

Assessing Children's Reading Improvement in Summer School Programs



CIERA

Center for the Improvement of
Early Reading Achievement

www.ciera.org

Research Team

- **University of Michigan**
 - Scott Paris, Robert Carpenter, Jennifer DeGroot, Jonathan Flukes, Alison Paris
- **Michigan State University**
 - David Pearson, Gina Cervetti, Joseph Martineau
- **Ingham Intermediate School District**
 - Kathy Humphrey, Tamara Bashore-Berg

Goals

- To evaluate the effectiveness of K-3 summer reading programs in sample sites in Michigan
- To develop assessment tools for K-3 literacy as part of the MLPP
- To provide suggestions to improve K-3 summer reading programs

Three One-Year Projects

- Year 1: Summer 1998
 - 6 sites, no control group
- Year 2 : Summer 1999
 - 12 sites, experimental design
- Year 3: Summer 2000
 - 4 model sites to document programs

Year 1: Summer 98

- K-3 programs in 6 MI cities evaluated
- Main data from 185 children in grades 1, 2, & 3 who had complete pre- and post-test data (49% third graders, 45% males, 55% Caucasian)
- QRI-II given individually at beginning and end of summer school using the same passages
- DRA and DRP given during summer school
- Classroom observations & teacher logs

Year 1: Results on QRI-II

Significant improvements on:

- Words read correctly in lists
- Oral reading accuracy
- Oral reading rate
- Comprehension
- Retelling

Additional Questions

- Which children improved their reading and how?
- Did summer school programs help children's reading skills generally or selectively?

