all five classrooms calculators were being used on a pilot basis. So, while they were used at least once a week, most of the children's math lessons did not involve calculators. Teachers used the Silver-Burdett first and second grade math books for most of their lessons (Orfan, Vogeli, Krulik, & Rudnick, 1987).

In the calculator lessons, children were involved in solving meaningful problems; there was no drill and practice. So there were no calculator problems such as, "Add these five numbers: ... ," and no calculator worksheets.

Near the end of the spring semester, the teachers were asked to give their children a test of addition of whole numbers. The test is shown in Figure 1. The purpose of the exercise was not to test their understanding of multidigit numbers, but to test their skill in adding a list of numbers using a calculator, independent of their meaning. Each child had a calculator during the tests, and children were instructed to work alone and to record their answers on their sheets. Teachers were asked not to provide help to the children. The test was given to 21 children in class 1a, 21 in 1b, 17 in 1c, 18 in 2a, and 19 in 2b.

Figure 1. Addition test given to children.
Notice that the sizes of some addends are larger
than those given in typical first and second grade problems.

Name:________________________________

5		12	2310
8	13	9	906
3	35	40	4
<u>4</u>	<u>86</u>	<u>77</u>	<u>5780</u>

1, 23, 46, 3, 5, sum = ______________

975310, 2481056, 120009, sum = ________________

Results and Discussion

The answer for each problem given by each child was read by at least two people. If there was doubt about what the child had written (in about 20 of the 576 answers), the two readers settled on one answer after a discussion or asked a third reader for an opinion in order to reach a final decision. The children's answers were entered into a data file for computer analysis. Table 1 presents the number correct for each problem and for each classroom. Overall, 375 of 576, or 65.1%, of the answers, were correct. The percentages correct were 41.3, 67.5, 74.5, 71.3, and 74.6 in classes 1a, 1b, 1c, 2a, and 2b respectively. No answers were recorded in 31 instances. All of these occurred in class 1a, and there were 3, 3, 5, 5, 7, and 8 answers left blank in problems 1 through 6 respectively.

Table 1. Number correct on addition test,
by problem (1-6) and by classroom
(1a, 1b, 1c, 2a, 2b)
(Number of children in each class who took the test in parentheses)

	Problem number:						
class:	1	2	3	4	5	6	total
1a (21)	13	10	8	10	6	5	52 or 41.3%
1b (21)	19	18	14	11	13	10	85 or 67.5%
1c (17)	14	17	13	13	10	9	76 or 74.5%
2a (18)	14	14	10	14	14	11	77 or 71.3%
2b (19)	18	15	16	9	17	10	85 or 74.6%
Total	78	74	61	57	60	45	375 of 576 or 65.1%

Table 2. Number of problems correct on 6-problem addition test, given by children in each class. (Scores are rank ordered) Median in each class is starred.

child's rank:	Class: 1a	1b	1c	2a	2b
1	6	6	6	6	6
2	6	6	6	6	6
3	5	5	6	6	6
4	5	5	6	6	6
5	5	5	6	6	6
6	4	5	5	5	5
7	3	5	5	5	5
8	3	5	5	5	5
9	3	4	5*	5*	5
10	3	4	5	5	5*
11	3*	4*	4	4	4
12	2	4	4	4	4
13	1	4	3	3	4
14	1	4	3	3	4
15	1	4	3	2	4
16	1	4	2	2	3
17	0	3	2	2	3
18	0	3		2	3
19	0	3			1
20	0	2			
21	0	0			
mean	2.48	4.05	4.47	4.28	4.72

Note: Classes 1a, 1b, and 1c are first grade; classes 2a and 2b are second grade.

Table 2 shows the number of problems that each child in each class answered correctly, and the mean and median for each class. Within a class, the scores in the table are rank-ordered. The data in Tables 1 and 2 show that the scores in classes 1b, 1c, 2a, and 2b are almost identical, and that class 1a scores significantly (over 25%) lower. Further, at least two children in each class score perfectly on the test, and at least five miss at most one problem.

