Tracking Effects of Curriculum Changes in Mechanics using FCI Data
An Action Research Report –––PER Conference, Guelph 2000

Robert A. Morse, St. Albans School, Washington, DC robert_morse@cathedral.org

In 1985 I began changing the way I taught two-dimensional motion by using
1. Computer simulations (1985-89)
2. Computer simulations in conjunction with lab experiences (1989-91)
3. Computer simulations, lab experiences and class discussions(1991-present)

Initially, I used my own tests to gauge the effect of the changes. (1985-1991)
Later, I used sub tests of the Force Concept Inventory (Hestenes, Wells & Swackhamer, 1992) (1992-2000)

In 1997, I changed my approach to teaching Newton's Third Law, and have used the new approach since (1997-2000)

I will consider four questions:

1. Are my students different in their FCI performance before instruction?

2. Are my students different in their FCI performance after instruction?

3. Does the effect of a curriculum change persist over time, even after becoming 'routine'?

4. Are differences in my student's FCI performance after instruction attributable to the curriculum change or merely to having me as their teacher?

Curriculum Changes in Two-dimensional Motion

Barbara White (1984) described the use of DiSessa's Dynaturtle in a set of computer games for teaching Newtonian Concepts of motion in two dimensions. I adapted this idea to my classroom, and programmed version of the games initially in Pascal on the Apple II+ and later in cT for the Macintosh. My initial use of the games as tracked by my own test questions was disappointing. Students seemed to switch into a "game mode"of thinking and did not draw generalizations from the tactics they used to reach the game goals.

An AAPT workshop with Priscilla Laws (1989) introduced me to the technique proposed by Redish and Taylor (1987) of using impulsive taps on real objects in the lab to see the effect of impulsive forces. This mirrored the impulsive forces applied to the Dynaturtle, and I introduced this technique to my teaching program. At the same time, I had changed the order of teaching topics so that I now developed Newton's second law in one dimension before teaching two dimensional motion (Morse, 1992). This allowed me to present the argument that since a force had to be applied to cause a change in direction of motion, there must be an acceleration in that direction. If some kind of force must be applied towards the center of a circle to cause circular motion, then there must be a center-pointing acceleration as a consequence. Finally, I integrated the computer simulations, the lab experiences, and thought experiments in class discussion into a coordinated unit.

As I made the changes, I tried to keep some track of their effect by comparing test question results from year to year,using questions I made up myself. For example on one question about acceleration in circular motion, correct responses increased as follows:

1986 63%, 1988 77%, 1990 95%,

giving me some reason to think my curriculum changes were successful.

Comparison with other curricula using the Force Concept Inventory

In 1992, the Force Concept Inventory was published along with detailed comparison data on student responses. I began using this as a post-test, and later as a pretest and post-test for my physics classes. One question I was interested in was the effect of my two-dimensional motion curriculum. Also, in 1993 Barbara White published results of a study using her ThinkerTools computer simulations to teach similar concepts to sixth graders. Several of her test questions were similar to FCI questions, and allowed me to make a second comparison with my curriculum.

In 1992 and 1993, I compared my students post-test scores with data from these sources on 11 questions from the FCI. Seven of the FCI questions dealt directly with aspects of two-dimensional motion addressed by my teaching approach, and four more are problems which students should be able to answer by transfer of the ideas to a new context. The two sets of questions are thus divided into two sub-tests, 2-dim direct and 2-dim transfer.

Table 1 compares FCI responses on the 11 questions for my 1992 and 1993 students with two groups using standard curricula, Arizona regular (N=612) and Arizona honors (N=118) and two groups using modeling method curricula, Malcolm Wells honors (N=30) and Greg Swackhamer honors (N=63), using the data reported in the 1992 FCI paper.

Table 2 compares sixth grade responses to four ThinkerTools questions with my 1993 students responses to four similar FCI questions.

From these results I concluded in 1993 that my changes were successful.