Posthoc Groups Based on Accuracy and Comprehension Scores

Good Readers

n = 43, Males = 19 Females = 24

First = 3, Second = 11, Third = 29

Mean Student Grade Level = 2.6

Mean Passage Level = 3.7

Word Callers

n = 44, Males = 18 Females = 26

First = 3, Second = 11, Third = 30

Mean Student Grade Level = 2.6

Mean Passage Level = 4.4

Gap Fillers

n = 42, Males = 22 Females = 20

First = 14, Second = 8, Third = 19

Mean Student Grade Level = 2.1

Mean Passage Level = 2.1

Poor Readers

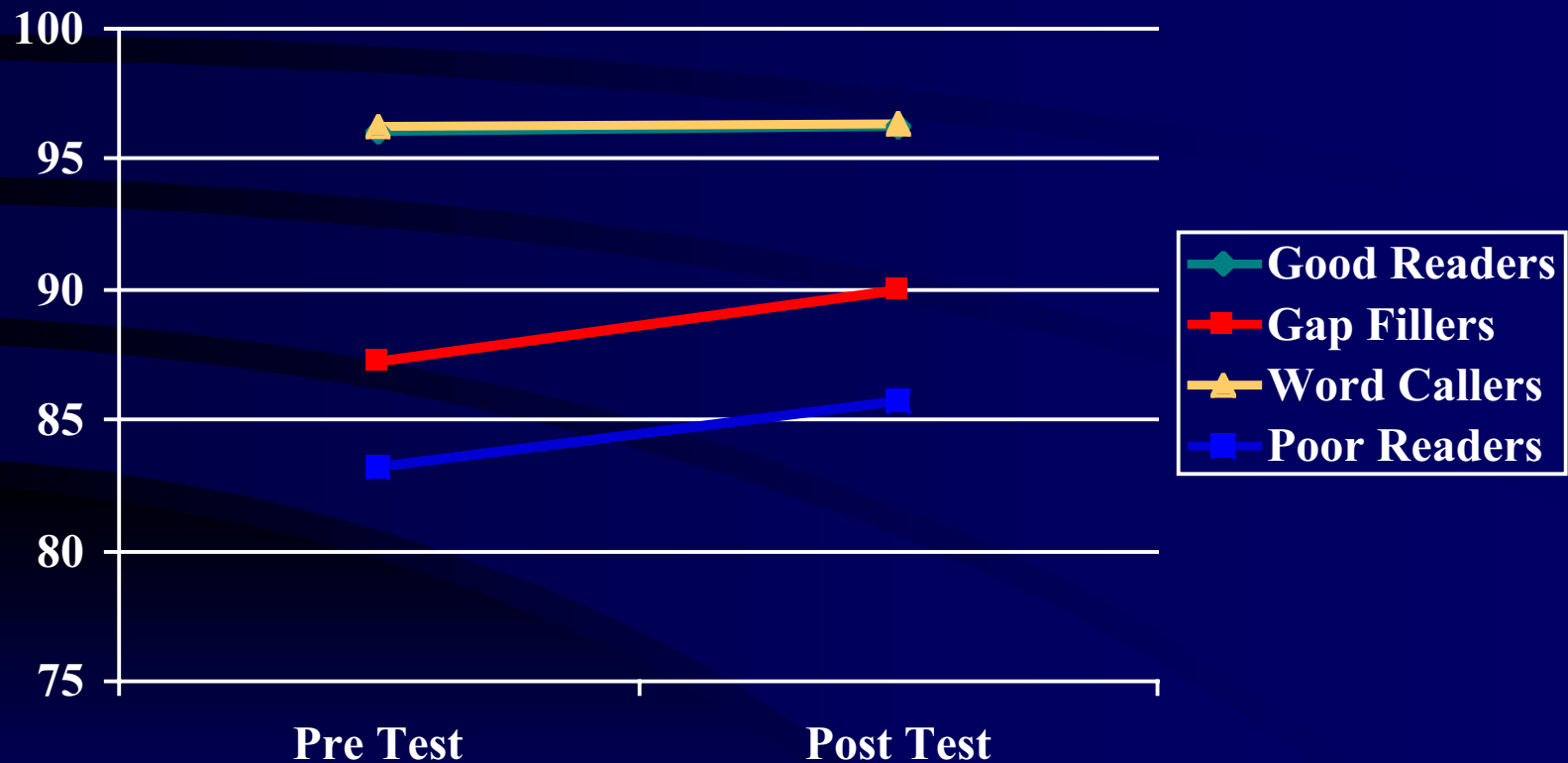
n = 50, Males = 22 Females = 28

First = 27, Second = 12, Third = 11

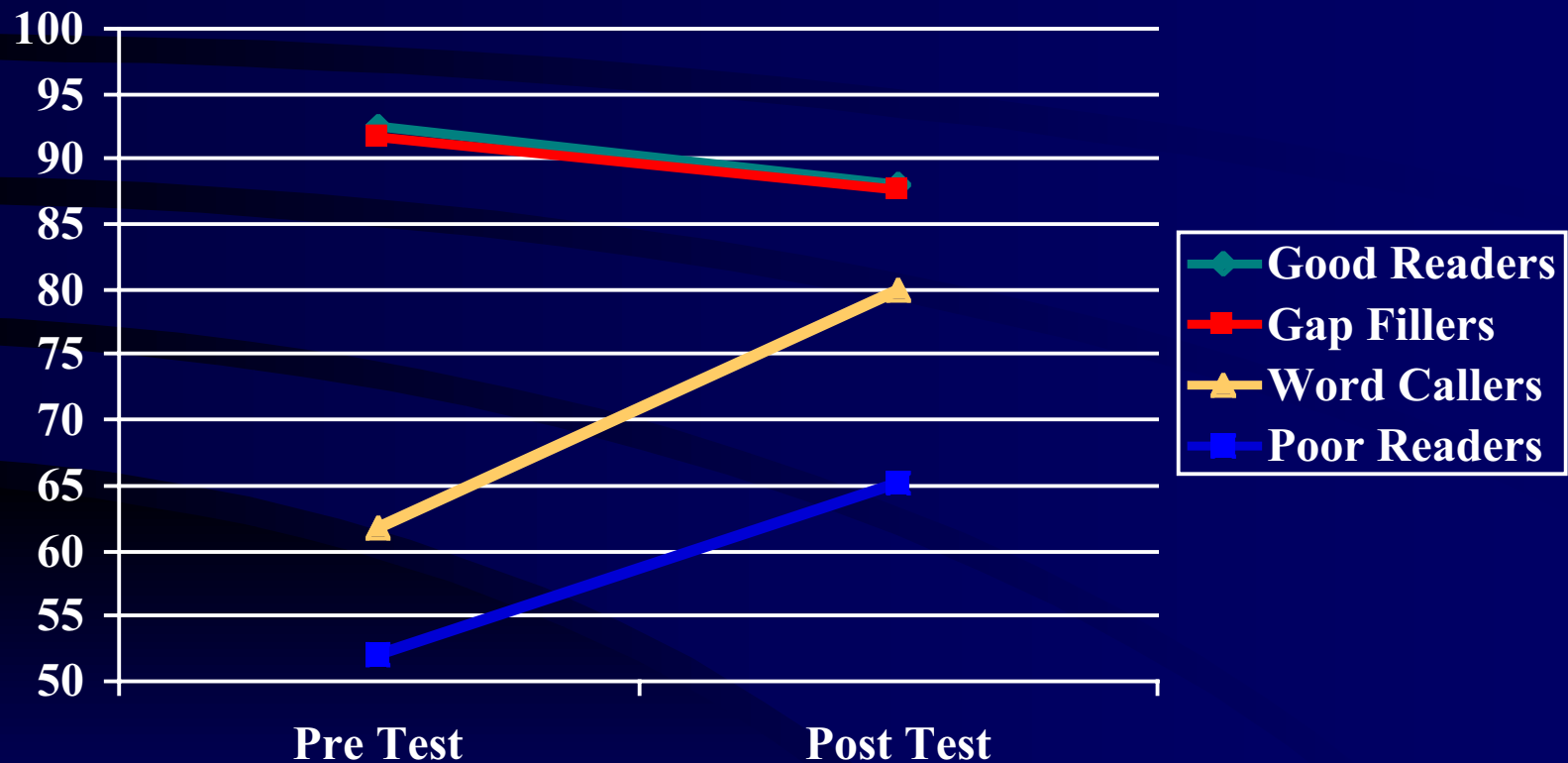
Mean Student Grade Level = 1.7

Mean Passage Level = 2.6

% Gain Scores in Oral Reading Accuracy on the QRI by Group



% Gain Scores on QRI Comprehension by Group

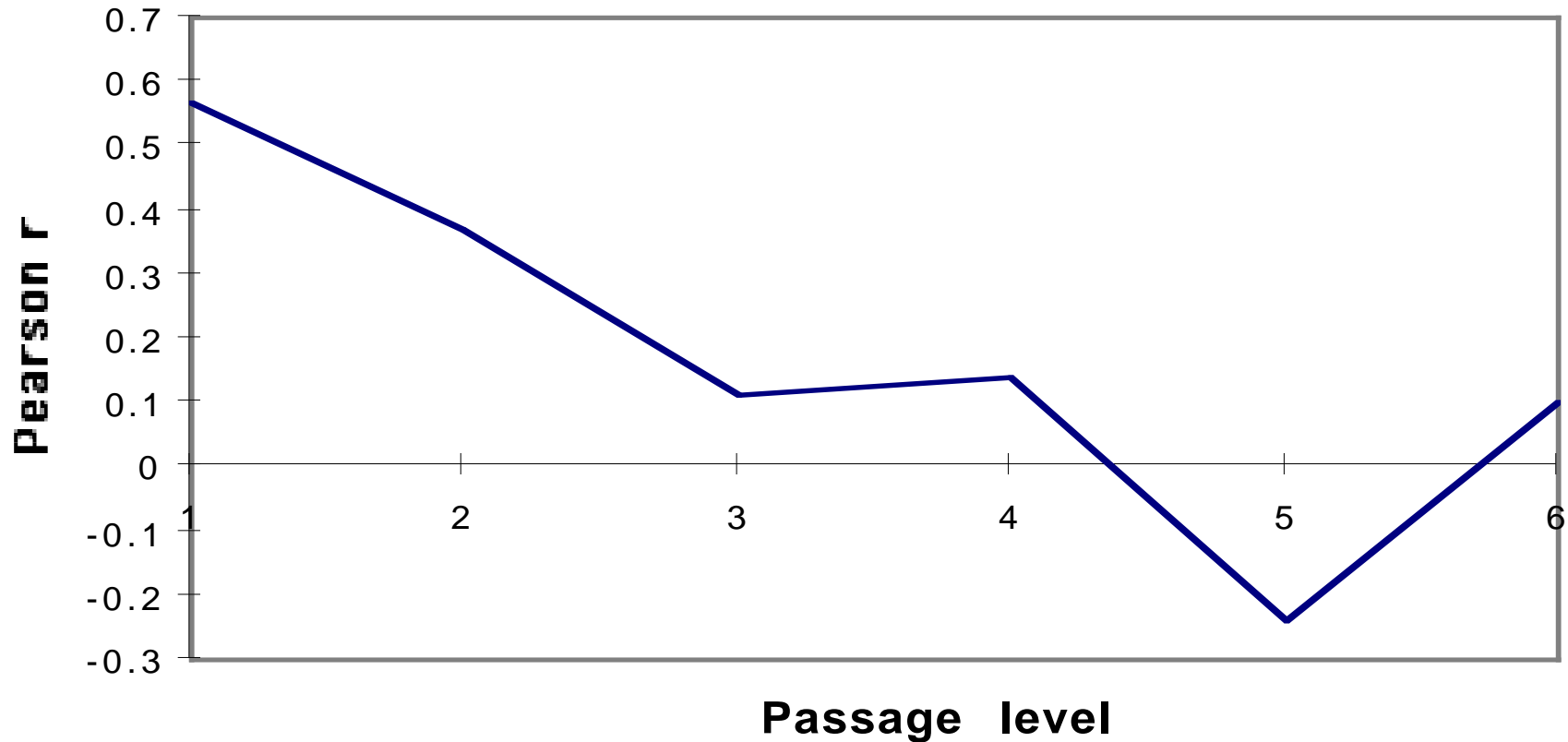


Are Oral Reading Accuracy and Comprehension Independent?