Of the 576 possible answers, 170 (29.5%) contained errors, namely, the answers children wrote did not match the correct answers. We wanted to explain and classify these errors; this is the topic of the next section.

Method of error analysis: Keystroke and transcription errors.

When a child wrote an answer that did not match the correct answer given on the calculator display, we wanted to determine what went wrong. All the wrong answers given for each of the six problems, and how frequently they occurred, are given in Tables 3 through 8. For each problem, we looked for two kinds of errors: keystroke errors (errors in pressing keys) and transcription errors (errors in reading the display or in copying from it).

We first looked for keystroke errors. The diagnosis of these errors for each problem was based on a comparison of two procedures (two sequences of keystrokes). The first procedure gives the correct answer to the problem, and the second gives the actual answer written down by a child. The keystrokes in one procedure were matched with those in the other in order to indicate the mismatches. The wrong answer could then be attributed to one or more specific mismatches.

We note that the tests were administered when we were not present, and the actual keystrokes children made were not recorded in any way, so we do not have information on what they really did while solving the problems. There are many possible procedures which yield a given answer, but only some classes were considered as candidates, as follows.

Our assumption was that a correct procedure was addend + addend + ... + addend =, where addends could be entered in any order. So a correct procedure for a problem with addends a, b, c, and d is [a] [+] [b] [+] [c] [+] [d] [=] . Note that any procedure which gives the correct answer (such as [a][+][b][=][+][c][=][+][d][=], or [a][M+][b][M+][c][M+][d][M+][MRC]) cannot be distinguished from the procedure we assumed.

We also considered only cases in which the calculator was cleared between problems, or in which the last operation in the preceding problem was [=], and in which the operation was [+]. If the actual keystroke sequence a child pressed was different, e.g., if the child used [M+] and did not clear memory between problems, or used [-], such an error was not detected at all and would fall into our "not explained" category.

The second type of error we looked for was a transcription error, namely, a error in reading the display or copying a number on the display to the paper. For example, a 2 on the display might be written as a 5 by a child (or a 5 written as a 2); or '83' on the display might be recorded as '38' by a child.

Diagnosing keystroke and transcription errors was done with the help of an interactive program. All answers given for each of the six problems were entered into the program. Tables 3 through 8 give our analyses of the six problems. In the tables, D means deleted, I means inserted, and R means replaced by. Hypothesized keystrokes or transcription errors take the form xD, yI, xRy. So 12[3D]45[0I]67[8R00]9 means that the sequence 123456789 changed into 1245067009. When the correct answer is 20, and the child recorded 50, this is indicated by [2R5]0, and counted as a transcription error. As another example, consider problem 3 on the test. Its correct keystroke sequence is [12][+][9][+][40][+][77][=], giving the answer 138. One child recorded 205 (see Table 5), which could be gotten as follows:
[1D]2 + 9 + 40 + 77 = [=I], which is 2 + 9 + 40 + 77 = =. (Note that on the TI-108 calculator, the repetition of = adds an extra 77 at the end.)

The best way for the reader to become familiar with the errors is to follow the keystrokes as indicated in Tables 3 through 8, preferably with a calculator in hand. One should realize that the keystroke sequences given for the TI-108 calculator may give different answers when executed on other brands or models of calculators.

A question mark by an answer in the tables means either that we cannot explain what happened, or that, although we can find an explanation, we are less sure that the pattern was actually given by a child.

Table 3. Analysis of problem 1 on test.

problem 1:

```
5
8
3
4
```

result:	keystrokes:	display:	# of cases:		type of error:
20	5 + 8 + 3 + 4 =	20	78 (of 96)		
0	?		1	not explained	U
6	?		1	not explained	U
12	5 + [8D] + 3 + 4 =	12	1		T
16	5 + 8 + 3 + [4D] =	16	1		T
17	5 + 8 + [3D] + 4 =	17	1		T
25	5 + 8 + [3R8] + 4 =	25	1		P
47	5 + 8 + 3 [+D] 4 =	47	1		T
50	5 + 8 + 3 + 4 =	[2R5]0	2		Tr
79	?		1	not explained	U
92	5[+D]8 + 3 [+D] 4 = or 5 + 8 [+D] 3 + 4 =	92	4		T
5834	5834 [+ omitted]		1		C
--	? no answer		3		

error types:
U - unexplained;
P - perceptual (e.g., type 8 instead of 3);
T - typographical error not judged to be perceptual (mispress key(s));
Tr - transposition error: misread display, or miscopy number from display to paper;
C - conceptual error (see Table 10).