Table 1: Force Concept Inventory Results Per Cent Correct

Question Number New(Old)		Standard Teaching			Experimental Group-RAM			High -Acheiving Comparison Groups	
Direct		AR	AH		92RM	93RM		MW	GS
8	(6)	58	63	<	82	100	≈	97	60
9	(7)	33	46	<	48	52	≈	52	40
10	(8)	53	64	<	85	89	≈	90	71
21	(24)	22	41	<	61	52	>	45	30
22	(25)	38	56	<	67	67	>	62	52
23	(26)	42	48	<	79	78	>	60	41
24	(27)	77	84	≤	91	85	≈	97	83
Transfer									
7	(4)	72	66	<	91	85	≈	83	90
6	(10)	88	91	≈	85	85	≈	93	87
12	(16)	67	83	≈	79	78	≈	76	65
14	(23)	47	59	<	63	64	<	69	75

AR = scores reported for Arizona high school regular students (Hestenes et al, 1992)
AH = scores reported for Arizona high school honors students (Hestenes et al, 1992)
92 RM = Force concept inventory questions given on 1992 final exam by R. Morse
93 RM = Force concept inventory questions given on 1993 final exam by R. Morse
MW = scores reported for Malcolm Wells Honor students (Hestenes et al, 1992)
GS = scores reported for Gregg Swackhamer Honor students (Hestenes et al, 1992)

Old Question numbers from 1992 29 question version of FCI.
New Question from present 30 question version of FCI.

Table 2 : ThinkerTools Test Comparison
White (1993)

ThinkerTools question #	FCI Equivalent New(old)	93RM Percent correct		TT
2	6 (10)	85	>	49
3	12 (16)	78	=	83
5i	21 (24)	52	=	51
5ii	23 (26)	78	=	83

93RM = FCI questions given on 1993 final exam by R. Morse.
TT = Thinker Tools students scores on transfer test 9.

Question 1. Are my students different in their FCI performance before instruction?

I teach in a private boys school with a competitive admissions policy. One confounding variable could be that my students are already different before instruction. In my 1992 and 1993 administrations of the FCI, on which I based my previous conclusion, I used the FCI only as a post-test. Since 1994, I have given the FCI as both pretest and post-test.

Table 3 shows the FCI pretest results for my physics students in 1994, 1997, 1999, and 2000 on each question of the 2-Dim Direct, 2-Dim Transfer and Third Law sub tests of the FCI. In addition, means for each sub test are given. For comparison, the same question and sub test results are given for the Arizona regular, Arizona honors, Malcolm Wells honors and Greg Swackhamer regular groups described earlier.

(The “odd-ball” data selection occurs because a flood in my office destroyed the records for individual FCI responses for 1995 and 1996, and 1998 data were not available.)

Inspection of the table suggests that my students do a little better on the 2-Dim Direct sub test than the regular groups (AZ-R & GS) and the same or worse than the honors groups (AZ-H & MW). They are closer to the regular groups on the 2-Dim transfer sub test and the Third Law sub test. In all three cases they do not do as well as the MW honors group.

Overall, my students are fairly comparable on pretest scores to the students for which I have comparison data.

FCI Question #		Pretest Results							
New	Old	2-D Direct Sub test							
		AZ-R	AZ-H	MW	GS	RM-00	RM-99	RM-97	RM-94
2	6	33	39	47	31	38	31	50	31

Tab

FCI Question #			Pretest Results						
New	Old		2-D Direct Sub test						
		AZ-R	AZ-H	MW	GS	RM-00	RM-99	RM-97	RM-94
8	**6**	33	39	47	31	39	21	50	31
9	**7**	27	27	34	28	49	45	27	33
10	**8**	18	26	50	25	46	32	37	48
21	**24**	19	20	25	19	22	39	30	27
22	**25**	30	31	61	28	17	32	40	33
23	**26**	32	44	58	25	20	18	40	39
24	**27**	52	60	83	58	63	61	63	67
MEAN		**30**	**35**	**51**	**31**	**37**	**35**	**41**	**40**
			2-D Transfer Sub test						
7	**4**	42	48	47	39	56	45	40	60
6	**10**	63	70	72	53	83	74	63	67
12	**16**	33	42	61	27	29	47	43	33
14	**23**	26	31	42	38	24	21	17	29
MEAN		**41**	**48**	**56**	**39**	**48**	**47**	**41**	**47**
			Third Law Sub test						
4	**2**	21	22	17	14	17	16	17	19
28	**11**	8	13	22	5	2	13	13	12
15	**13**	8	6	14	8	7	3	7	13
16	**14**	35	39	56	30	22	24	37	35
MEAN		**18**	**20**	**27**	**14**	**12**	**14**	**18**	**20**