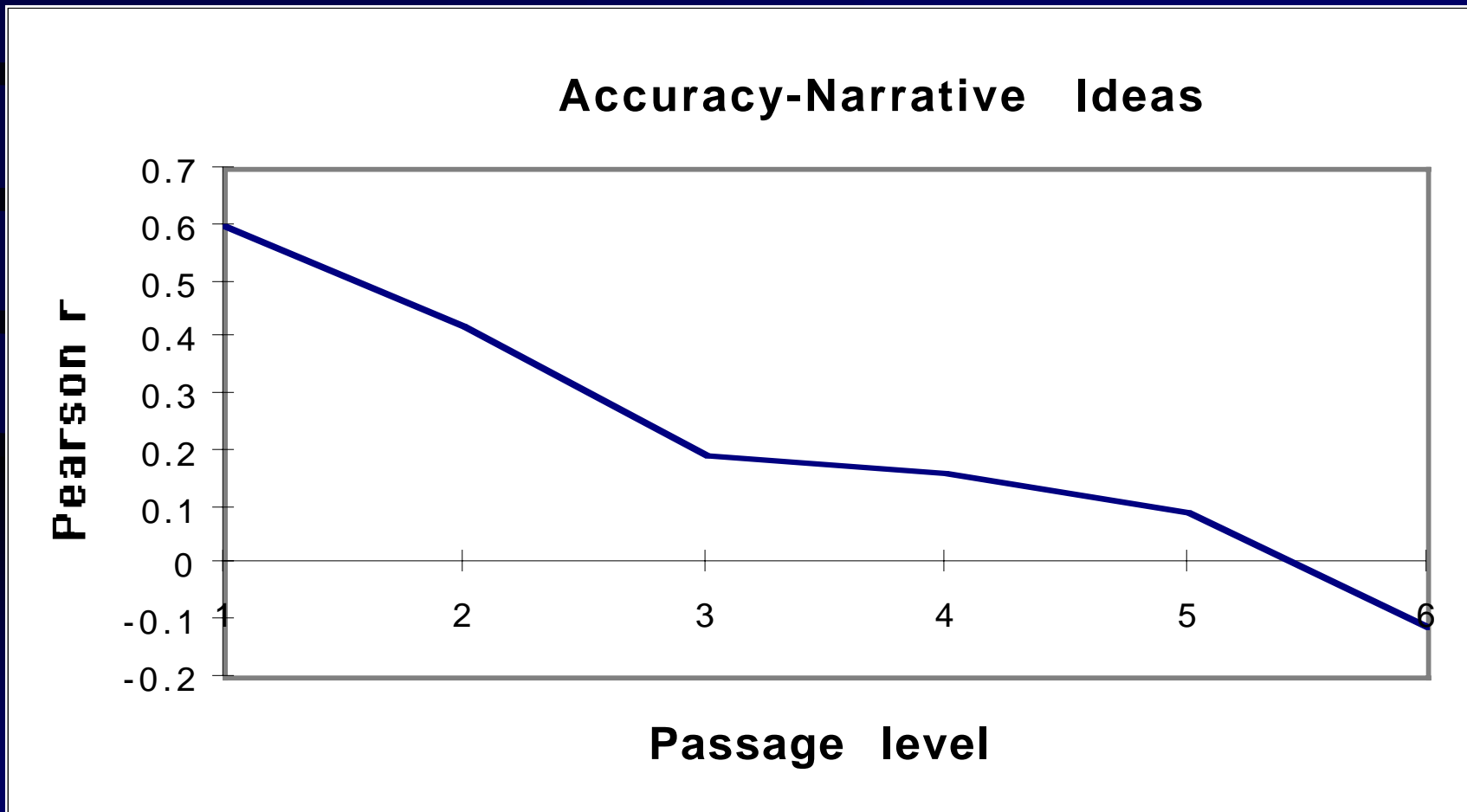
- Factor analysis revealed two independent factors
- Gains from pre-test to post-test appear specific, not general
- Accuracy is NOT correlated with comprehension for more difficult passages

Correlations By Passage Between Accuracy and Comprehension

Accuracy-Comprehension



Correlations By Passage Between Accuracy and Recall of Ideas



Why do correlations between accuracy and comprehension/recall decrease with increasing passage difficulty?

Possible Explanations

- Nature of the text
 - More difficult passages
 - More difficult questions
- Nature of oral reading
 - Takes all cognitive resources
 - Attention to self-presentation
- Nature of the readers
 - Older children have less practice reading aloud
 - More word callers in remedial programs

Conclusions About Oral Reading

- Oral reading accuracy may become less related to comprehension among older, better readers
- Measures of oral reading accuracy, like running records and miscue analyses, may be limited to word identification
- Fluency involves more than accuracy and may be related better to comprehension

Conclusions From Year 1

- **Good news:**
 - Children read same passages better after summer school
 - Observations and teacher logs revealed features of effective programs
- **Worries:**
 - No control for practice or maturation
 - No control group w/o summer school

Another Concern: Validity

- We converted children's scores on the QRI-II and the DRA to DRP text difficulty units and computed the correlations among them:
 - DRA & QRI: $r = .75$
 - DRP & DRA: $r = .59$
 - DRP & QRI: $r = .41$
- Conclude: modest concurrent validity

Recommendations for Instruction in Year 2 Programs

- 📖 Provide daily opportunities to:
 - 📖 read familiar texts
 - 📖 read challenging materials with teacher guidance
 - 📖 write authentic texts
- 📖 Provide explicit instruction on decoding, comprehension, and writing skills
- 📖 Enhance student motivation
- 📖 Promote parent involvement

Other Desirable Characteristics

- ★ Manageable class sizes and teacher-student ratios
- ★ Knowledgeable and experienced staff
- ★ Professional staff development opportunities
- ★ High quality instructional materials
- ★ Effective use of libraries/media resources
- ★ Minimum 60 hours instructional time on reading
- ★ Thematic & connected instruction
- ★ Connections with regular curriculum & teachers
- ★ Program leadership and accountability