summary:
78 correct
3 no answer
12 explained
3 not explained

Table 4. Analysis of problem 2 on test.

problem 2:

13
35
<u>86</u>

answer:	keystrokes:	display:	number of cases:	type of error:
134	13 + 35 + 86 =	134	75 (of 96)	
15	?		1 not explained	U
26	1[+I]3+3[+I]5+8[+I]6=	26	1	C
41	35 + 6 =	41	1	T
99	13 [+35D] + 86 =	99	1	T
116	13 + 35 + [86R68] =	116	1	P
135	13 + 3[5R6] + 86 =	135	1	T
137	13 + 35 + 8[6R9] =	137	1	P
152	[13R31] + 35 + 86 =	152	2	P
154	13 + [3R5]5 + 86 =	154	1	T
155	[13R31] + 35 + 86 =	15[2R5]	1	P,Tr
182	?13 [+ = =I] + 35[= = +I]86 =	182	1	T
302	13[0I] + 86 = =	305	2	T
305	13[0I] + 86 = =	30[5R2]	1	T,Tr
400	?		1 not explained	U
494	138+356={columns added}	494	1	C
133586	133586 [plus ignored]	133586	1	C
--	? no answer		3	

summary:
75 correct
3 no answer
16 explained
2 not explained

Table 5. Analysis of problem 3 on test. (correct answer is 138):

12
9
40
<u>77</u>

answer:	keystrokes:	display	# of cases:	type of error:
138	12 + 9 + 40 + 77 =	138	61 (of 96)	
1	?	1	1 not explained	U
9	? 9 {one addend recorded}	9	1	C
30	1[+I]2+9+4[0D]+7[+I]7=	30	1	C
	or 12 + 9 [+ 40 + 77 D]= =			
61	12 + 9 + 40 [+77D] =	61	1	T
66	1[+I]2+9+40+7[+I]7 =	66	1	C
68	12 + 9 + 40 + [7D]7 =	68	1	T
75	12 + 9 + 40 + 7 [+I] 7 =	75	1	T
84	12 + 9 =[=I]+ 40 + 7[+I]7 =	84	1	T
90	12 + 9 [+40D] + 77 =	9[8R0]	1	T
	or [12+D]9 + 4[0D]+77=	90		
98	12 + 9 [+40D] + 77 =	98	2	T
102	12 + 9 + 4[0D] + 77 =	102	1	T
109	12 [+9D] + [4R2]0 + 77 =	109	1	T
111	[1R3]2 + 9 + 40 + [77D] =	111	1	T
128	12 + 9 + [4R2]0 + 77 =	128	1	T
129	12 [+9D] + 40 + 77 =	129	2	T
130	12 + [9R1] + 40 + 77 =	130	1	T
134	12 + [9R5] + 40 + 77 =	134	1	T
139	1[2R3] + 9 + 40 + 77 =	139	1	T
140	1[2R4] + 9 + 40 + 77 =	140	1	T
150	77 + 40 + 9 + 12 = [=I]	150	2	T
183	12 + 9 + 40 + 77 =	1[38R83]	1	Tr
205	12 [+D] 9 [+40D] + 77 =	20[6R5]	1	T,Tr
	or [1D]2 + 9 + 40 + 77=[=I]	205		
246	12 [+D] 9 + 40 + 77 =	246	1	T
2297	?		1 not explained	U
4098	[12R21][+9D] + 40 [+D] 77 =	4098	1	T
	or 12 + 9 + 40 [+D] 77 =			
400770	40 0 77 0		1 not explained	U
	{random selection from the list}			
7740912	{plus omitted, list bottom up}		1	C
--	? no answer		5	

summary: 61 of 96 correct; 27 explained; 3 not explained; 5 no answer.