AZ-R =Arizona high school regular students AZ-H = Arizona high school honors students
MW = Malcolm Wells Honor students GS = Greg Swackhamer Honor students

RM-94 =FCI questions given on 1994 final exam by R. Morse
RM-97 =FCI questions given on 1997 final exam by R. Morse
RM-99 =FCI questions given on 1999 final exam by R. Morse
RM-00 =FCI questions given on 2000 final exam by R. Morse

Question 2 and Question 3

Graph 1 shows year by year results on the FCI 2-Dim direct post-test questions from 1992 through 2000. Graph 2 similarly shows results on the 2-Dim transfer test. Individual question averages are shown by bars, and the sub-test averages are shown by the line graph. In addition, the sub-test average for the same set of questions is shown for the four comparison groups, AZ-Reg, AZ-Hon, MW, and GS.

Question 2. Are my students different in their FCI performance after instruction?

Comparison of the post-test averages shows that my students performed comparably to those of Malcolm Wells and Greg Swackhamer, and noticeably better than the groups receiving standard instruction. My students are different from the standard instruction groups after instruction.

Question 3. Does the effect of curriculum change persist after becoming 'routine'?

Inspection of the post-test averages and the individual question averages shows some fluctuation with time, but no particular trend, so it seems that even though this curricular change has become a 'routine' part of my teaching, it continues to be effective with students.

Question 4. Are differences in my student's FCI performance after instruction attributable to the curriculum change or merely to having me as their teacher?

Because I did not use the FCI during the time from 1985 to 1992 while I was making the largest changes in my teaching sequence for two dimensional motion, it is possible to argue that my students' high scores on these sub-tests are due to some personal characteristic of my teaching rather than the particular curricular approach I was using. In 1997-98 school year, I changed my method of teaching the concepts associated with Newton's Third Law using a variation of exercises from the CPU Project mechanics curriculum (Goldberg, et al., 1999). For this change, I do have FCI data from years before and after the curriculum change. This permits investigating the extent to which changes in my students' performance depends on me or on the curriculum.

Graph 3 shows FCI post-test for a four question sub test of the FCI that deals with Newton's Third Law. Data are shown for 1994 and 1997, years in which I had used more or less standard instruction, and 1999 and 2000, years in which I used the CPU Project unit. (1995 and 1996 data were lost, and 1998 was not available at this writing). In addition, the sub-test average for the same set of questions is shown for the four comparison groups, AZ-Reg, AZ-Hon, MW, and GS.

The first two years show sub-test averages which are no different from those of the two groups receiving standard instruction. The second two years are noticeably different from the first two, and much closer to the post-test averages achieved by Greg Swackhamer's and Malcolm Wells's classes. This suggests that the difference is due to the curriculum unit and not to the personal characteristics of the instructor.

References:

Goldberg, F. (1999) The CPU Project Curriculum, http://cpu.sdsu.edu/CPU/

Hestenes, D. Wells, M. Swackhamer, G. (1992) *The Physics Teacher* **30**, 141-158

Laws, P., Thornton, R., Sokoloff, D. (1989) Teaching Physics as a Workshop Course, AAPT Summer Meeting, San Luis Obispo, CA