Year 2: Design

- K-3 summer programs in 12 MI districts
- Tested >1000 children who were eligible or recommended for summer school
- Pretest: Spring 1999
- Posttest: Fall 1999
- Delayed post test: Spring 2000
- Compare Experimental & Control students

Measures for Year 2

- Gates-MacGinitie Reading Tests for readers
- Johns Basic Reading Inventory (BRI)
- Literacy Habits
- Student Opinions About Reading (SOAR)
- MLPP tasks for non-readers
 - Concepts of Print, Phonemic Awareness, and Letter Identification

The Gates-MacGinitie Reading Tests

- Level PRE
 - Literacy Concepts, Reading Instruction & Relational Concepts, Oral Language Concepts, Letter-Sound Correspondences
- Level R
 - Beginning Consonants, Final Consonants, Vowels, Use of Context
- Levels 1, 2, 3
 - Vocabulary, Comprehension, Total

Pros & Cons of a Standardized Test

Benefits

- Group administered in about an hour
- Multiple forms for pre & post testing
- Subscores and scaled scores
- Administrators want/expect standardized test

• Liabilities

- Young children unfamiliar with format
- Children distressed
- Not aligned with curricula & instruction
- May measure ability not achievement

Level PRE Sample

N=138

Status

Experimental	58%
Control	42%

Gender

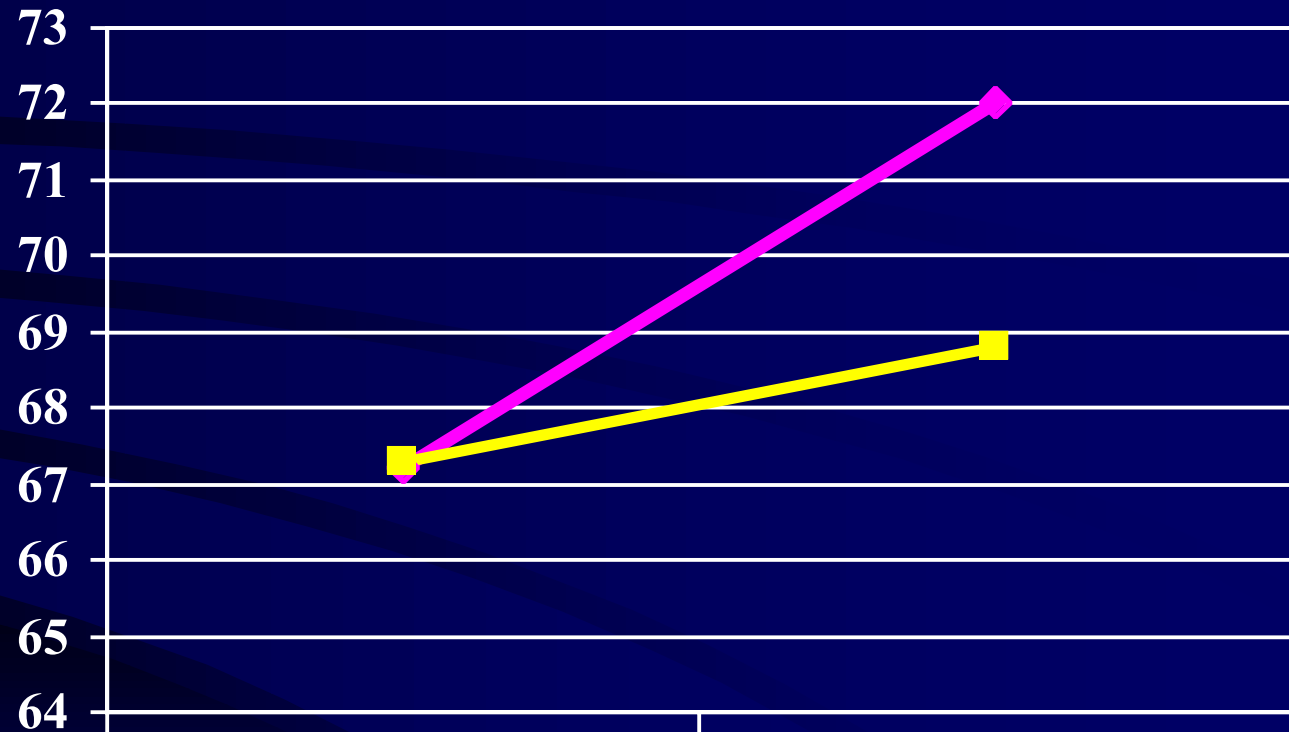
Male	62%
Female	38%

Race

Caucasian	57%
African-American	27%
Latino	9%
Other	7%

Level PRE Results

Raw Score



	Pretest	Posttest
Experimental	67.2	72
Control	67.3	68.8

Levels R-3

Modified Sample

We deleted...

- High achieving readers (>1.5 SD above mean on pretest)
- Students who were given different levels of Gates for pretest and posttest