Table 6. Analysis of problem 4 on test. (Correct answer is 9000.)

```
 2310
  906
    4
 5780
```

answer:	keystrokes and display:		# of cases		type of error:
9000	2310 + 906 + 4 + 5780 =	9000	57		
135	23[+I]10+9[0D+I]6+4+5[+I]78[0D]=	135	1		C
261	23+1+90+6+4+57+80=	261	1		C
400	2310 + 906 + 4 + [5D]780 =	400[0D]	2		T,Tr
458	23+1+90+6+4+57+80=	458	1		C
787	?		1	not explained	U
820	2310+[9R1]06+4+5780=	820[0D]	1		T,Tr
900	2310 + 906 + 4 + 5780 =	900[0D]	2		Tr
1000	[2R3]310 + 906 + 4 + 5780 =	1000[0D]	1		T,Tr
3300	2310 + 9[0R8]6 + 4 [+5780D] =	3300	1		T
5784	2310+ 906 [+R=] 4 + 5780 =	5784	1		T
6106	[2D]3[1R2]0 + [90D]6 +5780 =	6106	1		T,P
6686	2310 [+R=] 906 + [4D] + 5780 =	6686	1		T
6713	23[10D]+ 906 + 4 + 5780 =	6713	2		T
6920	231[0D] + 90[6R9] + 4 + 5780 =	6920	1		T,P
8100	2310 + [9D]06 + 4 + 5780 =	8100	1		T
8184	2310 + 90[6D] + 4 + 5780 =	8184	1		T
8996	2310 + 906 + [4D] + 5780 =	8996	4		T
9003	2310 + 90[6R9] + 4 + 5780 =	9003	1		P
9888	2310 + 906 + 4 + 5780 =	9[000R888]	1		Tr
10034	2310[+132I]+ 906[= =I] + 5780=	10034	1		T,P
11310	5780 + 4 + 906 + 2310 =[=I]	11310	1		T
18000	2310 + 906 + 4 + 5780[+=I]=	18000	1		T
29764	2310[I0]+90[6R6]+4+57[8R6]0=	29764	1		T
48996	2310 + 906 + 4 [+D] 5780 =	48996	2		T
93735	2310+906[00I]+4[+D]5[+I]780=	93735	1		T
2315484	? 231 ... {first addend?}		1	not explained	U
2310906..	{= ignored, calculator not used}		1		C
--	no answer		5		

summary:
57 correct; 32 explained; 2 not explained; 5 no answer

Table 7. Analysis of problem 5 on test.

problem 5: 1, 23, 46, 3, 5, sum = ______ (correct answer is 78)

answer:	keystrokes:	display:	# of cases:	type of error:
78	1 + 23 + 46 + 3 + 5 =	78	59	
23	{one addend recorded}	23	1	C
36	1 + 23 + 4[6D] + 3 + 5 =	36	1	T
73	1 + 23 + 46 + 3 + [5D] =	73	2	T
75	1 + 23 + 46 [+3D] + 5 =	75	2	T
77	[1+D] 23 + 46 + 3 + 5 =	77	2	T
79	1 + 2[3R4] + 46 + 3 + 5 =	79	1	T
	or 1[+1I] + 23 + 46 + 3 + 5 =			
81	1 + 23 + 4[6R9] + 3 + 5 =	81	1	P
	or 1 + 23 + 46 + 3 +			
	[= =I]+ 5 =			
100	[1R23] + 23 + 46 + 3 + 5 =	100	1	T
	or 23 + [= =I]+ 46 + 3 + 5 =			
102	1 + 23 + 46 + 3 [+D] 5 =	10[5R2]	1	T,Tr
105	1 + 23 + 46 + 3 [+D] 5 =	105	3	T
156	1 + 23 + 46 + 3 + 5 [+=I]=	156	1	T
178	1 [+D] 2[3R4] + 46 + 3 + 5 =	178	1	T
204	1 [+D] 23 + 46 + 3 [+D] 5 =	204	1	T
212	1 [+D][2R1]3 + [46R64]			T,Tr
	+ 3 [+D] 5=	212	1	T,P
	or 1 [+D] 23 + 46 [+35I]			
	+ 3 + 5 =			
492	1 + 23 + 46 [+D] 3 + 5 =	492	1	T
900	? {copied from problem 4?}	9000	1 not explained	U
2108	?1[+2030I]+ 23 + 46 + 3 + 5 =	2108	1 ?	T
4698	? 23 + 40 + 4635 =	4698	1 ?	T
23469	? 1 + 23463 + 5 =	23469	1 ?	T
123468	? 123463 + 5 =	123468	1	C
	{combined strategy}			
1234635	? {plus ignored}		3	C
1294011	? {random typing}		1 not explained	U
51813141	? {random typing}		1 not explained	U
--	? no answer		7	