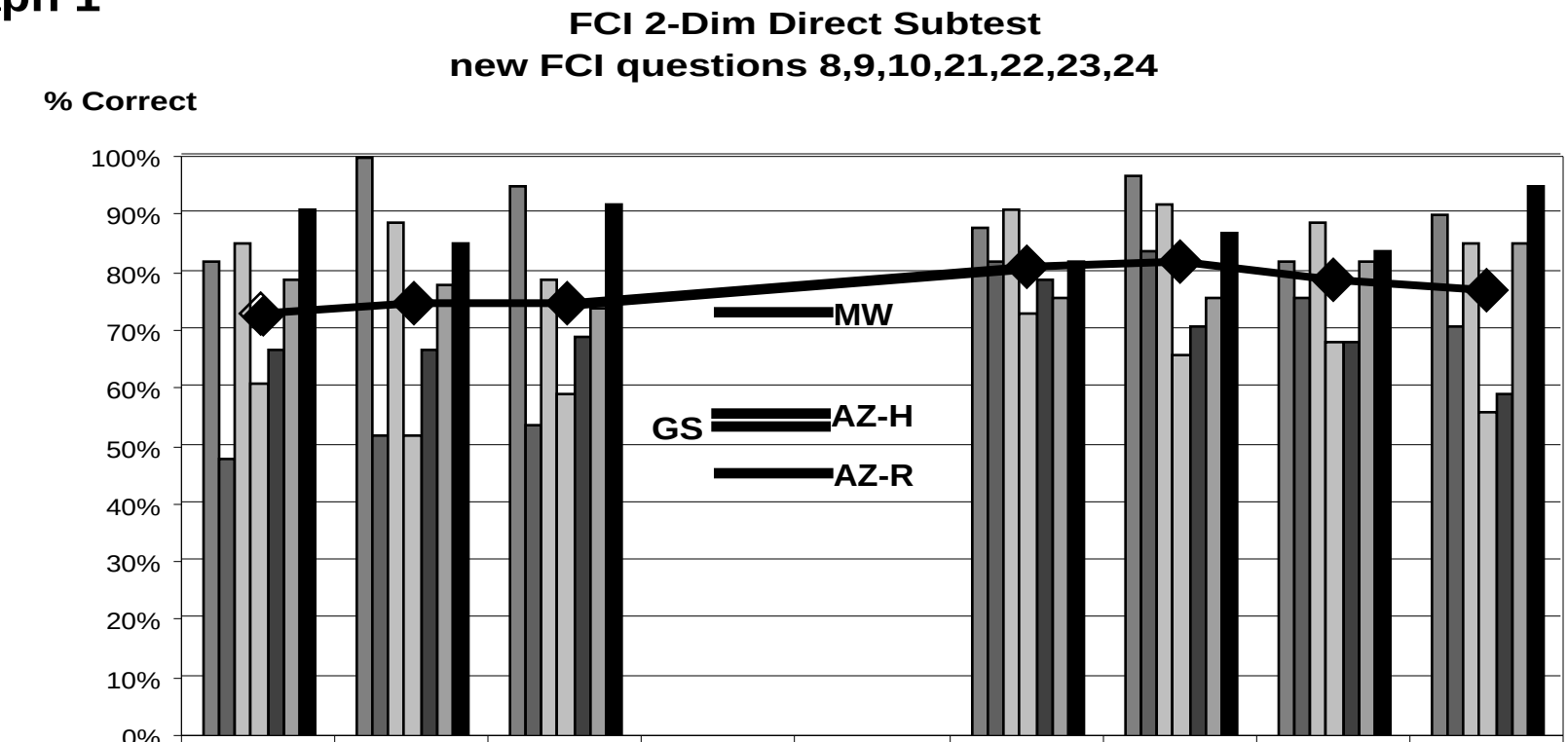
Morse, R. A. (1992) A Teachable Sequence for Mechanics. In New Mechanics Advisory Conference, Tufts University, Medford MA