Levels R-3 Sample

N=566

Status

Experimental	55%
Control	45%

Gender

Male	55%
Female	45%

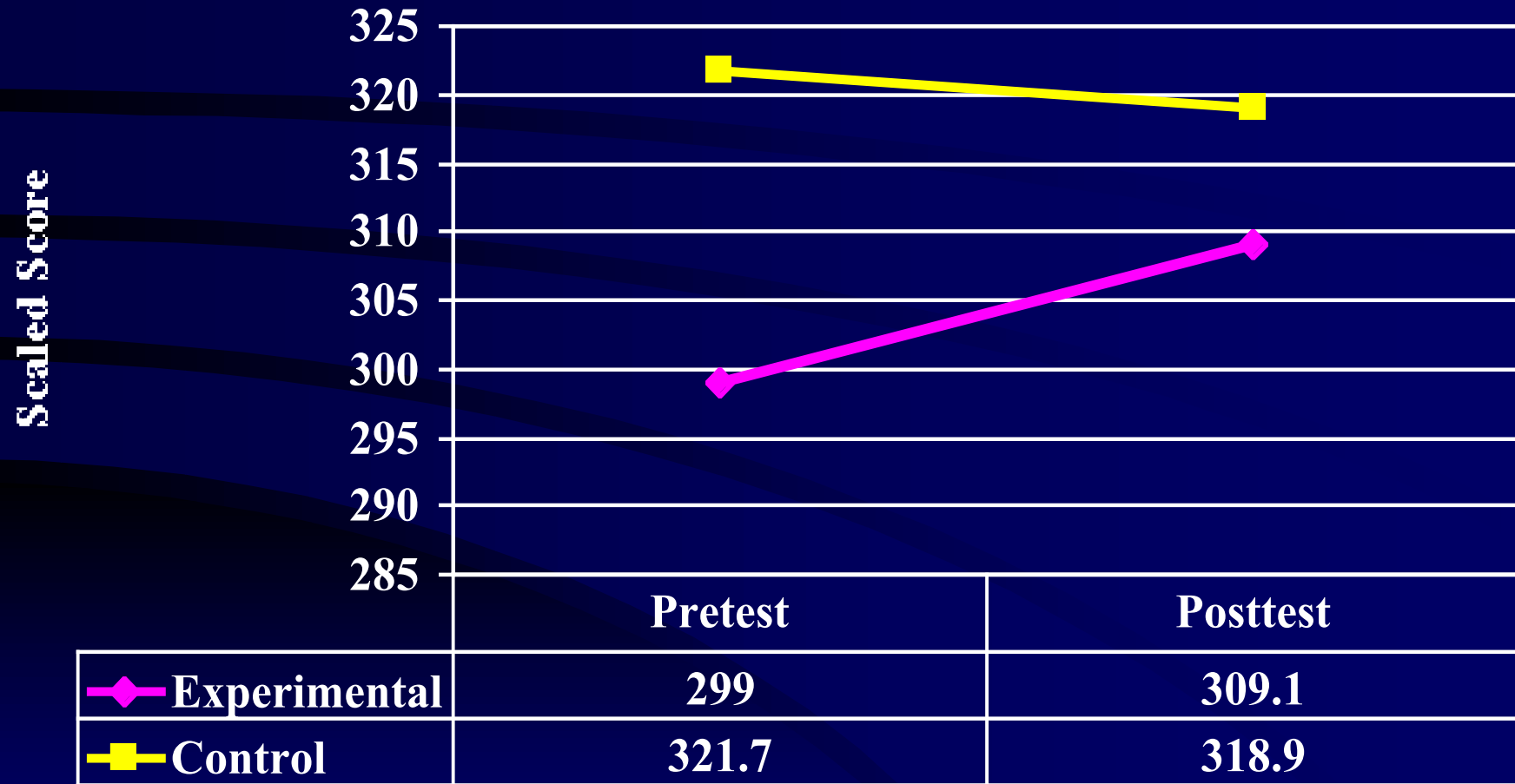
Race

Caucasian	59%
African-American	20%
Latino	13%
Other	8%

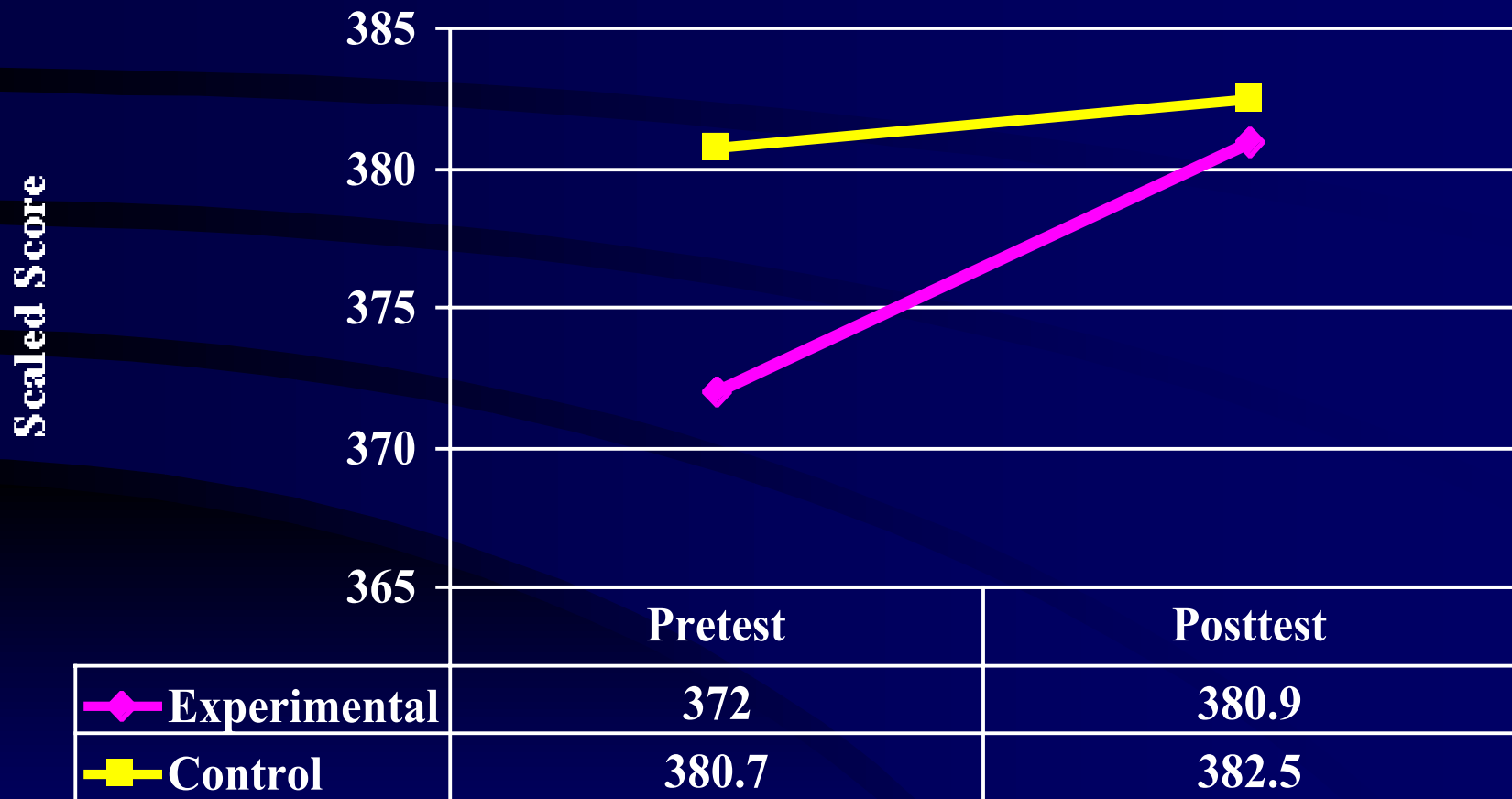
Grade

K	0.4%
1	38.8%
2	35.9%
3	25.8%

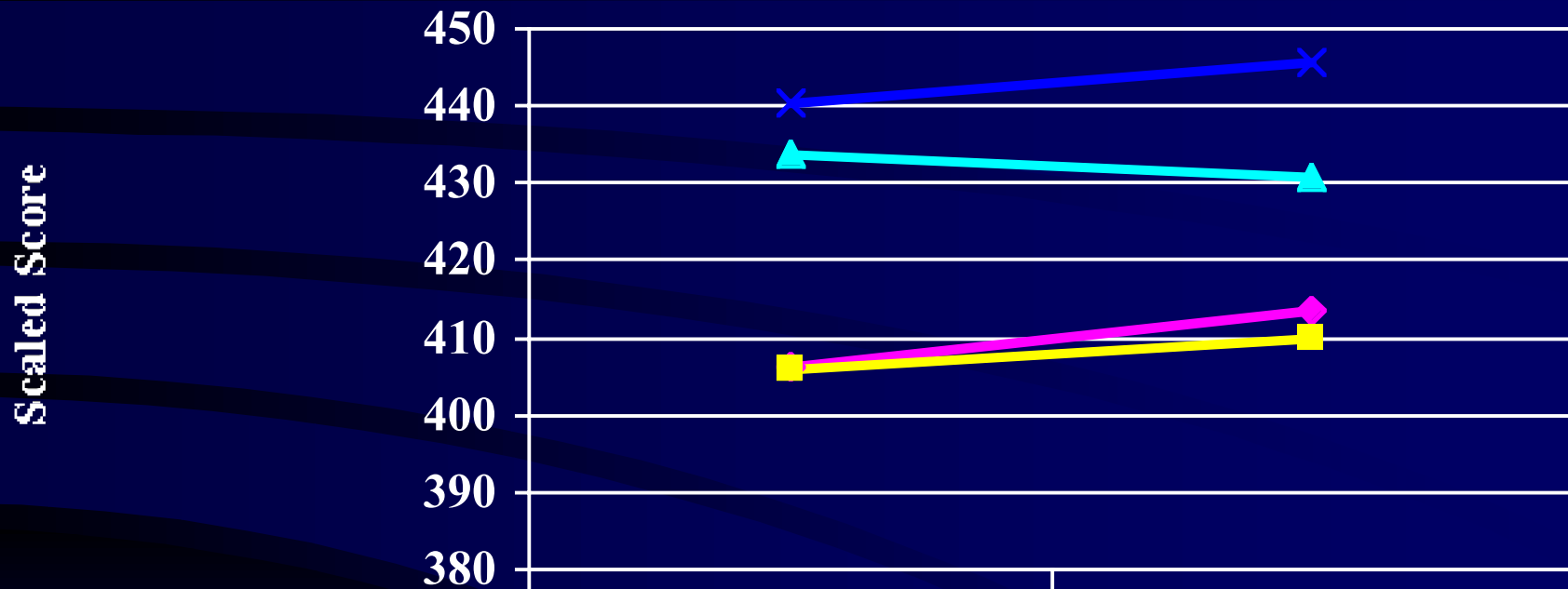
Level R Results



Level 1 Results



Levels 2 & 3 Results



	Pretest	Posttest
Experimental - 2	406.1	413.5
Control - 2	405.8	409.8
Experimental - 3	433.7	430.7
Control - 3	440.3	445.7

Year 2: Conclusions for Gates-MacGinitie Spring-Fall 99

- No gender differences at any level
- Greatest benefits for beginning or struggling readers, usually the youngest children
- No gains in standardized test scores for better/older readers

Oral Reading Measures

Benefits

- Aligned with daily instruction