summary: 59 of 96 correct; 27 explained; 3 not explained; 7 no answer

Table 8. Analysis of problem 6 on test.

problem 6

975310, 2481056, 120009, sum = ______________

correct answer: 3576375

answer:	keystrokes:	display:	# of cases:	type of error:
3576375	975310 + 2481056 + 120009 =	3576375	45	
171	9+7+5+3+1+2+4+8+1+5+6+120=	171	1	C
1001	975+3+1+2+4+8 = =	1001	1	C
1335	975+240+120=	1335	1	C
3575	975310 + 2481056 + 120009 =	35[763D]75	1	Tr
134325	975[3D]10+248[1D]0[5D]6+1200[0D]9 =	134325	1	T
217820	97531 + 280 + 120009 =	217820	1	T
242129	97[5D]310 + 24810[56D] + 120009 =	242129	1	T
252612	?		1 not explained	U
270610	9[7D]5310 + 24[8D]10[56R70] + [1D]20009 =	270610	1	T
377475	9[75D]310 + 2481[0D]56 + 120009 =	377475	2	T
458575	[9753R5]10 + 248[1D]056 + [12R21]0009 =	458575	1	T,P
975387	? 9753 87 {first addend ...}		1 not explained	U
1140143	?		1 not explained	U
1223415	975310 + 248105[6+120009D]=	1223415	1	T
1336375	975310 + 24[8D]1056 + 120009 =	1336375	1	T
1343475	975310 + 2481[0D]56 + 120009 =	1343475	1	T
1376375	975310 + 2[4D]81056 + 120009 =	1376375	1	T
2601065	975310 [+R=] 2481056 + 120009 =	2601065	1	T
2698369	97[5D]310 + 248105[6R0] + 120009 =	2698369	1	T,P
2698575	975[3D]10 + 2481056 + 120009 =	2698575	2	T
2698682	975[3D]1[0R7] + 2481[0R1]56 + 120009 =	2698682	1	T
3276372	975310 + 2481056 + 120009 =	3[5R2]7637[5R2]	1	Tr
3276375	975310 + 2481056 + 120009 =	3[5R2]76375	1	Tr
3456366	975310 + 2481056 [+120009D] =	2698682	1	T
3468375	975310 + 2481056 + 1200[0D]9 =	3468375	1	T
3506375	9[7R0]5310 + 2481056 + 120009 =	3506375	1	T
3576345	975310 + 24810[5R2]6 + 120009 =	3576345	1	P
3576372	975310 + 2481056 + 120009 =	357637[5R2]	2	Tr
3576383	97531[0R8] + 2481056 + 120009 =	3576383	1	T

```
 3576453    975[3R4]1[0R8] + 24810[5R2]6 + 120009 =      3576453    1                  T,P
 4656375    975310 + 2481056 + 12000[0I]9 =              4656375    1                  T
 4704422    975310 + 2481056 + 12[0009R48056] =          4704422    1                  T
 4799741    ? 975310[+95310I]+2481056+12[0009R48065]=    4799741    1 not explained    U
13135186    ?                                                       1 not explained    U
14751918    ?                                                       1 not explained    U
23454823    97[5R3]310 + 2481[05R50]6 + 12000[00I]9 =   23454823    1                  T,P
25905875    975310 + 248105[5I]6 + 120009 =             25905875    1                  T
97651065    975310 [+24810D]56 + 120009 =               97651065    1                  T
975310...   ? {plus ignored, calculator not used}                   2                  C
      --    ? no answer                                             8

summary:
 45 of 96 correct
 37 explained
  6 not explained
  8 no answer
```

We were not able to explain 19 of the 170 wrong answers, or 11.2%. Table 9 shows in which problems, and in which classrooms the unexplained wrong answers occurred. The actual wrong answers can also be found in Tables 3 through 8.