Redish, E.F., Taylor, E. F. (1987) *AAPT Announcer* **17**, 82

White, B. Y. (1984) *Cognition and Instruction* **1**, 69

White, B. Y. (1993) *Cognition and Instruction* **10**, 1-100

Graph 1

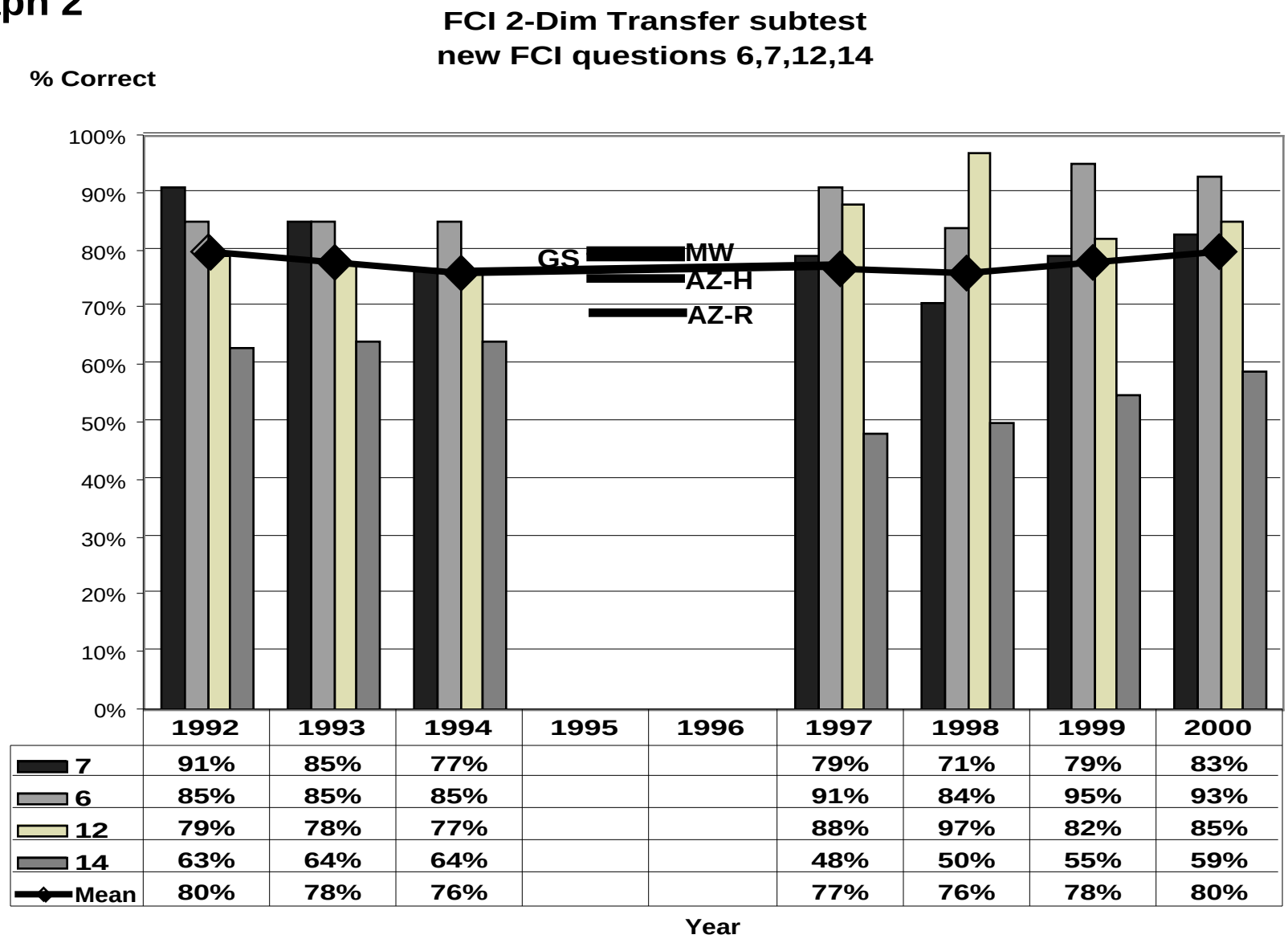


	1992	1993	1994	1995	1996	1997	1998	1999	2000
8	82%	100%	95%			88%	97%	82%	90%
9	48%	52%	54%			82%	84%	76%	71%
10	85%	89%	79%			91%	92%	89%	85%
21	61%	52%	59%			73%	66%	68%	56%
22	67%	67%	69%			79%	71%	68%	59%
23	79%	78%	74%			76%	76%	82%	85%
24	91%	85%	92%			82%	87%	84%	95%
Mean	73%	75%	75%			81%	82%	79%	77%

Year

Comparison Sub test means from Hestenes, Wells & Swackhamer 1992
MW 73% GS 53% AZ-H 57% AZ-R 46%