- Multiple measures of fluency & comprehension collected simultaneously
- Diagnostic immediately

Liabilities

- Requires expertise to administer & interpret
- Requires multiple passages to assess reading level
- Accuracy (i.e., running records & miscues) are insufficient by themselves
- Teachers may “teach” commercial materials

Oral Reading Data: BRI

- Graded Word lists
- Reading Time & Words/Minute Read Correctly on Grade Level Passage
- Miscues & Self-Corrections
- Propositions & Key Ideas Recalled
- Comprehension Questions

BRI Results

- Reading Rate: Exp > Control
- Word lists : Exp > Control
- No differences on oral reading measures

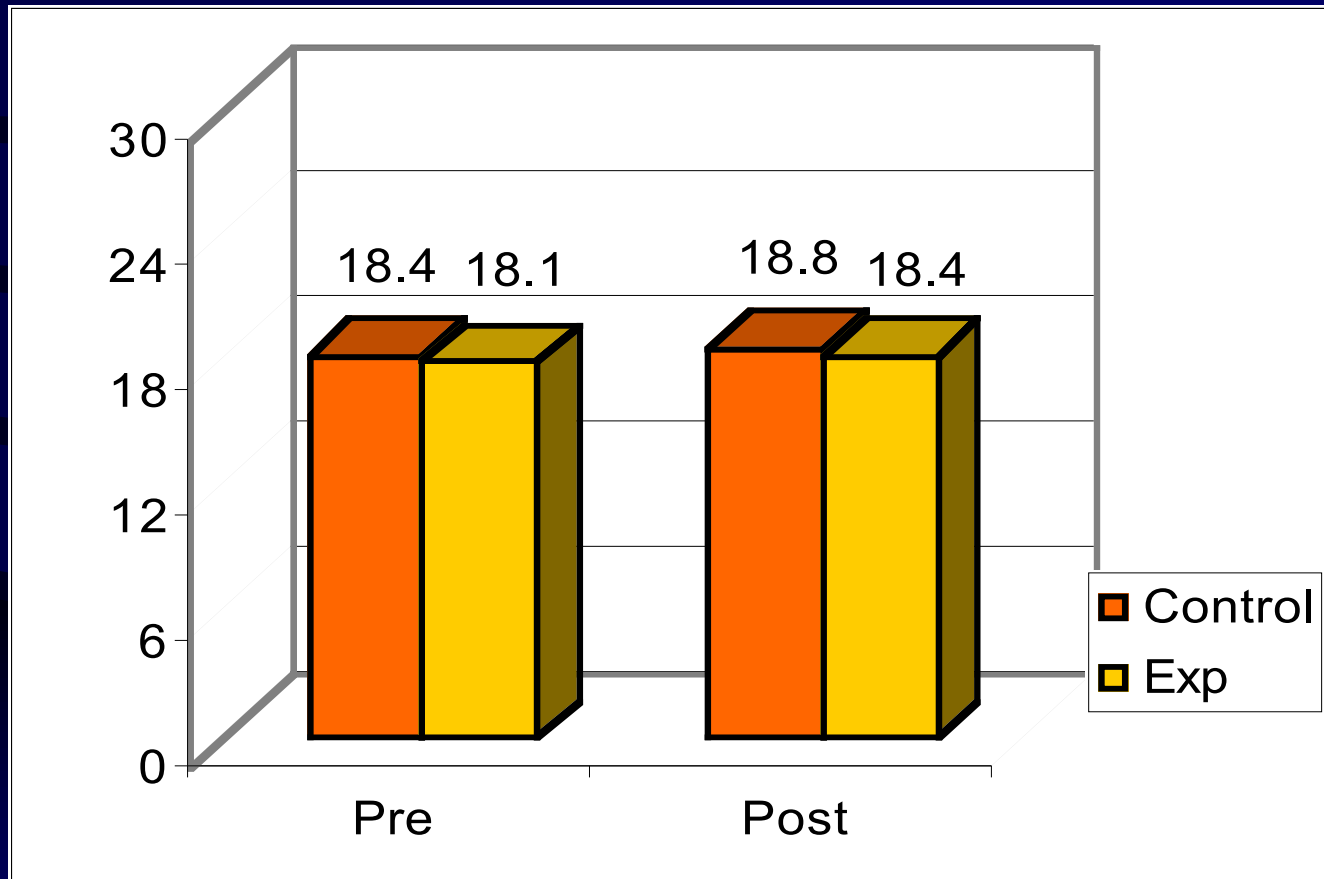
BUT

- BRI differences between Forms interacts with Groups
- Groups not equivalent at pre-test