Table 9. Number of unexplained errors, by class and by problem (See also Tables 3 through 8.)

	Problem number:						
class:	1	2	3	4	5	6	total
1a	1	1	2	0	3	1	8
1b	2	1	0	2	0	1	6
1c	0	0	0	0	0	2	2
2a	0	0	0	0	0	1	1
2b	0	0	1	0	0	1	2
total	3	2	3	2	3	6	19

Of the 151 wrong answers we could explain, we categorized them as follows:

(1) a conceptual error: the child does not understand the process of addition. Table 10 gives a description of the five types of conceptual errors and their frequency. The number of these errors by problem and by classroom, is given in Table 11. There were 22 such errors; they are marked 'C' in Tables 3 through 8.

(2) a perceptual error; this was a keystroke error in which, for example, a 3 would be replaced by an 8, or 12 would be replaced by 21. There were 17 such errors, marked by P in Tables 3 through 8.

(3) a typographical error; this was most often an insertion or deletion of a digit (a keystroke error) which was not judged to be perceptual. There were 110 such errors; Table 12 lists them by problem and by classroom; and they are marked 'T' in Tables 3 through 8.

(4) a transcription error; this was an error in recording; for example, a 0 on the display might be recorded as an 8, or a 5 as a 2. There were 21 transcription errors, marked Tr in Tables 3 through 8.

Table 10. Hypothetical explanations of conceptual errors (indicating the child does not understand the concept of addition) in whole number addition with a calculator, and number of instances of each error type.

1. 'Adding numbers' means writing them together (concatenating them).
Example: As an answer to problem 1 (to add 5, 8, 3, and 4), the child writes 5834. (See Table 3.)
Example: As an answer to problem 3 (to add 12, 9, 40, and 77), a child writes 7740912 (the plus is omitted, and the list is given bottom up). (See Table 5.)
Number of errors of this type: 10.

2. Treating single digits as separate numbers in multidigit numbers.
Example: In adding 13, 35, and 86 (problem 2, Table 4), a child writes 26:
[1][+][3][+][3][+][5][+][8][+][6][=] display 26.
Number of errors of this type: 7.

3. (Similar to 2.) Breaking multidigit numbers into two or more not necessecarily single digit numbers, and adding these newly created numbers. (The child does not realize that numbers on a given line are single numbers.)
Example: Child writes 135 as answer to problem 4, with addends 2310, 906, 4, and 5780.
[23][+][10][+][9][+][6][+][4][+][5[+][78][=] display 135.
Number of errors of this type: 2.

4. Concatenating digits in a column and treating them as a number.
Example: In adding 13, 35, and 86 (problem 2, Table 4), a child writes 494:
[138][+][356][=] display 494.
Number of errors of this type: 1.

5. Recording just one of the addends.
Example: In adding 12, 9, 40 , and 77 (problem 3, Table 5), a child writes 9.
Number of errors of this type: 2.

Table 11. Number of conceptual errors (errors showing a lack of understanding), by class and by problem (See also Tables 3 through 8.)

class:	Problem number: 1	2	3	4	5	6	total
1a	1	1	1	1	1	3	8
1b	0	0	0	0	2	0	2
1c	0	1	1	0	1	1	4
2a	0	1	1	2	0	1	5
2b	0	0	1	1	1	0	3
total	1	3	4	4	5	5	22

Table 12. Number of typographical errors, by class and by problem

class:	Problem number: 1	2	3	4	5	6	total
1a	1	4	5	3	4	3	22
1b	0	1	6	6	6	7	26
1c	2	1	3	4	5	4	19
2a	4	1	7	2	4	4	22
2b	1	2	1	7	2	8	21
total	8	9	22	24	21	26	110

Some wrong answers contained two types of errors (e.g., perceptual and typographical, or typographical and transcription), so the number of errors is greater than the number of wrong answers.