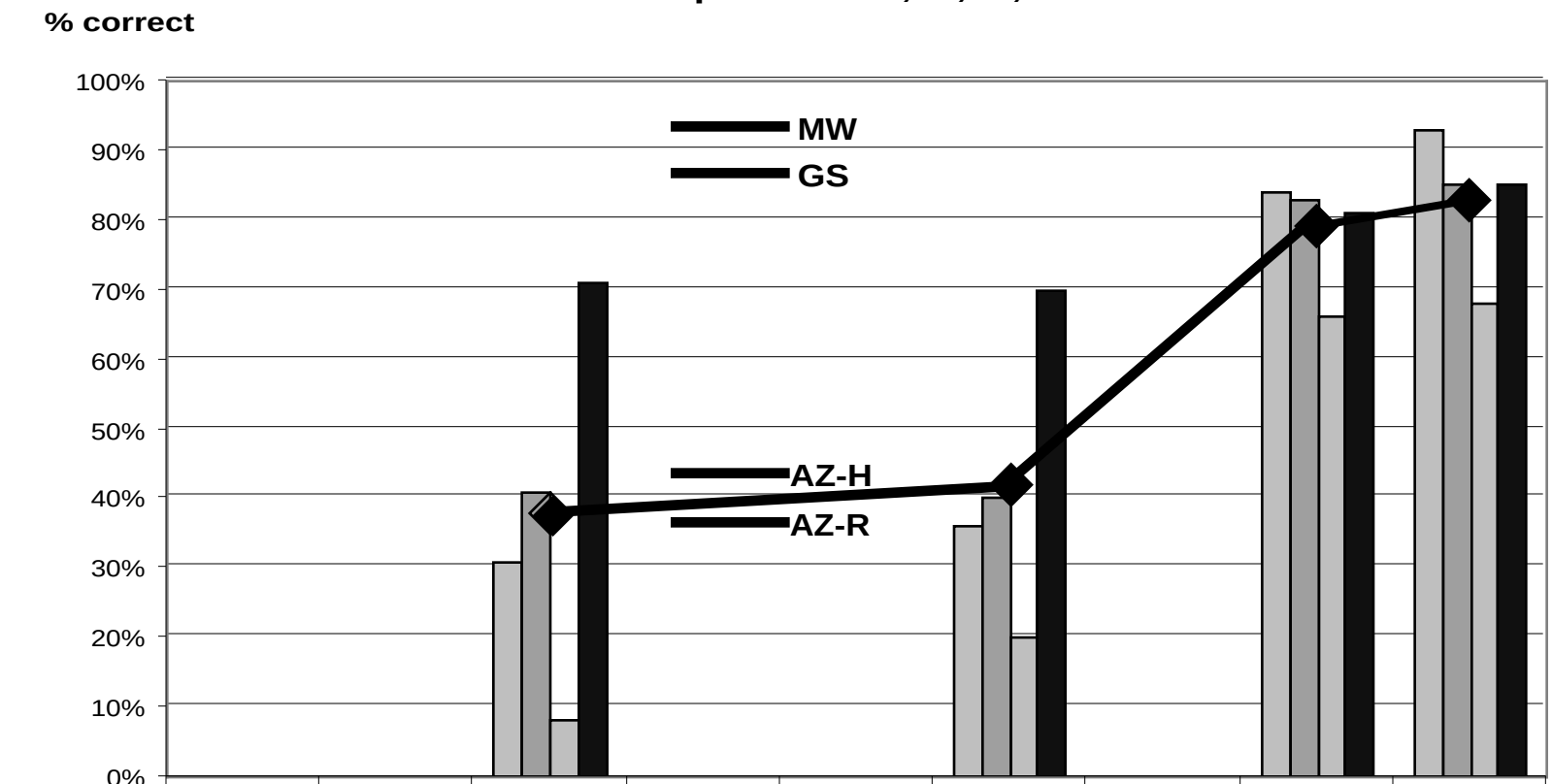
Graph 2



	1992	1993	1994	1995	1996	1997	1998	1999	2000
7	91%	85%	77%			79%	71%	79%	83%
6	85%	85%	85%			91%	84%	95%	93%
12	79%	78%	77%			88%	97%	82%	85%
14	63%	64%	64%			48%	50%	55%	59%
Mean	80%	78%	76%			77%	76%	78%	80%

Comparison Sub test means from Hestenes, Wells & Swackhamer 1992
MW 80% GS 79% AZ-H 75% AZ-R 68%

Graph 3



	1992	1993	1994	1995	1996	1997	1998	1999	2000
4			31%			36%		84%	93%
28			41%			40%		83%	85%
15			8%			20%		66%	68%
16			71%			70%		81%	85%
Mean			38%			42%		79%	83%

Year

Comparison Sub test means from Hestenes, Wells & Swackhamer 1992
MW 92% GS 86% AZ-H 43% AZ-R 37%