Literacy Habits Items

- How often do you...
 - visit the library or Bookmobile?
 - write stories or poems at home just for fun?
 - read at home for fun?
 - read books or stories at bedtime?
 - your parents help you read or write at home
- Response options
Hardly ever About once a week Almost every day

Literacy Habits



K-1 SOAR Items

- Opinions (8 items)
 - I can say and understand many words.
 - I like to read books with other people
 - I try hard to do my best when I read
 - I really like to look at books even when I don't have to

- Response options

Not like me

A little like me

A lot like me

(smiley face scale)

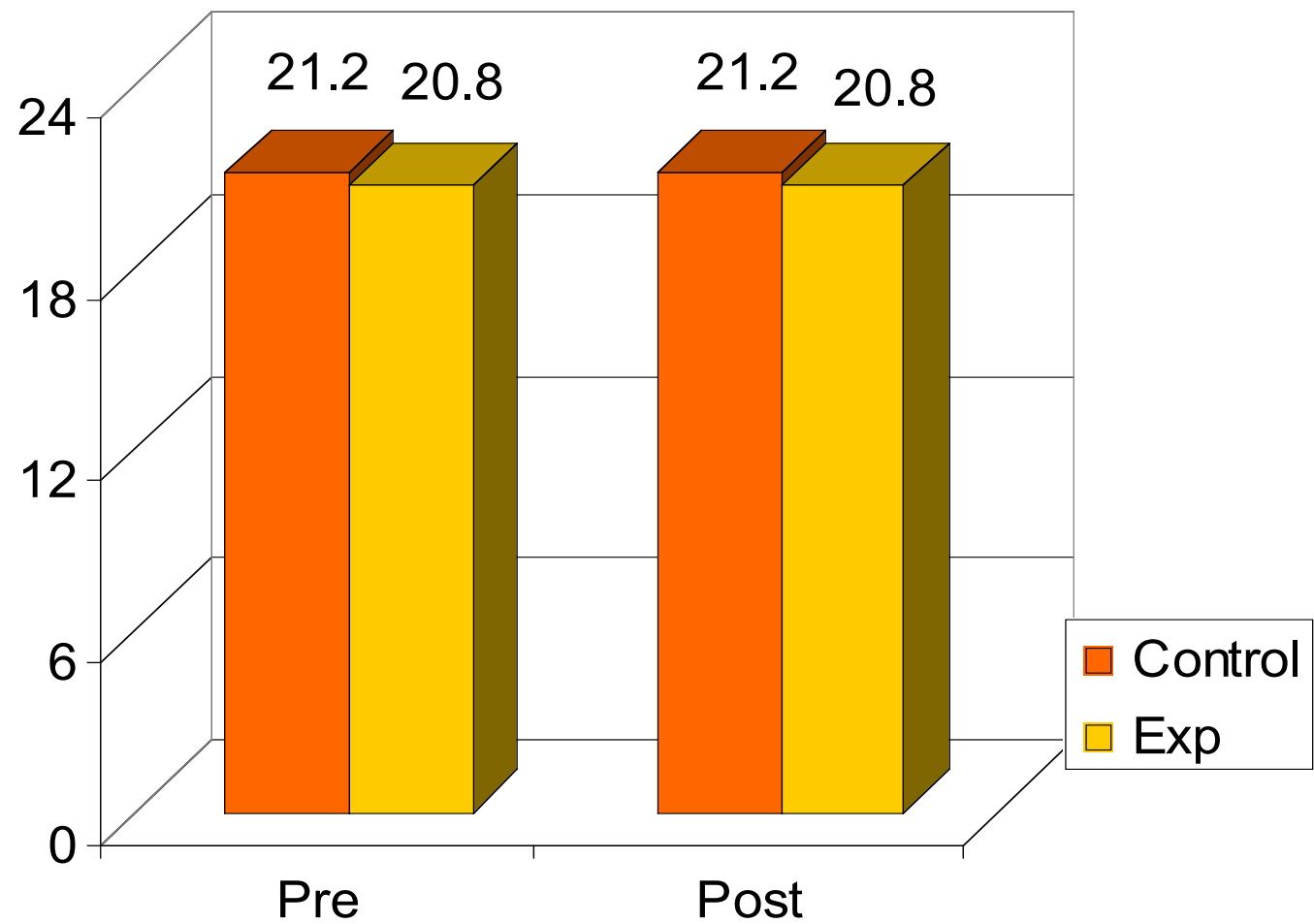
Grade 2-3 SOAR Items

- Opinions (16 items)
 - I can read out loud in class without making many mistakes.
 - I pay attention when I read in class.
 - I choose to read things that really make me think.
 - I think reading is fun.