The data show that class 1a scored lowest on all but two measures; it was lowest on percentage correct (41.3%), number of answers left blank (31; no other class left any blank); errors we were not able to explain (8, or 42.1%). conceptual errors (8, or 36.4%). and transcription errors (8, or 38%). The other first grade class, 1b, made the most typographical errors (26, or 23.6%; class 1a made 22, or 20%). Class 1b also made the most perceptual errors (mostly reversals of keystrokes): 7 of 17, or 41.2%; class 1a made 4, or 23.5%. We do not know why the performance of class 1a was lowest overall; it may be that the teacher did not have enough time to give the test, and so rushed the children (as evidenced by the 31 problems they left blank). Or the difficulty could be associated with the fact that it was the teacher's first year to teach, while the other four teachers were more experienced.

Classes 1b, 1c, 2a, and 2b are very similar in percentage correct and in error patterns (see Tables 1, 2, 9, 10, and 11). First graders with experienced teachers do not make more conceptual errors than second graders. We have no evidence of a significant difference in performance among the four classes.

There were also no differences in scores for problems presented vertically (numbers 1 through 4) vs. horizontally (numbers 5 and 6) on the test. This result contrasts to results found when children do computations without calculators. Baroody (1987, p. 227) states that children, especially those with learning difficulties, who are working without calculators, may use procedures correctly only when problems are in a familiar form (e.g., column addition), and not on novel assignments (e.g., horizontal addition).

The number of keystrokes required for a problem was related to the percentage correct. Keystrokes required were 8, 9, 11, 16, 12, and 22 for problems 1 through 6 respectively. The percentage correct for each problem is rather closely predicted by the formula $p = [20/(n+16)]*100$, where n is the number of keystrokes required for the problem. Table 13 gives the predicted and actual percentages correct for the six problems. If this formula would hold for larger numbers of keystrokes, it would indicate that the chance of a correct answer is inversely proportional to the number of keystrokes in the calculation.

Table 13. Predicted and actual percentage correct on the six problems, for the four classes 1b, 1c, 2a, and 2b (class 1a is omitted). Predicted value is computed as [20/(n+16)]*100, where n = number of keystrokes in the problem.

problem number	number of keystrokes	percentage correct	predicted percentage correct
1	8	86.7%	83%
2	9	85.3%	80%
3	11	70.7%	74.1%
4	16	62.7%	62.5%
5	12	72%	71.4%
6	22	53.3%	52.6%

Of the 170 incorrect answers given on the test, 143 were different. Of these 143, we could give a fairly straightforward and simple explanation for 132 of them, or 92.3%. The advantage of our error analysis is that, looking at a child's incorrect answer, we can explain in nearly all cases how the error was created. While it is not the objective of this paper, the technique can give an individual diagnostic and the basis for personalized remediation. (Each error can be analyzed individually, and the pattern of errors for a child can be observed.) For example, a second grade child gave the following answers to the six problems:
(1) 20, (2) 134, (3) 138, (4) 900, (5) 78, (6) 3468375.
The child answered problems 1, 2, 3, and 5 correctly.
9000 is the correct answer for problem 4, and the child recorded 900, omitting a 0 from the answer (a transcription error).
For problem 6, the correct answer is 3576375. We think the child's answer came about as follows:
975310 + 2481056 + 1200[0D]9 = display: 3468375

Again the child omitted a 0, but this time it was a typographical error rather than a transcription error. That the child sometimes omits 0 in reading and recording can provide a basis for remediation.

Let us compare the diagnostic capability of a four-alternative multiple choice test. In this case, there are only three wrong answers possible, and therefore the information about what error the child actually made is extremely limited. A simple application of information theory (Shannon, 1949; Goldie & Pinch, 1991) indicates that an answer on a four-alternative multiple choice test provides only at most two bits of information, whereas a number n, written freely, provides $\log_2(n)$ bits (the number of digits when the number n is written in base 2, or binary form). So, for example, in problem 4 (Table 6), when most answers are in the range of several thousand, each answer provides more than 10 bits of information ($\log_2(1000)$), which is five times as much as a four-alternative forced choice test could provide.