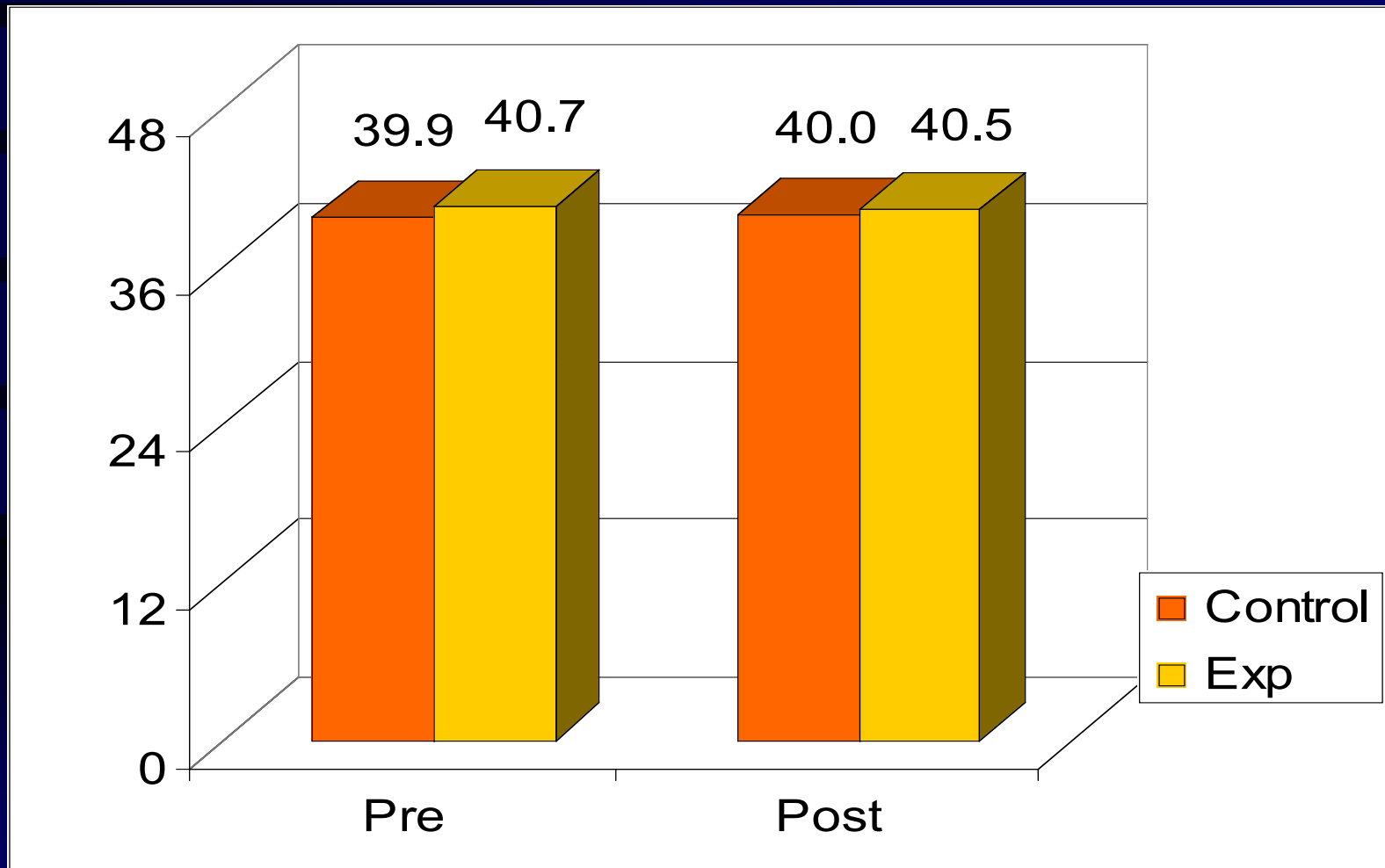
- Response Options

Not like me A little like me A lot like me

K-1 SOAR



Grades 2-3 SOAR



Parents' Survey

- 25 item survey mailed to parents
- Includes 15 items about reading habits with 4 response options about frequency of activities
- Includes 10 items about attitudes with 4 options for agree/disagree
- Exp N = 319; Control N = 120

Parents' Reports of Habits

- Parents of children in Exp and Control groups report equal habits except for 3rd grade Exp>Control
- Parents report MORE good reading habits among younger children
- Parents report MORE good reading habits among girls>boys

Parents' Reports of Habits

Parents reported Exp > Control for:

- Visit library
- Attend special activities at bookstores
- Go to bookstore
- Read books during summer
- Ask for help to read/write at home

Parents' Reports of Habits

Parents of 3rd graders reported Exp > Control for:

- Write letters to people
- Write stories/poems for fun
- Read newspapers/magazines

Parents' Reports of Attitudes

- Parents report MORE positive attitudes among younger children
- Parents report MORE positive attitudes among girls > boys
- No differences between Exp and Control groups

Conclusions

- Children's responses to Literacy Habits & SOAR did not differ by group
- Girls and younger children may have better habits and attitudes
- Treatment too brief?
- Measures too complicated?
- Parental reports more accurate?

Nonreaders: Phonemic Awareness Change Scores Show Exp > Control

<u>Tasks</u>	<u>Experimental</u> <u>N = 104</u>	<u>Control</u> <u>N = 106</u>
Blending	.5	.1
Rhyming	.1	.1
Segmenting	.5	.3
Total PA	1.0	.6

*Did Summer School Help
Children Read Better?*

Yes

No

Hard to tell

Evidence for No Advantage

- BRI Oral Reading
- Literacy Habits
- SOAR (Attitudes)

Evidence for Yes Advantage

- Gates-MacGinitie Reading tests
- Reading rate and Word identification
- Phonemic awareness tasks
- Case studies of teachers who elicited High versus Low gains from students on the Johns BRI

Hard To Tell Because of Design Factors

Control Groups

- Not random; teacher nominated
- Why did they not attend summer school?
- What did they do during summer?
- Higher Gates pretests than Exp Ss so not equivalent groups at pretest
- Need to equate groups or match Ss - feasible?
possible?

Hard To Tell Because of Design Factors

Treatment Groups

- Not random; recruited/enlisted Ss
- Hawthorne effects, positive & negative
- Diverse etiology of reading problems
- Treatments/curricula vary by school & teacher
- Large teacher differences and teacher x treatment interactions
- Cost factors > curricula factors
- Assessments do not match curricula & instruction at each site

Hard To Tell Because of Design Factors

Assessment Factors

- Did we have poor tasks?
- Were tasks ill-suited to curricula?
- Why did Gates-MacGinitie scores change?
- Why didn't Habits and Attitudes change?
- How do you reconcile changes on some measures and not others?

Conclusions About Designs

- Traditional Pre-Post x Exp-Control design is inadequate & impractical
- Need to examine the fit between assessment tasks and curriculum in each program
- Need to develop alternative evaluation designs

Practical Concerns & Models

Schools need practical evaluation methods:

- Limited budgets & time for evaluation
- Accountability quickly and in useful terms
- Small, nonrandom, idiosyncratic samples
- No comparable control Ss
- Idiosyncratic curricula unless test prep
- Variable teacher expertise
- Longitudinal improvement more important than comparative achievement

A Design To Assess Progress

Assess personal growth in:

- Oral reading fluency (rate, accuracy, prosody)
- Comprehension & recall
- Vocabulary & word identification

Using:

- Reading same passages & word lists pre & post
- Optional comparison group for:
 - Pre-post differences
 - Practice effects
 - Norms or standards

Year 3: Summer 2000

- 4 model sites with different demographics
- Build local capacity for self-assessment
- Integrate into local assessment & curriculum
- Menu of options for what info to collect from students, teachers, parents, administrators
- Handbook of forms and info + CD with video clips for professional development