This research has uncovered four interesting findings:

(1) For rather complex whole number addition, first and second graders with experienced teachers performed equally well. (A first grade class with a first year teacher performed significantly poorer than all others on almost all measures.)

(2) A fairly straightforward analysis gave a hypothetical explanation for 151 (out of 170) wrong answers that children gave.

(3) Only 22 out of the 151 explainable wrong answers were conceptual, meaning that the child showed a lack of understanding of addition.

(4) With calculators, children in the early elementary grades are able to perform successfully arithmetical calculations that are beyond their hand computation ability.

Our findings are not the same as those of Brown and vanLehn (Brown & VanLehn, 1980; VanLehn, 1982) for hand calculations in multidigit subtraction. For the most part, they claimed that their children were correctly executing an incorrect procedure. Using their terminology (e.g. vanLehn, 1980, p. 12) our students are mostly making "slips" (they know the correct procedure but make small errors in it), and not "bugs," perturbations indicating the procedure is incorrect. We also think their children had harder (more stable) errors, while children in our study were making soft (unstable) errors; our children had not been drilled. Further, the kinds of errors found in hand computation (e.g. the bug of 'smaller from larger' in subtraction; or alignment difficulties resulting in incorrect positioning of digits in addition) are not found when calculators are used.

While analysis of errors made in hand calculation is fairly well established in the literature (e.g., Brown et al, 1980; VanLehn, 1982; vandeWalle, 1990, p. 164), analysis of errors made in calculator computation is new. Our rather simple error analysis was able to account for almost 89% of the errors children actually made (151 of 170). We suggest that, when calculators are used on standardized tests, a more informative approach about the processes children follow would be for children to write their answers, rather than selecting from four predetermined answers, and then to subject their answers to an analysis such as ours. In this way, an individual diagnostic and plan for remediation could be provided.

Acknowledgement. We thank the National Science Foundation for support of this work under grant number ESI-9353068, titled "Incorporating Calculators into Elementary Mathematics."

Bibliography

Baggett, P. & Ehrenfeucht, A. (1992). Do children require drill in order to acquire skill? Talk presented at thirty-third Psychonomic Society Annual Meeting, St. Louis, MO. November 1992.

Baroody, A. (1987). Children's mathematical thinking. New York: Teacher's College, Columbia University.

Brown, J.S. & VanLehn, K. (1980). Repair theory: a generative theory of bugs in procedural skills. Cognitive Science 4 (4) (Oct.- Dec.), 379-426.

Chall, J. (1983). Stages of reading development. New York: McGraw-Hill Book Company.

Clay, M. (1991). Becoming literate: The construction of inner control. Portsmouth, NH: Heinemann Educational Books, Inc.

Curriculum and Evaluation Standards (1989). Reston, VA: National Council of Teachers of Mathematics.

Fey, J. & Hirsch, C. (1992). Calculators in mathematics education. 1992 Yearbook. Reston, VA: National Council of Teachers of Mathematics.

Goldie, C. & Pinch, R. (1991). Communication theory, Cambridge; New York: Cambridge University Press.

Hughes, M. (1986). Children and number. Oxford, UK: Basil Blackwell Limited.

Orfan, L.J., Vogeli, B.R., Krulik, S., & Rudnick, J.A. (1987). Silver-Burdett Mathematics, Morristown, NJ: Silver Burdett & Ginn, 1987.

Piaget, Jean. (1952). The child's conception of number. New York: Humanities Press.

Shannon, C. (1949). The mathematical theory of communication. Urbana, IL: University of Illinois Press.

Stallard, A. (1982). Children's understanding of written arithmetical symbolism. Unpublished MA thesis, University of Edinburgh.

VanLehn, Kurt (1982). Bugs are not enough: Empirical studies of bugs, impasses and repairs in procedural skills. Journal of Mathematical Behavior 3 (2), 3-71.